THE AMERICAN HEALTH CARE SYSTEM: WHERE HAVE ALL THE DOLLARS GONE?

Marc J. Roberts*
Ted Bogue**

The failure of the American health care system to deliver services of high quality at a reasonable cost has prompted increasing concern. Congress, however, seems unable to reach agreement on national health insurance legislation. Messrs. Roberts and Bogue argue that this inaction is largely the result of new perceptions of what is wrong with the health care delivery system. Anxiety over rises in costs, utilization of services, and the proportion of GNP devoted to health is accompanied by doubts about the capacity of the health care system alone to improve significantly the health status of the population, given the newly recognized impact on health of other social and environmental factors.

The authors trace these problems to characteristics of financing, production, and consumption unique to the health sector, and to the uncritical use by health providers of therapeutic procedures and technology, the effectiveness of which has never been demonstrated. The final section suggests legislative strategies to re-direct medical care providers toward achieving greater health impact for each dollar spent, primarily by encouraging the development of a more competitive health care market.

Introduction

After decades of optimism, the society is coming to recognize fundamental structural inadequacies in the system by which health

*Professor of Political Economy and Health Policy and Academic Director, Executive Programs in Health Policy and Management, Harvard School of Public Health; A.B., 1964, M.A., 1967, Ph.D., 1969, Harvard University.

**Member of the Class of 1976 at Harvard Law School. The authors would like to acknowledge the assistance of numerous members of the faculty and staff of the Harvard School of Public Health, whose ideas, suggestions, and observations are to various degrees reflected in the arguments in this article. Particular thanks are due to Dr. Howard H. Hiatt, Dr. Robert Haggerty, Mr. Herbert Sherman, and Mr. Martin Pernick.

care is provided. In the past, as particular difficulties were recognized, the society turned haphazardly to various responses, ranging from incentive schemes and public payment, on the one hand, to appeals to private altruism and calls for professional ethics on the other. The result is a poorly designed, often counterproductive set of institutional arrangements that cannot satisfy legitimate public purposes unless subjected to systematic and in some ways fundamental reorganization. It is far from clear, however, given existing social customs and the current balance of political power, whether the existing arsenal of regulatory and organizational techniques can achieve an effective solution. All the most obvious policies seem either politically or administratively infeasible or promise to be largely ineffective. Given its limited capacity for analysis and innovation, the inaction of Congress in such a context is not inexplicable.

This article has four purposes. The first purpose is to provide historical perspective on how the definition of and responses to "the health care problem" have evolved since the beginning of the twentieth century. The second purpose is to argue that in the past several years the apparent lack of success of attempts at public intervention has resulted in a redefinition of the problem and a growing awareness of the obstacles to reform, which together help explain the current congressional impasse with regard to national health insurance. The third purpose is to demonstrate that underlying structural and economic characteristics of the health industry are largely responsible for the failure of government programs to improve the delivery system. The final purpose is to outline several congressional policy initiatives which might help to reverse the direction of the undesirable trends which now plague the health system.

In particular, in order to slow or reverse the recent rapid rise in medical care costs, the government must discourage overutilization of some services and encourage the more efficient provision of others. The former means only providing those services clearly worth the resources required to make them available. The latter will require either using less capital or cheaper labor to provide a given medical service or substituting less expensive capital for more expensive labor. Neither can be achieved until the medical care system is organized so that providers have the incentives, information, and legal freedom to adopt such cost control measures.

I. Evolving Definitions of "The Health Care Problem"

Although it does some violence to the richness of historical detail, one can usefully distinguish three definitions of the health care problem corresponding roughly to three chronological periods since the turn of the century.[1] The first phase, which extended until the end of World War II, was concerned with quality. The response was to create institutional arrangements, such as more advanced medical schools and professional licensing, that presumably insured high quality. The second phase, which extended into the early seventies, focused on access. Numerous efforts were made to extend the system on both the supply and demand sides. The third phase, which will be discussed in section II, is just now being defined. The dominant concern is cost and, to a lesser extent, effectiveness, although quality and access problems remain as well.

A. *The Flexner Period*

The defining document of the first period is the well-known Flexner Report of 1910.[2] The Council on Medical Education of the American Medical Association had requested the Carnegie Foundation to sponsor a study of medical education and make recommendations for its improvement. The report, authored by Abraham Flexner (neither a physician nor a scientist), urged the creation of what became the modern system of medical education in the United States. The model for the system was the medical school of Johns Hopkins University, then the only one in the

---

1 For a good general history of medical care in the United States see R. Stevens, American Medicine and the Public Interest (1971). For a telling view of earlier years see C. Rosenberg, The Cholera Years (1962).

2 A. Flexner, Medical Education in the United States and Canada (1910).

country which offered a program limited to those with undergraduate degrees.[3]

Flexner took as an article of faith the widespread belief that, through the development and application of scientific knowledge, the effective quality of medical care would be increased. Research was to be intimately linked to patient care. The curriculum he advocated combined strong training in basic science and extensive clinical involvement, an approach which necessitated elaborate laboratory facilities as well as connections between the medical school and appropriate hospitals. Given the costs of such facilities, Flexner vigorously urged that most extant medical schools be closed because they would never be able to raise the capital that "adequate" facilities would require.[4] Eliminating medical schools attracted much support in the higher status levels of the medical community, whose members had been saying for years that an oversupply of unqualified physicians was lowering the earnings and prestige of the profession as a whole.[5]

Medical education, however, was only one of many areas in which similar reforms were undertaken in this period. Others included the rise of specialized medical societies and the development of licensing procedures.[6]

Running through all these changes was the assumption that the consumers of medical care cannot make rational decisions about what kind of care they need or who is competent to provide it. Instead, only those with scientific training have enough expertise to judge the skill of practitioners, the adequacy of educational programs, or the needs of patients. The responsibility for monitoring and maintaining "quality" must therefore be given to the medical profession itself, which in effect comes to use the authority of the state to enforce its own standards.

Ironically, many current problems have their roots in the Flex-

---

3 *Id.* at 28.

4 At the time there were approximately 150 medical schools in the country; his ideal plan allowed for only 30. *Id.* at 146.

5 Markowitz & Rosner, *Doctors in Crisis: A Study of the Use of Medical Education Reform to Establish Modern Professional Elitism in Medicine*, 25 AM. Q. 83, 84 (1973).

6 R. STEVENS, *supra* note 1, at 149-71.

nerian reforms. What were once solutions have now become difficulties. First, despite the efforts to upgrade quality through professional licensure, today we have almost no legal restrictions on what tasks any doctor can perform or on what new procedures he can introduce, once he receives his initial license.[7] This is in sharp contrast to, for example, the controls on the introduction of new drugs,[8] or the elaborate, equipment-specific licensing required of commercial airline pilots.[9] The surprising lack of external controls is perhaps consonant with the widespread respect physicians command from the public. The same sense of consumer incompetence which made market processes seem unreliable as a guarantor of medical quality has also served to limit the willingness of laymen to try to regulate such quality through the public sector. Second, although limiting the number of medical schools and increasing the scientific content of the curriculum was designed to improve the quality of medical care, it also had the effect of limiting access and raising costs.[10] Third, a greater irony is that former problems have now become solutions. In an effort to lower costs and enhance access the society is now trying to expand the supply of what Flexner would call "unqualified" doctors. Now, however, they are called trained providers, paramedicals, physician assistants, physician extenders, nurse practitioners, etc.[11]

---

7 *See, e.g.*, Ch. 362, §§ 2-3, Mass. Acts (1975). However, reasonable nondiscriminatory restrictions by hospitals on medical staff privileges related to professional qualifications are permissible. Holder, *Restriction of Hospital Privileges*, in THE BEST OF LAW & MEDICINE '70-'73, at 191, 192 (1974). There have been proposals for relicensure of physicians, although most of them would not require procedure-specific examination. Mueller, *Continuing Assessment of Medical Performance: Proposals for a Recertification Examination Structure*, 284 NEW ENG. J. MED. 1379-80 (1971).

8 21 C.F.R. § 310 (1975).

9 14 C.F.R. § 61.65 (1975).

10 Clearly one effect of the drastic reduction in the number of graduates Flexner urged was, in the long run, a smaller number of doctors and hence reduced access to care. See note 4 *supra*. Such a restriction of supply, together with the increased costs of training that Flexner's model curriculum imposed, similarly tended to lead to price increases.

11 See A. SADLER, B. SADLER & A. BLISS, THE PHYSICIAN'S ASSISTANT — TODAY AND TOMORROW (1972); Kissam, *Physician's Assistant and Nurse Practitioner Laws: A Study of Health Law Reform*, 24 U. KAN. L. REV. 1 (1975) for surveys of the paramedical movement. See text accompanying notes 61-78 *infra*.

B. *The "Access" Period*

The second phase of public concern with the medical care system was the post-World War II period during which attention shifted from quality to access. In part because the Flexner reforms had been successful in restricting the supply of physicians a generation earlier, additional or larger medical schools now seemed required. Furthermore, if the high technology medicine produced by the Flexnerian schools was to be widely available, then the hospital, which was the appropriate setting for utilizing that technology, needed to be made more widely accessible. The Hill-Burton program of hospital construction and modernization grants[12] was initiated in 1946. The whole hospital system was to be manned by doctors trained in medical schools which Federal dollars from a series of programs[13] helped to support.

Simultaneously the prestige of science was further enhanced in medicine and elsewhere. The experience of World War II, of radar and the atomic bomb, reinforced the technological optimism characteristic of much of American public policy. Science would solve health problems, along with everything else.[14] Thus, in this era the National Institutes of Health, which supported medical

---

12 Hospital Survey and Construction Act, ch. 958, 60 Stat. 1040 (1946), *adding* Public Health Service Act §§ 601-35 (codified, *as amended*, at 42 U.S.C. §§ 291 to 291*o*-1 (1970)).

13 The government's active involvement with construction grants for medical schools and loans for medical students began in 1963. Health Professions Educational Assistance Act of 1963, Pub. L. No. 88-129, 77 Stat. 164. Broadening legislation was passed in 1965 (Health Professions Educational Assistance Amendments of 1965, Pub. L. No. 89-290, 79 Stat. 1052), 1968 (Health Manpower Act of 1968, Pub. L. No. 90-490, 82 Stat. 773), and 1971 (Comprehensive Health Manpower Training Act of 1971, Pub. L. No. 92-157, 85 Stat. 431). Formula grants for health professions schools began with the 1965 Amendments and were converted into their current form of capitation payments in the 1971 Act. All of the statutory provisions concerning federal support for medical schools and medical students are now contained in Parts A-F of Title VII of the Public Health Service Act §§ 701-86, 42 U.S.C. §§ 292 to 295g-23 (1970 & Supp. III, 1973).

For general descriptions and evaluations of federal financing of medical education see R. FEIN & G. WEBER, FINANCING MEDICAL EDUCATION (1971); H.R. REP. No. 266, 94th Cong., 1st Sess. 12-70 (1975); U. REINHARDT, PHYSICIAN PRODUCTIVITY AND THE DEMAND FOR HEALTH MANPOWER (1975); J. Winsten, Health Manpower — Public Policy Issues and Federal Legislation, Oct. 29, 1975 (unpublished paper on file with Harvard Journal on Legislation).

14 Price, *Money and Influence: The Links of Science to Public Policy*, 103 DAEDALUS 97, 99-100 (1974).

research and the rapidly growing university medical centers, underwent vast expansion.[15]

Toward the end of this second phase it became increasingly clear that merely expanding the supply of medical care was not necessarily going to guarantee broad access to it, despite the creation and expansion of health insurance schemes. Some people could not afford the increasing expense of high technology, hospital-based medicine delivered by scientifically trained doctors. The straightforward solution of the Kennedy-Johnson years was to purchase health care for those who otherwise would be deprived — the aged and the poor.[16] Relatively little thought was devoted, however, to whether or not the system could satisfy the expanded demands that the resulting Medicare[17] and Medicaid[18] programs made upon it. Instead, these new financing arrangements were launched with the self-confident view that "what society needs in the health area is more of the same."

II. New "Problems" and New Responses

The obvious final step in this historical development would have been comprehensive national health insurance. And for a period of time, some years ago, it seemed as if such a development would be inevitable "in the next few years." But those "next few years" have begun to pass, with such action becoming no more, and perhaps less likely. Changing economic conditions and political leadership have no doubt contributed to this change in mood, but new perceptions of what is wrong with the system of health care delivery are also important. In Congress and in the medical care sector itself, "more of the same" no longer seems a sufficient response. Instead, Congress is considering national health insurance in a context in which not only is quality too uneven and

---

15 For example, the budget of the National Institutes of Health has gone from $256,000,000 in 1960, Statistical Abstract of the United States: 1973, at 524 (1973), to $1,636,000,000 in 1975, Statistical Abstract of the United States: 1975, at 549 (1975).

16 T. Marmor, The Politics of Medicare *passim* (1970).

17 Social Security Act §§ 1801-75, 42 U.S.C. §§ 1395-95*ll* (1970).

18 *Id.* §§ 1901-08, 42 U.S.C. §§ 1396-96g.

access too difficult but, in addition, costs seem too high and "effectiveness"[19] too uncertain.

Since 1972, Congress has passed or considered several major pieces of health care legislation other than national health insurance which reflect a new formulation of the health care problem. Two distinguishable but related economic issues — cost and effectiveness or efficacy — have emerged. But there is also growing recognition of the political and institutional factors which constrain the competence of the federal government effectively to develop and implement reforms.

A. *The Current Problems*

1. Economic Problems

The most obvious and highly publicized reason for recent public and congressional anxiety is physician and especially hospital costs which have been inflating much faster than the overall economy.[20] The budgets for Medicare and Medicaid have increased faster than even their most cost-conscious critics were predicting only a few years ago.[21] This rise of overall health care costs has two distinct components: increasing costs per service (*i.e.*, efficiency) and increasing utilization of services.

Health cost inflation is a double-edged problem. It imposes an

---

19 See text accompanying notes 22-25 *infra*.

20 Feldstein, *The Rising Price of Physicians' Services*, 52 REV. ECON. & STAT. 121 (1970); Iglehart, *Health Report/Explosive Rise in Medical Costs Puts Government in Quandary*, 7 NAT'L J. REP. 319, 1320-21, 1325 (1975); N.Y. Times, Apr. 26, 1976, at 1, col. 5, 7. General Motors recently announced that it spends more for worker health benefits than it does for steel. Washington Post, Mar. 16, 1976, at A8, col. 1.

21 The Senate Finance Committee Report on the Social Security Amendments of 1972 included the following passage:

> According to recent estimates the costs of the medicare hospital insurance program will overrun the estimates made in 1967, by some $240 billion over a 25-year period. The monthly premium costs for part B of medicare — doctors' bills — rose from a total of $6 monthly per person on July 1, 1966, to $11.60 per person on July 1, 1972. Medicaid costs are also rising at precipitous rates.

S. REP. No. 1230, 92d Cong., 2d Sess. 254 (1972). More recent statistics reported by the Social Security Administration reveal no moderation in this trend. Fiscal year 1975 expenditures were 25% higher for Medicaid and 30% higher for Medicare than in fiscal 1974. Combined federal and state spending for the two programs in fiscal 1975 was $27.8 billion. N.Y. Times, Nov. 26, 1975, at 37, col. 4. *See* Iglehart, *supra* note 20, at 1319, 1324-26.

increasing financial burden on those who must pay privately for their health care; simultaneously, it increases congressional reluctance to expand government financing, which would relieve citizens' distress only eventually to increase their taxes or reduce other benefits.

In addition to the cost concerns, however, there is growing questioning of the effectiveness of medical care. Researchers are beginning to assess the impact of health care by looking at outcomes (how a particular service affects the patient's health status) rather than only examining the process of care (how it is delivered).[22] Evidence of the delivery of ineffective or unnecessary services has begun to accumulate. Such overutilization includes unnecessary surgery, over-hospitalization, and over-prescribed or misprescribed medication.[23] Moreover, using a criterion of empirical efficacy as measured by changes in health status rather than the norm of prevailing medical practice might result in a higher percentage of services being categorized as unnecessary. Perhaps the most provocative finding is that such overuse often seems to result from oversupply (*e.g.*, excessive surgery is associated with too many surgeons relative to population).[24]

The ineffectiveness of many health services seems even more striking in light of growing awareness of the relationship between personal health and events entirely outside the health care "system." Several recent popular works have focused on the critical role in determining life expectancy of environmental and behavioral factors — from murder, suicide, and auto accidents to the role of smoking, diet and air quality in causing heart disease and cancer.[25]

It is only partially correct to say that the cost aspect of current concerns is the older problem of access from a slightly different

---

22 *E.g.*, Brook & Appel, *Quality-of-Care Assessment: Choosing a Method for Peer Review*, 288 NEW ENG. J. MED. 1323 (1973); Rutstein, Berenberg, Chalmers, Child, Fishman & Perrin, *Measuring the Quality of Medical Care: A Clinical Method*, 291 NEW ENG. J. MED. 582 (1976).

23 Hiatt, *Protecting the Medical Commons: Who Is Responsible?*, 293 NEW ENG. J. MED. 235, 236-38 (1975) and sources cited therein.

24 Blacksone, *Misallocation of Medical Resources: The Problem of Excessive Surgery*, 22 PUB. POLICY 329 (1974).

25 *E.g.*, V. FUCHS, WHO SHALL LIVE? (1974); Iglehart, *supra* note 20, at 1327.

perspective — or that "effectiveness" is merely "quality" redefined. Instead, the linkage between the two introduces the novel possibility that some apparently desirable but only marginally effective techniques and procedures may not be worth the cost of making them widely accessible. In other words, the problem is being viewed as one of resource allocation.

In fiscal year 1975, according to the Social Security Administration, the U.S. spent $118 billion, or 8.3% of its Gross National Product, on medical care. This level of spending represents a dramatic increase after three years of the economic stabilization program, during which the percentage was stable at about 7.8%.[26] Health experts are paying more attention to the fact that, given scarce resources, every additional dollar spent on health care is a dollar which cannot be spent on housing, education, transportation, environmental control, etc. The belief is spreading that perhaps the country should begin to shift resources now invested in health care to production of alternative goods and services which yield a higher social "return."[27] Many contend that if a well-informed consumer of average income were required to pay the full cost of each health service delivered to him at the time of delivery, less economic resources would be devoted to health care than is now the case.[28]

2. Political and Implementation Problems

The response that the "access" era had demanded from Congress — providing funds for both producers and consumers — was simple, popular, and relatively easy to implement. Varying amounts of money were dispensed to different groups at different times (hospitals, medical schools, researchers, patients).[29] Health

---

26 Iglehart, *supra* note 20, at 1319, 1320; N.Y. Times, *supra* note 21.

27 Iglehart, *supra* note 20, at 1319, 1327; N.Y. Times, *supra* note 21. See Havighurst & Blumstein, *Coping with Quality/Cost Trade-offs in Medical Care: The Role of PSROs*, 70 Nw. U.L. Rev. 6-20 (1975) for an analysis of the allocation problem.

28 Havighurst & Blumstein, *supra* note 27, at 19-20. See text accompanying notes 80-85 *infra*.

29 See text accompanying notes 11-18 *supra*.

was an issue with wide appeal since in most cases both the narrow professional and industry interest groups and the more diffuse, issue-oriented "public" constituency favored the production of more services. Pork barrel legislation was apparently in the public interest. As a congressman, one could "sell out" and be virtuous simultaneously.

Now, however, Congress is presented with a different set of problems. If the society is to be able to afford broad publicly financed access to health care, costs must be diminished or at least contained by both providing individual services in a less costly manner and insuring that only effective and appropriate services are in fact utilized. But to accomplish such ends, Congress must overcome several obstacles. First, it must act in the face of conflicting objectives and perceptions — which makes it difficult to organize the broad legislative coalition needed to pass major programs.[30] Second, complex regulatory and structural initiatives must be reduced to statutory language, which strains limited staff resources in such a highly technical area.[31] Congress could delegate details to the Department of HEW, but such broad statutes would entail an increase in HEW's decision-making power over the health system vis-à-vis Congress. Third, because of the complexity of the problems, even the experts are divided or uncertain about the optimal solutions.[32] Finally, such actions are likely to offend powerful interests which have both large stakes in the outcome and the technical skills and organizational resources to have a major impact on the legislative or regulatory process.[33]

---

30 *See, e.g.*, Wilson, *The Politics of Regulation*, in SOCIAL RESPONSIBILITY AND THE BUSINESS PREDICAMENT (J. McKie ed. 1974).

31 As a result of such limited staff resources, special interests often have a major impact on legislation. *See* E. REDMAN, THE DANCE OF LEGISLATION (1973); R. BAUER, I. POOL & L. DEXTER, AMERICAN BUSINESS AND PUBLIC POLICY (1963).

32 Consider the differing views on hospital rate regulation offered by Cohen, *State Rate Regulation*, in INSTITUTE OF MEDICINE, NAT'L ACADEMY OF SCIENCES, CONTROLS ON HEALTH CARE 123 (1975) [hereinafter cited as CONTROLS ON HEALTH CARE]; Noll, *The Consequences of Public Utility Regulation of Hospitals*, in CONTROLS ON HEALTH CARE 25 (1975); A. SOMERS, HOSPITAL REGULATIONS: THE DILEMMA OF PUBLIC POLICY (1969). Similarly, for a survey of differing perspectives on national health insurance see K. DAVIS, NATIONAL HEALTH INSURANCE: BENEFITS, COSTS AND CONSEQUENCES (1975).

33 See note 31 *supra*.

B. *Responses*

It is not immediately obvious why health care cost-effectiveness should be a "problem" requiring positive public action. Many products and services are more costly and not as effective as we would like. But, if a manufacturer places on the market a costly and ineffective (badly designed) forty foot cruising sailboat, no one calls for public action to lower the cost (so that "everyone" can afford one) or to limit the future activities of the designer. Public policy toward health care, in contrast, reflects a widely shared ethical assumption that somehow it is not just another good or service whose distribution is to be determined by ordinary market principles. Instead, it is widely believed that in a good society access to a certain "minimum" amount of "adequate" medical services ought to be guaranteed by the government.[34]

Accordingly, the response to current problems has been to intervene ever more extensively in the structure of the health care system. Four federal actions in recent years illustrate this trend. Each of these programs represents a legislative response to the cost effectiveness problem (though some are motivated by considerations of quality and access as well). However, each has encountered some or all of the problems of formulation and implementation discussed above and therefore has experienced limited success at best.

1. Health Maintenance Organizations

Health Maintenance Organizations (HMO's)[35] differ from the conventional fee-for-service delivery system in several ways. They derive their revenue from fixed prepaid fees collected from members. They usually provide a comprehensive range of services (*i.e.*, inpatient and outpatient, hospital and physician), own or contract with the facilities in which they deliver care, and employ salaried physicians. Most are at least formally (if not in practice) controlled

---

34 Fried, *Rights and Health Care — Beyond Equity and Efficiency*, 293 NEW ENG. J. MED. 241 (1975).

35 For background and history concerning HMO's see Rosoff, *Phase Two of the Federal HMO Development Program: New Directions After a Shaky Start*, 1 AM. J.L. & MED. 209, 210-12 (1975); Schneider & Stern, *Health Maintenance Organizations and the Poor: Problems and Prospects*, 70 NW. U.L. REV. 90 (1975).

by consumer members rather than medical and hospital professionals. The chief economic advantage of HMO's is that they have no financial incentive to provide excessive treatment since their revenues do not vary with the volume of services they provide, *i.e.*, the provider is at risk for the costs of added treatment. Instead, they are encouraged to contain costs by eliminating unnecessary services and delivering others more efficiently.

The ostensible purpose of the Health Maintenance Organization Act of 1973[36] was to encourage the growth of a lower cost form of care. Yet, in an effort to encourage "quality" and "access," federal aid was made available only if HMO's offered a wide range of benefits,[37] and, within certain limits, accepted high risk subscribers[38] at normal rates[39] — provisions which have raised HMO costs and limited their ability to compete with health insurance plans, thus defeating the legislative purpose.[40]

2. Professional Standards Review Organizations

Professional Standards Review Organizations (PSRO's), as provided for under the Social Security Amendments of 1972,[41] are regional organizations of physicians designed to review the medical necessity and quality of institutional care provided to Medicare and Medicaid patients in their area.[42] The physicians themselves are to develop and enforce objective standards to assure that Medicare and Medicaid do not reimburse providers for services which are either unnecessary or of substandard quality.[43] The medical necessity of hospital admissions is to be reviewed just

---

36 Pub. L. No. 93-222, 87 Stat. 914, *adding* Public Health Service Act §§ 1301-15 (codified at 42 U.S.C. §§ 300e to 300e-14 (Supp. III, 1973)).

37 Public Health Service Act §§ 1301(b)(1), 1302(1), 42 U.S.C. §§ 300e(b)(1), 300e-1 (1) (Supp. III, 1973).

38 *Id.* § 1301(c)(3)-(5), 42 U.S.C. § 300e(c)(3)-(5).

39 *Id.* §§ 1301(b)(1)(B)-(C), 1302(8), 42 U.S.C. §§ 300e(b)(1)(B)-(C), 300e-1(8).

40 INSTITUTE OF MEDICINE, NAT'L ACADEMY OF SCIENCES, HMOs: TOWARD A FAIR MARKET TEST 44-45 (1974). These problems were not unforeseen at the time the legislation was being written. *Hearings on S. 703, S. 837, S. 935, S. 1182, S. 1301 Before the Subcomm. on Health of the Senate Comm. on Labor and Public Welfare*, 92d Cong., 1st & 2d Sess., pt. 6, at 2616 (1972). See text accompanying notes 127-30 *infra*.

41 § 249F, Social Security Act §§ 1151-70, 42 U.S.C. §§ 1320c to 1320c-19 (Supp. III, 1973).

42 Social Security Act § 1152, 42 U.S.C. § 1320c-1 (Supp. III, 1973).

43 *Id.* §§ 1155(a), 1156, 1158, 42 U.S.C. §§ 1320c-4(a), 1320c-5, 1320c-7.

after the patient enters the hospital. Stays beyond a pre-determined number of days are to be reviewed before the end of the initial period. If the PSRO disapproves, the patient is notified that the government will not pay for further care and has the choice of leaving the hospital or making private payment. Quality is to be reviewed retrospectively against pre-selected standards. When unjustifiable deviations from the standards appear, the PSRO is supposed to take remedial steps to correct the problem.[44]

Conflicting objectives and interests have interfered with the effective implementation of PSRO's. Congressional advocates of this legislation saw it primarily as a cost control measure to prevent overutilization of medical procedures and hospital beds.[45] Yet if PSRO's really do review and improve "quality," they could raise costs rather than lower them by enforcing adherence to standards requiring greater use of more sophisticated facilities and ancillary services — no doubt much to the dismay of some early backers.[46]

Also, it seems clear that some group outside the hospital would be best able to monitor critically and objectively the institution's performance. However, at the urging of the American Hospital Association, the Senate Finance Committee amended the PSRO bill to allow hospitals the option of taking over their own PSRO activities.[47]

Control of PSRO's was delegated to the physicians themselves because they are viewed as the skilled practitioners of a difficult but important craft whose skills cannot be fairly judged by those outside the guild. However, the very processes which have been used to try to guarantee quality — particularly the development of medical schools that are simultaneously major centers for scientific research — raise doubts about the wisdom of physician

44 For a detailed description of the PSRO hospital review system, see Goran, Roberts, Kellogg, Fielding & Jessee, *The PSRO Hospital Review System*, 13 MED. CARE, Apr. 1975 Supplement.

45 *See* S. REP., *supra* note 21, at 254 (1972).

46 Havighurst & Bovbjerg, *Professional Standards Review Organizations and Health Maintenance Organizations: Are They Compatible?*, 1975 UTAH L. REV. 381, 394-98, 401-06 (1975); note 48 *infra*.

47 Social Security Act § 1155(e)(1), 42 U.S.C. § 1320c-4(e)(1) (Supp. III, 1973); Turner, *Health Report/HEW Begins Medical Review; AMA, Hospitals Mount Opposition*, 6 NAT'L J. REP. 90, 101 (1974).

self-regulation. PSRO standards of care are being developed for each medical specialty based largely on extant patterns of practice, despite the lack of research in such medical centers as to the empirically measurable health impact of the procedures required.[48] The profession has every reason to support the view that the existing orientation of medical research produces important and valuable medical progress. Thus, PSRO's, given their secure location within the medical community, are not likely to raise questions about the efficiency and effectiveness of new methods and technological advances.

3. Health Planning

Since the early 1960's the federal government has directly supported institutional arrangements designed to provide health planning at the local level.[49] The principal purpose of these or-

---

48 The PSRO program has made some verbal effort to encourage the use of outcome-validated process criteria. Bureau of Quality Assurance, Health Services Ad., Dep't of HEW, PSRO Transmittal no. 17, at 4 (of Attachment 2), Feb. 18, 1975. However, an expert on assessing quality of care has warned that if "outcome" criteria are developed subjectively by physicians, who tend greatly to overestimate the effectiveness of their care, rather than through empirical research, the result may be the addition of more process criteria. Thus, the delivery of even more services of questionable effectiveness will be encouraged in an effort to achieve artificially inflated outcome expectations. R. BROOK, QUALITY OF CARE ASSESSMENT 56-57 (1973); Brook & Appel, *supra* note 22, at 1328.

49 Local organizations had been involved in health facilities planning in some areas since the 1920's. For useful general histories of health planning in the U.S. see Gottlieb, *A Brief History of Health Planning in the United States*, in REGULATING HEALTH FACILITIES CONSTRUCTION 7 (C. Havighurst ed. 1974); J. MAY, HEALTH PLANNING: ITS PAST AND POTENTIAL (1967).

Statewide plans for health facilities were first required from the states as a condition for hospital construction financing under the Hospital Survey and Construction Act, ch. 958, § 2, 60 Stat. 1043-44, *adding* Public Health Service Act § 623 (*as amended,* Public Health Service Act § 604) (codified, *as amended,* at 42 U.S.C. § 291d (1970)) (popularly known as the Hill-Burton Act).

In 1965 the so-called "Regional Medical Programs" (RMP's) were established as the planning and coordinating component of a federal program aimed at combatting heart disease, cancer, and stroke. Heart Disease, Cancer, and Stroke Amendments of 1965, Pub. L. No. 89-239, § 2, 79 Stat. 927, *adding* Public Health Service Act §§ 903-905 (codified, *as amended,* at 42 U.S.C. §§ 299c-99e (1970)).

The same Congress in 1966 created the Comprehensive Health Planning (CHP) program, which provided federal grants to support both state and areawide (local) agencies designated to perform comprehensive health planning. Comprehensive Health Planning and Public Health Services Amendments of 1966, Pub. L. No. 89-749, § 3, 80 Stat. 1181, *amending* Public Health Service Act § 314 (codified, *as amended,* at 42 U.S.C. § 246 (1970)). The CHP program was modified to provide ganizations has been to improve the efficiency and lower the price of health services by preventing the construction of unneeded or duplicative facilities and the acquisition of unnecessary equipment.

The National Health Planning and Resources Development Act of 1974[50] will create a nationwide system of regional Health Systems Agencies.[51] These HSA's are not only required to formulate plans[52] but also will play a key role in approving all new hospital construction programs, through their participation in "certificate of need" procedures.[53] They also will control the distribution of Federal health funds in their regions.[54] However, Congress was so unsure of how to organize this activity that it left many critical questions incompletely resolved — including the relationship HSA's were to have to general purpose local governments. The law allows for various organizational forms[55] and reserves for HEW the power to choose among applicants in accordance with statutory criteria.[56] This ambiguity has re-

---

for more local government input in 1967. Partnership for Health Amendments of 1967, Pub. L. No. 90-174, § 2(b)(2), 81 Stat. 533, *amending* Public Health Service Act § 314(b)(1)(A) (codified at 42 U.S.C. § 246(b)(1)(A) (1970)). The requirement that a majority of the members of each areawide health planning council be consumers and that RMP's be represented on such councils was enacted in 1970. Public Health Service Act Amendments of 1970, Pub. L. No. 91-515, § 230(5), 84 Stat. 1305, *adding* Public Health Service Act § 314(b)(2)(A) (codified at 42 U.S.C. § 246 (b)(2)(A) (1970)). Another title of the same Act broadened the scope of the Regional Medical Programs. Heart Disease, Cancer, Stroke, and Kidney Disease Amendments of 1970, Pub. L. No. 91-515, §§ 105-07, 84 Stat. 1299-1300, *amending* Public Health Service Act §§ 903-05 codified at 42 U.S.C. §§ 299c-99c (1970)). The federal statutory scheme for health planning is outlined in D. Neuhauser & F. Wilson, Health Services in the United States: An Introduction 52-63, Apr. 3, 1973 (available from Harvard School of Public Health).

50 Public Health Service Act §§ 1501-1640, 42 U.S.C.A. §§ 300k-t (Supp. 1976). The Act essentially integrates the three existing federal health planning programs — the Comprehensive Health Planning, Regional Medical, and Hill-Burton programs — into one new program.

51 *Id.* § 1512, 42 U.S.C.A. § 300*l*-1.

52 *Id.* § 1513(b), 42 U.S.C.A. § 300*l*-2(b).

53 *Id.* §§ 1513(f)-(g), 1523(a)(4)-(6), 42 U.S.C.A. §§ 300*l*-2(f)-(g), 300m-2(a)(4)-(6). Under the Act, state health planning and development agencies are responsible for operating a state and in some cases federal certificate of need program which is designed to prevent all health facility capital expenditures except those for which the state agency has certified the need. The agency must consider any recommendations submitted by HSA's on particular proposals.

54 *Id.* § 1513(e), 42 U.S.C.A. § 300*l*-2(e).

55 *Id.* § 1512(b), 42 U.S.C.A. § 300*l*-1(b).

56 The Department of HEW must find that the applicant is capable of performing the statutorily-required planning functions, must consult with the Governor of

sulted in a major political battle over the process of writing the regulations implementing the law.[57]

At the same time, the limited available evidence does not conclusively demonstrate that state certificate of need programs, much less health planning without sanctions, have effectively reduced hospital construction or hospital costs.[58] Thus it is far from clear that the creation and strengthening of a heterogeneous series of local planning and review processes will accomplish congressional goals.

4. Health Manpower

Although there had been substantial federal funding for medical schools in the sixties,[59] by the early seventies there was widespread concern over the "doctor shortage," particularly in inner cities and rural areas. Congressional response to the manpower issue is an apt illustration of changing or even inconsistent perceptions and objectives in the health field. The initial theory was the medical equivalent of the "trickle-down" approach to housing — train enough new doctors and some will be forced out of the major cities by market pressure and come to practice in underserved areas. But this policy failed to recognize the ability of doctors to raise prices and simultaneously ensure that demand increases enough to support even rapidly growing supplies.[60]

Taking another tack, Congress for several years considered and implemented various programs, including student loan forgiveness and the deployment of Public Health Service doctors, to

---

the state, and must give priority to existing local CHP and RMP agencies. *Id.* § 1515, 42 U.S.C.A. § 300*l*-4.

57 Iglehart, *Health Report/State, County Governments Win Key Roles in New Program,* 7 NAT'L J. REP. 1533 (1975).

58 One recent statistical study concludes that certificate of need programs have had some impact in keeping bed levels below what one would expect in their absence but have also led to a largely offsetting increase in assets invested per bed, Salkever & Bice, *The Impact of Certificate-of-Need Controls on Hospital Investment,* MILBANK MEMORIAL FUND Q. (forthcoming). A somewhat more optimistic view of the Massachusetts experience, but without sophisticated statistical analysis, is given by Bicknell & Walsh, *Critical Experiences in Organizing and Administering a State Certificate of Need Program,* 91 PUBLIC HEALTH REPORTS, Jan.-Feb. 1976, at 29.

59 See note 13 and accompanying text *supra.*

60 See text accompanying notes 24 *supra* and 103-06 *infra.*

induce medical schools and medical students to practice in underserved areas,[61] but the problem persisted. The Comprehensive Health Manpower Training Act of 1971, besides increasing the amount of loan forgiveness, specifically provides financial support for the training of physician assistants.[62] Physician assistants and other non-physician practitioners were favored by the Nixon administration not only because it was believed they would increase the availability of medical services in shortage areas but also because they were far less costly to train and compensate than physicians and could handle many of the routine tasks which only physicians previously performed.[63]

Two bills currently pending in Congress[64] constitute the most

---

61 In 1965 a health professions student loan program, which provided forgiveness for physicians, dentists, and optometrists who practiced in shortage areas, at the rate of 10 percent per year of practice in such areas up to a 50 percent maximum, was enacted. Health Professions Educational Assistance Amendments of 1965, Pub. L. No. 89-290, § 4(b)(2), 79 Stat. 1057, *amending* Public Health Service Act § 741(f) (codified, *as amended*, at 42 U.S.C. § 294a(f) (Supp. III, 1973)). The program was amended and made more generous in 1966 (Allied Health Professions Personnel Training Act of 1966, Pub. L. No. 89-751, § 4(a), 80 Stat. 1230, *amending* Public Health Service Act § 741(f) (codified, *as amended*, at 42 U.S.C. § 294a(f) (Supp. III, 1973))) and 1971 (Comprehensive Health Manpower Training Act of 1971, Pub. L. No. 92-157, § 105(b)(1), 85 Stat. 449-50, *amending* Public Health Service Act § 741(f) (codified, *as amended*, at 42 U.S.C. § 294a(f) (Supp. III, 1973))). The effectiveness of the program, however, is open to doubt, since few chose to participate, and many of those who did might have practiced in underserved areas anyway. J. Winsten, *supra* note 13, at 49-51.

In 1970 Congress established the forerunner of the National Health Service Corps, which compensates and provides other support to physicians who practice in underserved areas as employees of the Public Health Service. Emergency Health Personnel Act of 1970, Pub. L. No. 91-623, §§ 1-2, 84 Stat. 1868-70, *adding* Public Health Service Act § 329 (codified, *as amended*, at 42 U.S.C. § 254b (Supp. III, 1973)). The program was expanded in 1972 (Emergency Health Personnel Act Amendments of 1972, Pub. L. No. 92-585, §§ 1-2, 86 Stat. 1290-92, *amending* Public Health Service Act § 329 (codified, *as amended*, at 42 U.S.C. § 254b (Supp. III, 1973))) and again in 1975 (Public Health Service Act Amendments of 1975, Pub. L. No. 94-63, §§ 801-03, 89 Stat. 353-54, *amending* Public Health Service Act § 329 (codified at 42 U.S.C.A. § 254b (Supp. 1976))). The NHSC program offers scholarships, with one year of obligated service required for each year of support. Public Health Service Act § 225, 42 U.S.C. § 234 (Supp. III, 1973). A similar scholarship program for non-NHSC doctors gives priority to students who are from low-income families and residents of the shortage area in which they agree to serve. Public Health Service Act §§ 784-86, 42 U.S.C. §§ 295g-21 to 295g-23 (Supp. III, 1973). The NHSC program is discussed by J. Winsten, *supra* note 13, at 51-55.

62 Public Health Service Act §§ 772(a)(3), 774(a)(1)(C), 42 U.S.C. §§ 295f-2(a)(3), 295f-4(a)(1)(C) (Supp. III, 1973).

63 A. SADLER, B. SADLER, A. BLISS, *supra* note 11, at 10-16.

64 H.R. 5546, 94th Cong., 1st Sess. (1975); S. 3239, 94th Cong., 2d Sess. (1976).

recent attempts to solve the health manpower problem. Once again, however, the issues have been redefined. The "doctor shortage" seems to have gone the way of the "missle gap." The focus has shifted to maldistribution by specialty and by geographic area; health experts now believe that if physicians were more appropriately located and more were trained for primary care, no areas would have to be underserved.[65] Both bills provide for some degree of federal control over the number of residencies for each specialty[66] (these provisions of H.R. 5546, which were intended to delegate the designation of residencies to a private group, were, however, amended out before passage by the House),[67] but they take different approaches to encouraging medical school graduates to locate in underserved areas.[68]

The manpower issue of the seventies, however, is the foreign medical graduate (FMG),[69] whose numbers have grown to the point that substantially more FMG's are admitted to the U.S. each year than new physicians are graduated from American medical schools.[70] The concern of Congress and the medical profession is supposedly the poor quality medicine practiced by FMG's.[71] Because FMG's are attracted to the U.S. by the opportunities for

---

H.R. 5546 has passed the House. 121 CONG. REC. H6661 (daily ed. July 11, 1975). The Senate Labor and Public Welfare Committee, after substituting the provisions of S. 3239 for the provisions passed by the House, ordered H.R. 5546 favorably reported to the floor on April 7, 1976. 122 CONG. REC. D483 (daily ed. Apr. 7, 1976).

65 *See* H.R. REP., *supra* note 13, at 38-43 (1975).

66 H.R. 5546, 94th Cong., 1st Sess. § 801 (1975) (which would have added § 1703 (a)-(b) to the Public Health Service Act); S. 3239, 94th Cong., 2d Sess. §§ 501, 802 (1976) (which would add §§ 798, 799(a), 771(b)(2), (7) to the Public Health Service Act).

67 121 CONG. REC. H6655-59 (daily ed. July 11, 1975).

68 H.R. 5546 would require any medical school graduate to pay back to the government an amount equal to the total capitation payments the medical school received on his behalf, unless he agrees to practice in an underserved area. H.R. 5546, 94th Cong., 1st Sess. § 502 (1975) (which would amend Public Health Service Act § 771(a)(3)). S. 3239 would require each medical school, as a condition of eligibility for capitation payments, to establish mechanisms to reserve a required percentage of places for students who are receiving National Health Service Corps scholarships, which obligate them to serve in a shortage area. The required percentages for 1978, 1979, and 1980 are, respectively, 25%, 30%, and 35%. S. 3239, 94th Cong., 2d Sess. § 802 (1976) (which would add §§ 771(b)(1), (4)-(6) to the Public Health Service Act).

69 Mick, *The Foreign Medical Graduate*, 232 SCI. AM., Feb. 1975, at 14.

70 H.R. REP., *supra* note 13, at 46-47 (1975).

71 *Id.* at 53-54.

post-graduate medical education, both bills would limit the total number of residencies in the U.S. to slightly more than the number of graduates of U.S. medical schools in a given year.[72]

Both bills also authorize money for the training of physician assistants and "expanded function dental auxiliaries"[73] because of their promise of improving productivity, reducing costs, and enhancing both availability and quality.[74] It is noteworthy that, while the House committee report mentions that HEW has spent $22.5 million on training and research for non-physician providers, it cites no empirical evidence that paramedicals have actually increased the availability of medical care in underserved areas. The report does mention two studies which concluded that the quality of care provided by paramedicals is comparable to that of physicians.[75] The advocates of paramedical providers have failed to clarify why such personnel would not respond to the same incentives that push doctors toward pleasant, cosmopolitan, overserved major urban areas. Also, the use of paramedicals, especially physician's assistants, has been hindered by political efforts at the state level mounted by medical societies and professional nursing groups, who wish to restrict or control the activities of physician's assistants to prevent them from competing with doctors and nurses.[76]

On the other hand, in the course of condemning the quality of care provided by FMG's, the Committee notes that they have "become one of the primary sources of inexpensive medical manpower for many U.S. hospitals, long term care institutions, mental health hospitals, and prisons,"[77] while earlier in the report it is argued that FMG's have contributed to geographic maldistribu-

---

72 H.R. 5546 would eventually set the limit at 125%. H.R. 5546, 94th Cong., 1st Sess. § 801 (1975) (which would have added § 1701(a) to the Public Health Service Act). S. 3239 would eneutally set the limit at 110%, S. 3239, 94th Cong., 2d Sess. § 501 (1976) (which would add § 798(a)(1) to the Public Health Service Act), and also impose immigration restrictions on FMG's, *id.* § 601.

73 H.R. 5546, 94th Cong., 1st Sess. § 609 (1975) (which would add § 789 to the Public Health Service Act) (authorizes $90 million over a three year period); S. 3239, 94th Cong., 2d Sess. § 1201 (1976).

74 H.R. REP., *supra* note 13, at 59-60.

75 *Id.*

76 Kissam, *supra* note 11, at 19-20, 31-32. See note 160 and accompanying text *infra*.

77 H.R. REP., *supra* note 13, at 50.

tion by locating primarily in urban areas.[78] If "U.S. hospitals" includes inner city hospitals, the Committee has listed four of the most medically underserved facilities in the health care system. One wonders whether the Committee believes the solution is to send the "inferior" FMG's home to practice in their native countries (which need their services but lack sophisticated enough health systems to utilize their skills) and replace them with American paramedicals (who have much less training), since the compensation for serving in such facilities is considered inadequate by U.S. physicians. It is difficult to reconcile congressional rejection of low cost/low skilled care in one context with praise for it in another.

C. *Conclusion*

This review of the inadequacies of recent congressional attempts to reform the supply of health care demonstrates the difficulty of legislative solutions. In this context the apparent recent decline in enthusiasm for national health insurance is understandable. The Medicare/Medicaid experience appears to imply that cost-reducing, effectiveness-increasing reforms are required before solutions to the access problem can be adopted at a politically acceptable cost. But in light of the failure or lack of promise of recent congressional initiatives, what strategy should Congress now pursue? Before making suggestions for new legislation designed to alter the economic behavior of health providers, a more precise analysis of the incentives at the root of the current problem is in order.

III. Roots of the Current Problem

What explains the apparent incongruity between high and rising medical care costs, which make access increasingly difficult, and increasing doubts about the efficacy of health care? Any analysis of the health care cost problem must begin with the bounteous funding which has been made available to health care providers since the beginning of the "access" period. This expansion of

78 *Id.* at 47-48.

demand began well before Medicare and Medicaid with the increase in private health (especially hospital) insurance, partly through more generous employee benefits programs.[79] The inflation which resulted is attributable to unique financial incentives and behavioral factors present on both the demand (consumer) and supply (provider) sides of the health care market.

A. *Inflationary Incentives*

1. Consumers: Demand

The typical hospital patient has little financial incentive to be concerned with costs. In non-profit, non-government hospitals, only about 10% of all costs were borne directly by private patients in 1971.[80] The rest came from various public and private third party payors. When his insurance or the government covers the costs, the patient — anxious to get "the best possible care" — will utilize services for which he would be unwilling to pay the entire cost directly.[81] Even where he has to pay, his own cost consciousness is liable to be more limited than for non-medical services, given the aversion to low quality/low cost care.

A patient who does make an effort to get the most for his health care dollar is confronted with a market structure which makes "shopping around" very difficult. He usually enters the health care system by choosing a physician. Beyond this point, however, decisions about the type and volume of services utilized and where they are delivered are made largely by the physician.

This restriction on consumer choice is reinforced by the fact that in the typical community with two to five hospitals, medical staff privileges often do not overlap extensively.[82] Many doctors affiliated with one hospital do not have access to others. As a result, for any one doctor/patient purchaser, there may well be only one choice at the time when hospital care is required. This is

---

79 K. DAVIS, *supra* note 32, at 31-32; H. SOMERS & A. SOMERS, DOCTORS, PATIENTS, AND HEALTH INSURANCE 230-31, 249, 252, 256 (1961); Havighurst & Blumstein, *supra* note 27, at 13-14.

80 Pettengill, *The Financial Position of Private Community Hospitals, 1961-72*, 36 SOCIAL SECURITY BULL., Nov. 1973, at 3, 4 (Table 1).

81 Havighurst & Blumstein, *supra* note 27, at 15-18.

82 H. KLARMAN, THE ECONOMICS OF HEALTH 24 (1965).

because many buyers will return to a previously chosen physician once health problems arise, in which case competition will not operate to shift consumer choices. Of course, a very few sophisticated buyers may consider hospital privileges when choosing a doctor. But even fewer know, or give much weight to, the relative charges of hospitals with which alternate doctors are affiliated. Many other characteristics of doctors (and hospitals) — *e.g.*, reputation and academic affiliation — may be much more important to them. Rational consumers — even if they lacked health insurance — could not be expected to spend much time making such cost comparisons. Thus, almost all price competition vanishes, and with it many of the incentives to control costs by proper management.

With regard to physician services, patients rarely have readily available the information — fees, hospital affiliation, availability, billing procedures, etc. — which is necessary to informed choice.[83] In addition, many areas have few doctors among which to choose. In fact, in 1970 132 rural counties in the United States had no doctor at all.[84] Low income inner city dwellers who cannot afford physicians are often forced to obtain ambulatory care from the outpatient department of teaching hospitals or the local city or county hospital.[85]

In fact, the actual purchasers of hospital care are not the patients but the third party payors: the federal and state governments (through Medicare and Medicaid) and Blue Cross and other private health insurers. Any efforts toward cost minimization must come from them. Unfortunately, the system of cost-reimbursement fee-for-service financing they employ contains instead a powerful incentive for cost maximization: the more costs which are added into the reimbursement formula the greater the hospital's revenue. Hospital managers with the bias toward high "quality," high

---

83 This dearth of data is partially attributable to state statutes which prohibit in varying degrees advertising by physicians. For a compilation of such statutes plus an overview of the effort by consumer groups to compile directories of physicians, see Health Research Group, A Guide for Compiling a Consumers Directory of Doctors, May 1975 (statutes at 33) (available from Public Citizen's Health Research Group, Washington, D.C.).

84 J. Winsten, *supra* note 13, at 46.

85 *Id.* at 45-46.

technology care discussed below[86] thus can tap a relatively open-ended source of funds to pay for new wings and equipment and additional specialized, highly-trained personnel.[87] Of course, all of these costs are eventually passed on to health insurance subscribers or taxpayers.

Also, many services are not covered unless they are delivered in the hospital,[88] which supposedly is an assurance of quality (and a way of limiting the insurer's coverage) but also is much more expensive than outpatient delivery. Where a service can be delivered in either setting, the patient and his physician invariably choose the hospital.

2. Providers: Supply

One phenomenon at work on the supply side of the medical market is called the "Baumol-Bowen effect," after the authors of a well-known work on the economics of the performing arts.[89] Baumol and Bowen explored what happens in a society in which most jobs are in technically progressive industries where productivity (output per worker over time) is steadily rising. However, there are also some sectors in which there are minimal or no productivity increases but whose workers nonetheless generally succeed in matching the advancing wage rates of those in other lines of work. In order to pay these rising wages, costs and prices in the unproductive sector must continually increase relative to those of other parts of the economy. In contrast, prices in the technically progressive sector could remain relatively stable, with the higher wages "paid for" by increased productivity.[90]

In recent years it seems clear that medical care has been such a "non-progressive" sector — along with the arts, education, and many government services. In medicine, technological changes may have increased quality, but relatively few have been labor

---

86 See text accompanying notes 96-101 *infra*.

87 B. Ensminger, The $8 Billion Hospital Bed Overrun 43-47, 1975 (available from Public Citizen's Health Research Group, Washington, D.C.); Havighurst & Blumstein, *supra* note 27, at 14; Lee, *A Conspicuous Production Theory of Hospital Behavior*, 38 S. ECON. J. 48 (1971).

88 B. Ensminger, *supra* note 87, at 44.

89 W. BAUMOL & W. BOWEN, PERFORMING ARTS — THE ECONOMIC DILEMMA (1966).

90 *Id.* at 167-72.

saving.[91] For example, any one obstetrician can only deliver so many babies a year. If he is to keep his income position relative to a machinist — who now, thanks to automation, produces much more than he did ten or fifteen years ago[92] — his fees will have to go up relative to the prices of products made by machinists. Thus, where technical change does not increase productivity, medical care cost increases are inevitable unless providers accept lower relative incomes than they now enjoy.[93]

Some technological developments in medicine have saved effective labor (*e.g.*, use of technicians instead of doctors). But the gains have been small, especially in comparison to the *labor-using* impact of many new techniques. Indeed, what is ironic about many medical innovations is that they use more of *both* capital and labor (*e.g.*, intensive care units in hospitals) instead of substituting one for the other, or lessening the use of both.[94]

However, if significantly better (*i.e.*, more effective) medical care results from these innovations, then the usual cost statistics are systematically deceptive. Some of the cost increase can be attributed to higher quality. If one switches from a Chevrolet to a Mercedes, more has changed than the price he is paying for automobile transportation.[95] But suppose the new car is not a Mercedes but instead a fancy Edsel, *i.e.*, different and more expensive but not necessarily more effective. Unfortunately, the latter seems to be the more accurate description of what has been occurring in the health sector. A higher volume of more sophisticated, technological services are being delivered by more highly

---

91 There is no question that doctor productivity greatly increased in the period before and after World War II, as advanced technology, the use of the automobile and the decline in the practice of housecalls increased the number of patients a doctor could see. H. SOMERS & A. SOMERS, *supra* note 79, at 48-51 (1961). But by the mid-1960's this trend had apparently begun to reverse itself. H. SOMERS & A. SOMERS, MEDICARE AND THE HOSPITALS 106-07 (1967).

92 STATISTICAL ABSTRACT OF THE UNITED STATES: 1975, at 358 (1975) (Table No. 584).

93 In fact, the relative wages of hospital workers were often above the earnings of comparable non-hospital workers in 1969. M. FELDSTEIN, THE RISING COST OF HOSPITAL CARE 57-61 (1971).

94 In hospitals, where both assets and personnel per bed have increased, the failure of cost-saving technical change is especially evident. Lee, *supra* note 87, at 56. *See* M. FELDSTEIN, *supra* note 92, at 36-51.

95 The problem of correcting price-indexes for quality variations is a risky and difficult one. *See* PRICE INDEXES AND QUALITY CHANGE (Z. Griliches ed. 1971).

trained personnel without any discernible enhancement of the health status of the population.

The pattern of increasing technological sophistication is itself a result of the incentive structure of the health sector. Doctors acquire their reputations by scientific advances and new techniques or devices. Little credit goes to those who introduce cost-reducing innovations. The developers of cost-saving technology do not find a ready market among either doctors or hospital managers with other objectives.[96] Such behavior can be explained by the incentives which influence the key decision-makers in the largely nonprofit hospital system: hospital administrators and physicians.

Professional hospital administrators do not enjoy peer recognition or career advancement for instituting successful cost control efforts. The hospital Boards of Trustees who hire and fire administrators (generally "community leaders" who serve without pay) tend to measure the institution's success in terms of its growth and prestige, which they find both ethically necessary and personally gratifying. Hence, there are two courses of action which the administrator's peers and board will almost always view with favor. One is an effort to deliver a higher "level" of care by acquiring more highly-trained and specialized medical, nursing, and technical staff, and more modern and exotic equipment. Elaborate equipment often is necessary to attract a prestigious medical staff. (This is not meant to imply that administrators are not concerned with patient care; to the contrary, they firmly believe that such changes improve quality.) The other is a successfully orchestrated program of expansion. Both are effective pathways to a higher salary or a better job.[97]

As explained above,[98] because of cost-reimbursement financing,

---

96 In general, cost reducing innovations are made by an industry's equipment supplier. *E.g.*, Peck, *Inventions in the Postwar American Aluminum Industry*, in UNIVERSITIES — NAT'L BUREAU COMM. FOR ECON. RESEARCH, THE RATE AND DIRECTION OF INVENTIVE ACTIVITY 292-93 (1962).

97 B. Ensminger, *supra* note 87, at 42-43. At least two economists who have attempted to develop theories of hospital behavior emphasize the importance of such factors. Lee, *supra* note 87; Newhouse, *Toward a Theory of Nonprofit Institutions: An Economic Model of a Hospital*, 60 AM. ECON. REV. 64 (1970). While these models are no doubt oversimplications, they do illustrate widely shared perceptions about hospital behavior.

98 See text accompanying notes 85-88 *supra*.

the administrator can pass the cost of "quality" improvement on to the third party payor. He also can increase revenue by delivering more services — a tendency which is exacerbated by the unused capacity caused by new construction. Once new beds are built (or if old ones are not occupied), the administrator, in concert with his medical staff,[99] will encourage utilization in order to fill the new or empty beds — *i.e.*, to maximize the hospital's occupancy rate. When it approaches 100%, he may be able to justify yet another expansion of his hospital. In addition, administrators pay a price for failure to use the new capacity; beds which sit empty are a significant fiscal burden, since they cost a great deal to maintain, in terms of overhead and required staffing, whether or not occupied and paid for by a patient.[100] This cost, too, must be reimbursed by the third party.

Once the administrator has increased the number of beds, he and his medical staff have an incentive to generate the demand necessary to utilize the additional capacity and produce the revenue needed to pay for it. Indeed, the reality of this vicious circle is the rationale for health planning and certificate of need regulation of facility construction. Thus, the administrator is encouraged to increase costs not only by raising quality but also by increasing utilization.

In contrast, cost control offers every prospect of making enemies with important constituencies. The antagonism of doctors and employee unions leads to routine job harassment and may impair the administrator's relationships with his board as well as his broader reputation. Boards in particular dislike having to become involved in internecine disputes or having "their" institution subjected to the bad publicity caused by strikes or public complaints from the medical staff.[101]

---

99 The physician also has his own incentives to overutilize the hospital. See text accompanying notes 101-06 *infra*.

100 It has been estimated that an unoccupied bed costs anywhere from 35% to 75% as much as an occupied bed to maintain. *See* Lave & Lave, *Hospital Cost Functions*, 60 AM. ECON. REV. 379, 393-94 (1970).

101 *See generally* LeRocker & Howard, *What Decisions Do Trustees Actually Make?*, 94 MODERN HOSP., Apr. 1960, at 83. In private firms Boards of Directors often become involved only in a significant crisis. M. MACE, DIRECTORS, MYTH AND REALITY 39-42 (1971).

Even if he wanted to exercise some controls over cost, the administrator's power is seriously limited because utilization decisions are made by the doctor dealing with a particular case. For example, through good management an administrator might reduce the costs of providing a series of lab tests. But he still has virtually no control over how many series of tests are ordered by the physician because the medical staffs of most hospitals are affiliated doctors with staff privileges who are paid not by the hospital but by the patient.

The doctor has several reasons to overutilize hospital resources. First, patients are generally impressed by physicians who seem to indicate their concern with the patient's condition by ordering many ancillary services. Second, the physician's fear of subsequent malpractice claims may lead to defensive medicine — ordering additional services not because he believes they are medically necessary but to buttress a future legal defense.[102] (Actually, assuming that a particular physician will negligently cause injury in the performance of a percentage of all cases, the more services he orders the greater is his exposure to potential liability.) Third, doctors understand that the financial viability of the hospital, and hence its freedom to spend funds in ways that doctors would like, depends on maximizing the utilization of its beds and services. Such perceptions influence physician choices about length of stay, whether a procedure that is only marginally indicated should be performed, and whether a service should be performed on an inpatient or outpatient basis.[103] Finally, the doctor has a financial incentive to provide more of those services — notably surgery — for which he, rather than the hospital, is paid.[104]

Evidence is accumulating that doctors have quite a substantial

---

102 Simmons & Ball, *PSRO and the Dissolution of the Malpractice Suit*, 6 U. TOL. L. REV. 739, 753-57 (1975) and sources cited therein; Project, *The Medical Malpractice Threat: A Study of Defensive Medicine*, 1971 DUKE L.J. 939 (1971). There is no conclusive empirical evidence that defensive medicine alone is a significant cause of overutilization.

103 Occupancy rates often influence average lengths of stay, with stays increasing as occupancy rates fall to low levels. *E.g.*, L. Bilheimer, Hospitals in Arkansas 172-77, July, 1974 (unpublished thesis in Harvard University Archives).

104 Monsma, *Marginal Revenue and the Demand for Physicians' Services*, in EMPIRICAL STUDIES IN HEALTH ECONOMICS 145 (H. Klarman ed. 1970).

ability to influence the demand for their own services.[105] They can vary how soon (and under what conditions) they ask a sick patient to return for a follow-up check-up. They can vary how many surgical procedures and tests they perform themselves. And they directly control hospital admissions and lengths of stay. Of course the fact that more surgery is done where there are more surgeons may only mean that the populations being treated in the two areas differ in their need for surgery. But in two studies which took such variations into account, the evidence for the view that "supply creates demand" was nevertheless quite strong.[106]

The concentration of surgeons also seems to be related to the income surgeons derive from each operation. As the density of surgeons rises, procedures performed per surgeon — workload — actually declines,[107] even though more operations are done relative to population. However, there is also evidence that, as the concentration of surgeons relative to population rises, net income per operation increases,[108] which has the effect of at least partially offsetting the lower workload. Apparently, the greater the number of operations per surgeon, the less implicit pressure there is to maintain relatively high fee schedules.

B. *Economic Analysis*

Considered against this background, the increase in hospital costs of recent years involves more than just the effect of expanded demand (brought on by Medicaid and Medicare) on a supply schedule which would only allow a larger volume to be produced with higher costs. Rather, what seems to have occurred is that high levels of demand provided the surplus revenue which hos-

---

105 *Id.* at 147-48; Reinhardt, *Alternative Methods of Reimbursing Non-Institutional Providers of Health Services*, in CONTROLS ON HEALTH CARE, *supra* note 32, at 142-43; Hughes, Fuchs, Jacoby & Lewit, *Surgical Workloads in a Community Practice*, 71 SURGERY 315 (1972).

106 Wennberg & Gittelsohn, *Small Area Variations in Health Care Delivery*, 182 SCIENCE 1102 (1973); Wennberg & Gittelsohn, *Health Care Delivery in Maine I: Patterns of Use of Common Surgical Procedures*, 66 J. ME. MED. ASS'N, May 1975, at 123.

107 AM. COLLEGE OF SURGEONS & AM. SURGICAL ASS'N, SURGERY IN THE UNITED STATES 75 (1975) (Figure 14).

108 *Id.* at 66, 74, 75 (calculated by the authors from information presented in Figures 9, 13, and 14).

pitals used either to increase volume by expanding capacity or to raise "quality" by investing in more expensive plant, equipment, and personnel. For example, many hospitals abandoned old, supposedly high cost facilities for new ones whose capital costs were so high that total costs increased even though operating costs declined.[109] In other words, third party financing removed the financial burden from both patient and provider, thus allowing them to concentrate their attention on attempting to attain the highest possible quality of care without regard to cost.

Graphically, the picture economists usually draw to show the effect of a demand increase on output and price is given in Figure 1. Two alternative demand curves, which show how much con-

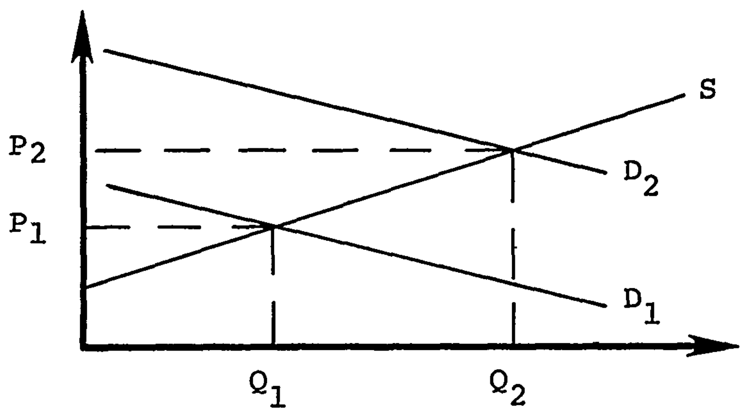

Figure 1

sumers will buy at each possible price, are labeled D_1 and D_2; the supply curve, which describes how much producers will sell at each price, is labeled S. When demand increases from D_1 to D_2,

---

109 Conversation with Dr. William Bicknell, former Massachusetts Commissioner of Public Health, in Boston, May 14, 1976. For example, a recent Massachusetts certificate of need application, which proposed renovating and upgrading a hospital, estimated that costs both per patient day and per outpatient procedure would be substantially higher if the remodeling was undertaken than if it was not. Certificate of Need Agency, Massachusetts Dep't of Public Health, Staff Summary for Determination of Need by the Public Health Council 13, 1975 (Waltham Hospital, Project Number 3-2464, on file with Mass. Dep't of Public Health, Boston, Mass.).

consumers eager to buy more bid up the price (from P_1 to P_2), while producers respond by expanding supply (from Q_1 to Q_2).

In medical care, in contrast, it may well be the case that, because of third party payment and consumer attitudes, demand is relatively rather insensitive to prices at current average price levels[110] (*i.e.*, an increase in price results in little or no reduction in the quantity of services demanded by consumers, so that the demand curve is almost vertical). At the same time, demand will vary in response to changes in various structural and situational variables like hospital occupancy rates, doctors per capita, etc. The demand curve is not well defined but rather a zone of possible demands whose boundaries cannot be precisely delineated, as shown in Figure 2.

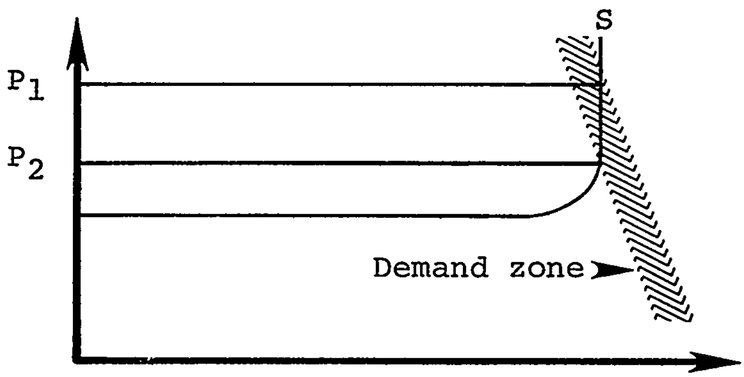

Figure 2

On the supply side, furthermore, it is very difficult to expand output beyond a certain point because new buildings or physicians take so long to produce (although if prices go too low, hospitals fail to cover out-of-pocket costs and are forced to close). Once the system reaches full capacity, only small expansions (*e.g.*, setting up beds in the hall) are possible in the short run. This leads to a supply curve with a relatively distinct corner, as shown in Figure 2.[111] As a result of these features, it is possible for the supply

110 H. KLARMAN, *supra* note 82, at 24-25.
111 *Id.* at 78-79, 104-6.

of and the demand for hospital care to be more or less equal at a wide range of prices (anywhere between P_1 and P_2 in Figure 2). Neither demand nor supply varies much in response to prices in the short run, while in the long run demand responds to non-price factors in a fashion which tends to ensure such equality.

The argument that supply changes will in turn modify demand because physicians control utilization of health services is illustrated in Figure 3. In the initial situation, supply and demand

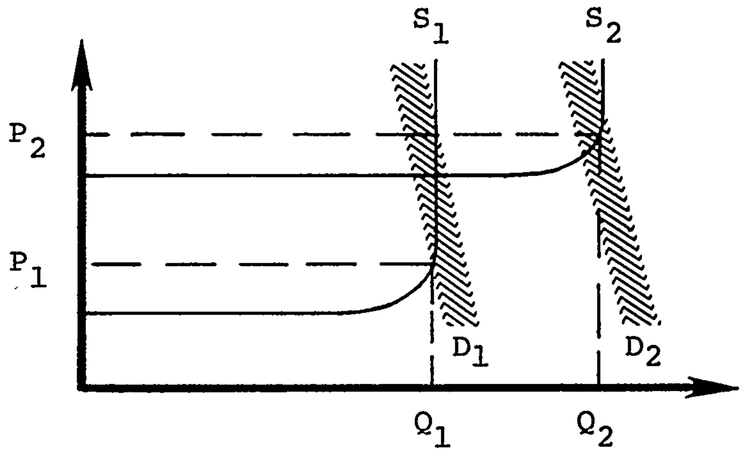

Figure 3

(S_1 and D_1) lead to price and quantity P_1 and Q_1. Then insurance and Medicare, etc., increase demand. Seeing the potential for cost reimbursement, hospitals expand both capital and labor, moving toward supply curve S_2, while prices rise steadily. To fill the new capacity, the system generates a still further increase in demand, to D_2, while prices continue to adjust until they reach P_2, and quantity becomes Q_2. Again, however, due to the imperfections in the way these markets function, a significant range of prices and outputs is compatible with apparent equilibrium. That is, demand may not increase enough to insure full utilization of all new facilities, while at the same time suppliers do not respond to such a surplus either by price cutting or by withdrawing from the market.

IV. Recommendations for Government Action

In order for the health care system to deliver efficacious services at a reasonable cost, it must undergo a major reconstruction. First, institutional arrangements which discourage the use of expensive services such as hospital care and encourage the development and use of cheaper alternatives must be created and promoted. Second, a comprehensive, skeptical, empirical evaluation of existing medical procedures and practices must be undertaken, together with a major research and development program to generate new technology to deliver necessary services more efficiently. Third, limitations on the production of surgeons and other specialists and removal of the legal and practical constraints which inhibit the cost-effective use of paramedical manpower must be initiated. Only if such changes are brought about can the increased demand which would be created by national health insurance[112] be absorbed without devoting a larger and larger proportion of our resources to a health care system which has little impact on people's health.

A. *Health Maintenance Organizations*

There is now wide agreement among health care reformers that Health Maintenance Organizations (HMO's) are the most promising answer to the need for economic incentives for more cost effective delivery of services.[113] The HMO strategy is also attractive for other reasons.

1. Advantages of HMO's

As explained above,[114] HMO's have no reason to overprovide

---

112 While this section will not discuss the details of how best to assure access to a minimum level of health services to everyone without regard to income, it is clear that such a financing scheme faces major political obstacles until Congress is confident that it will not just exacerbate the inflationary spiral and lead to intolerable demands on the federal treasury.

113 Havighurst & Bovbjerg, *supra* note 46, at 384-85 & nn. 13 & 14. One of the staunchest and most vociferous advocates of HMO's has been Dr. Paul Elwood, whose most recent paper is Elwood, *Alternatives to Regulation: Improving the Market*, in CONTROLS ON HEALTH CARE, *supra* note 32, at 49.

114 See text accompanying notes 35-36 *supra*.

services. If they achieve a surplus, money is available to lower their members' capitation (per person) fees in an effort to attract new subscribers, to expand or modernize their facilities, or to hire new personnel or pay them year-end bonuses. In order for the HMO's cost-minimizing behavior to benefit its members rather than the owners or managers, however, there must be enough population in the service area to support more than one, and preferably several competing providers (HMO's, group practices, or fee-for-service providers). Otherwise there will be no pressure on the HMO to limit its membership fees and thus no savings for the consumer.

While they vary in the services they provide, HMO's in general have been quite successful in saving costs — primarily by keeping their subscribers out of the hospital.[115] In some cases the staff physicians have been given a direct financial incentive to minimize hospitalization (by being compensated with salaries plus surplus-sharing rather than fee-for-service), while other HMO's have relied more on the selective recruitment and socialization of their medical staffs.[116] Since HMO's are usually "vertically integrated" (*i.e.*, they provide all "levels" of care to their members), they have an incentive to match patients to the most cost-effective delivery setting (home, outpatient clinic, nursing home, community hospital, teaching hospital, etc.).

The HMO does not eliminate the health professional's bias in favor of "high quality" services and expansion of facilities.[117] However, the HMO manager must persuade present or potential subscribers to pay premiums high enough to cover the associated costs, in a context where lower-cost options are available. In stark contrast to the cost-reimbursement fee-for-service system, market incentives now constrain professional behavior.

Another advantage of HMO's is that they are largely self-regulating. There should be little or no need for such regulatory mechanisms as fee or rate setting, certificate of need, or utilization

115 Elwood, *supra* note 113, at 58-60; Roemer & Shonick, *HMO Performance: The Recent Evidence*, 51 MILBANK MEMORIAL FUND Q. 271 (1973).

116 Conversation with Dr. Richard Nesson, former Medical Director, Harvard Community Health Plan, in Boston, Mar. 1976.

117 See text accompanying notes 96-97 *supra*.

review in whatever part of the market is covered by competing HMO's. The government, however, would still have an important role to play in such areas as financing to alleviate distributional problems, data collection, research and development, and quality control. Perhaps even some of these functions could be reduced or eliminated if health consumers had freer access to better evaluative data about medical providers and procedures.

If the payment mechanism between the government and the HMO is properly designed,[118] the consumer will be able to choose among varying degrees of comprehensiveness of coverage and varying prices. Allowing the consumer the flexibility to dispose of his income would, from the perspective of economic analysis, maximize his well-being. For example, a subscriber could decide not to buy coverage for heart surgery because the ability to spend the money saved on other goods and services is more important to him than the risk of heart disease. The consumer can also choose between a prepaid plan and the traditional fee-for-service system, depending on which he believes offers the most benefit for his expenditure.

2. Problems of HMO's

HMO critics often point to the fact that when the financial incentives which influence health managers are reversed, there is a tendency to underprovide services or to deliver services of low quality (either of which could increase the probability of patient injury) in an effort to save costs.[119] Such temptations are particularly strong in for-profit HMO's.

There are at least three factors which tend to mitigate the likelihood of underutilization or medical injury in an HMO. First, the possibility of malpractice liability is always a deterrent to the negligent performance of care as well as to a negligent decision not to deliver a service.[120] Second, professional norms in favor of high quality and high utilization still will exist,[121] although the

---

118 See text accompanying notes 133-37 *infra.*

119 Schneider & Stern, *supra* note 35, at 97-98.

120 Curran & Moseley, *The Malpractice Experience of Health Maintenance Organizations,* 70 Nw. U.L. Rev. 69, 89 (1975).

121 See text accompanying notes 101-04 *supra.*

attitudes of HMO physicians may be altered by financial incentives to minimize utilization. Also, the retrospective quality-oriented component of the PSRO review system should discourage incompetent delivery. PSRO's tend to increase the flow of evaluative data among physicians about the quality of practice. Comparative data on individual physicians could have the effect of improving the clinical performance of less competent doctors, especially if integrated with continuing medical education.[122] Third, assuming that HMO's will be in competition among themselves and with fee-for-service providers, an HMO which systematically underprovides or acquires a reputation for poor quality care will begin to lose members.

Another serious problem with the HMO approach is that incentives for cost minimization also encourage "skimming" or the avoidance of "bad risks." It is in the financial interest of HMO's (as it is for health insurers) to prevent those who will create high demand for covered services (*i.e.*, people in poor health) from enrolling in their plans for the standard fee. If those most in need of services become members, the HMO will either have to deny them necessary services, charge them higher rates, charge all members higher rates so that the healthy subsidize the sick (which tends to drive away the subsidizers), or suffer the financial consequences and eventually go out of business.[123] The only way to avoid such a dilemma is for the government to require open enrollment and accomplish the cross-subsidization through the tax system by adjusting reimbursement to HMO's acording to the mix of patients enrolled in the plan. Such a payment mechanism will be discussed below.[124]

---

122 The PSRO system for performing retrospective quality review is called Medical Care Evaluation. Data on a particular medical problem is collected and compared to predetermined standards of expected quality. If deficiencies are identified, the hospital must undertake a program of corrective action (often, but not always, an educational program) and then restudy the problem to see whether it was remedied. Goran *et al.*, *supra* note 44, at 15-22; Jessee, Munier, Fielding & Goran, *PSRO: An Educational Force for Improving Quality of Care*, 292 NEW ENG. J. MED. 668, 670 (1975).

123 For further discussion of skimming and experience rating, particularly with regard to their impact upon the poor, see Schneider & Stern, *supra* note 35, at 98-101, and sources cited therein.

124 See text accompanying notes 133-37 *infra*.

3. Implementation of HMO's

The fact that HMO's have rarely been successful enough to become financially self-sufficient[125] is disappointing and somewhat surprising, given their apparent desirability. Even some of the most established systems have experienced failures.[126] Middle class consumers are often reluctant to abandon their personal relationship with a particular physician (though it is not at all clear that physician-patient relations are more impersonal in a well-managed HMO).[127] But there are more important and pervasive reasons for the lack of success of the HMO movement, which suggest that the prepayment approach has yet to be given a "fair market test."

An experienced HMO manager has analogized the Health Maintenance Organization Act of 1973 to

> a hypothetical Congressional enactment which would attempt to stimulate sales of small cars by: (1) providing a minimal amount of federal money for the creation of new small car dealerships; (2) requiring all dealers, regardless of their affiliation, to stock a full supply of parts for all small cars; and (3) mandating that all small cars meet emissions and safety standards far more rigorous, and expensive, than those applicable to full-size cars.[128]

The problem, as discussed above,[129] is that the comprehensive benefit packages and long open enrollment periods required by the law make it nearly impossible for an HMO to obtain federal start-up assistance and still compete with health insurers who offer significantly lower rates because of narrower coverage and avoidance of bad risks. The differences in coverage but not the differences

---

125 Note, *The Role of Prepaid Group Practice in Relieving the Medical Care Crisis*, 84 HARV. L. REV. 887, 949-53 (1971) [hereinafter cited as *Prepaid Group Practice*].

126 For example, both the Kaiser Health Plan and the Health Insurance Plan (HIP) of New York have encountered substantial difficulties in raising capital for expansion. *Id.* at 949-50.

127 Curran & Moseley, *supra* note 120, at 81. In an informal, statistically nonrigorous survey of HMO's, Curran and Moseley found no evidence that the incidence of malpractice claims was higher for HMO's than for fee-for-service providers. *Id.* at 79-80.

128 Address by Thomas Pyle, administrator of the Harvard Community Health Plan, to the National Health Lawyers Ass'n, Boston, Mass., Apr. 26, 1974, paraphrased by Rosoff, *supra* note 35, at 215.

129 See text accompanying notes 36-40 *supra*.

in price tend to be obscured in the minds of consumers by the very imperfect information available to consumers in the health care market. It is encouraging that there is now pending in Congress legislation which would amend the 1973 Act to eliminate or relax these restrictions.[130]

The success which the established HMO's have enjoyed may be largely attributable to their being an embattled minority selectively recruiting those personnel who were relatively sympathetic to their goals and methods.[131] Many physicians are opposed to prepaid group practices and HMO's because they believe such approaches threaten the independence, prestige, and income they enjoy under the traditional solo practice, fee-for-service system. The opposition of organized medicine to HMO's has manifested itself not only in lobbying efforts against state and federal legislation supportive of HMO's but also in a variety of practices on the state and local level which have restricted the development of prepaid group health plans.[132] Thus, the development and ex-

---

130 H.R. 9019, 94th Cong., 1st Sess. §§ 4-6 (1975). H.R. 9019 would eliminate the requirement for an annual period of open enrollment, waive for five years the requirement for community rating of HMO premiums, and reduce the comprehensiveness of the benefits required of HMO's to qualify for federal funds under the 1973 Act. It passed the House by a large majority with no changes to these provisions. 121 CONG. REC. H10809-10 (daily ed. Nov. 7, 1975). The Senate counterpart to H.R. 9019 is S. 1926, 94th Cong., 1st Sess. (1975). As this article was going to press, a favorable report on S. 1926 by the Committee on Labor and Public Welfare was about to be filed. The reported bill has been amended, and there are technical differences between it and H.R. 9019. S. 1926 as amended would still require open enrollment, but the requirement is much more limited than in the 1973 Act and applies only to established, financially sound HMO's; community rating would be required after three years rather than five. The required benefits package would be narrowed, but the categories included and excluded are not all the same as in H.R. 9019. Telephone conversation with Ned Kelly, member of Senator Schweiker's staff, Washington, D.C., Apr. 30, 1976.

131 D. MECHANIC, THE GROWTH OF BUREAUCRATIC MEDICINE 83-98 (1976).

132 The constraints imposed by the medical establishment on HMO's were probably more overt in the past than they are now, but restrictive practices still exist in some areas, including professional ethical rules against advertising, denial of hospital staff privileges to HMO doctors, expulsion from or denial of admission to local medical societies, and refusal to refer patients to HMO physicians. Havighurst, *Health Maintenance Organizations and the Market for Health Services*, 35 LAW & CONTEMP. PROB. 716, 767-77 (1970) (discusses defensive, medical society-dominated prepayment plans and antitrust implications); *Prepaid Group Practice, supra* note 125, at 954-59. There are also state statutory and common law restrictions on the development of certain types of HMO's. Holley & Carlson, *The Legal Context for the Development of Health Maintenance Organizations*, 24 STAN. L. REV. 644, 653-62 (1972); *Prepaid Group Practice, supra* note 125, at 960-75; INSTITUTE OF MEDICINE, *supra* note 40, at 19-29. Most of these laws were overridden with

pansion of HMO's, at least in the near future, may only be realistic in those areas where the HMO can draw upon a pool of relatively young, idealistic, progressive doctors who are receptive to a non-traditional form of practice and where the rest of the local medical community does not actively erect barriers to the HMO's establishment and recruitment of members.

The mechanism by which HMO's are paid by the government (under Medicare, Medicaid or, eventually, national health insurance) or other third parties must preserve both market incentives and subsidization of the sick and the poor by the healthy and the affluent. The Social Security Act has been amended so that the Social Security Administration may enter into contracts with HMO's under which Medicare intermediaries make capitation payments to the HMO based on its annual operating budget and enrollment forecast for the next year.[133] Although it is encouraging that Congress has enabled Medicare and Medicaid beneficiaries to enroll in HMO's, two problems already discussed[134] are not addressed by such an approach. First, in order to make it possible for government beneficiaries effectively to purchase different levels of health care for different prices or to choose among competing HMO's, the system must allow the consumer to keep the money saved if he chooses a lower level of care or cheaper HMO and require him to pay more for a higher level of care or more expensive HMO as compared to a standard benefits package and an average HMO. Second, in order to make open enrollment financially feasible for the HMO, the government must somehow subsidize the sick, who need a higher level of services. Since existing levels of expenditure in most HMO's reflect the relatively healthy people who are now members,[135] a capitation payment based on current budgets will not adequately reimburse HMO's for Medicare and Medicaid recipients, who will demand more services than present members.

A payment mechanism which adjusts for both of these factors

---

respect to federally-qualified HMO's by the Health Maintenance Organization Act of 1973, § 2, Public Health Service Act § 1311, 42 U.S.C. § 300e-10 (Supp. III, 1973).

133 Social Security Act § 1876, 42 U.S.C. § 1395mm (Supp. III, 1973).

134 For discussion of the consumer choice problem see text accompanying notes 118-19 *supra*. For discussion of the skimming problem, see text accompanying notes 122-24 *supra*.

135 Schneider & Stern, *supra* note 35, at 99-100.

would look something like this: whatever is established as the base capitation payment, there would have to be provision for a rebate to reward those government beneficiaries who choose either a benefit package narrower than the standard plan or an HMO lower than average cost. The rebate could be paid by either the government or the HMO (where fewer benefits are chosen). It would be possible to have rebates of different amounts for "budget" benefit plans of varying expense. In this fashion the consumer receives the money which the government otherwise would have spent on a basic range of health services by taking the risk that he would not be covered for some services he may ultimately need or want (in which case he would have to make private payment). On the other hand, anyone who wanted coverage more comprehensive than the standard package or wanted to join a relatively expensive HMO would have to pay for it with his own funds, which means that the more affluent can have "high style" care if they bear the additional cost.[136]

It will be necessary to put some limits on such a scheme in order to prevent abuses. First, in a universal national health insurance system, the plan would be regressive because it would provide the same dollar value in benefits (services or rebates) for rich and poor; *i.e.,* it is the reverse of a "head tax." This undesirable effect would be avoided by either excluding the affluent from the government plan or including everyone but taxing the benefits as income. Second, it would probably be wise to specify a minimum package of health benefits, including emergency care and other basic hospital services, for which all government beneficiaries would have to be covered, and which would in turn put

136 Providers would be allowed to price different benefit plans however they wanted. Rebates would be equal to the difference between the price set by the HMO or other provider and a base rate set by the government financing agency. This rate would have to reflect what the government believes would be the cost of delivering a standard range of benefits for an efficient provider. If one assumes beneficiaries of average health, the cost to the government would always be the same. The only variable would be how much of the basic payment goes to the provider and how much to the beneficiary in the form of a rebate. Any costs above this base would have to be paid by the consumer from his own resources. If a provider lowers its price, it reduces the amount of government reimbursement but attracts consumers who are seeking rebates, *i.e.,* the system simulates a private market in which the consumer is spending his own money.

a ceiling on the amount of rebates. Otherwise, there might be a tendency for a poor person to forego all health care coverage and take the immediately available cash, reasoning that emergency services needed later would not be denied merely because he could not afford to pay cash. Of course, similar problems arise for rebates of any amount, but susceptibility to abuse is diminished if the beneficiary is not allowed to convert a health care financing system into a cash benefits-negative income tax program.

The problem of varying health conditions and thus varying need for care is more troublesome. In order to prevent the "sick" population (which will undoubtedly include a highly disproportionate number of the elderly) from being deprived of the opportunity to choose HMO's over fee-for-service providers, the government will have to subsidize utilization of services by the sick. This could be accomplished by sliding the base capitation payment to the HMO up or down according to the relative "healthiness" or need of its membership. The difficulty arises in devising a scheme for calculating need for services without re-introducing the cost reimbursement system. How does an HMO demonstrate a higher degree of "sickness" among its membership except by a retrospective analysis of services performed?

An HMO that wanted to manipulate such a system, however, could only do so by increasing utilization and hence cutting its profits or incurring a loss in the base year. Furthermore, it could only profit in the following year by cutting utilization, which would in turn result in a lower base for the year after that. Unfortunately, a non-profit HMO might be content to let cost and service levels drift upward over time, claiming that it had (but not actually having) an ever-sicker population and thus increasing its revenue. To the extent that the HMO is competing with other HMO's or health insurers, its subscribers might begin to switch to those with lower base rates, especially other HMO's. However, this does not prevent the government from over-subsidizing those that remain in the HMO which is overproviding services.

An alternative method of adjusting for sickness would be periodically to perform physical examinations on each member of the HMO, categorize expected demand for services, and establish a base rate for each person. Such a system could not be abused by increasing utilization, but it would be essential for the calculation of individual rates to be closely matched with costs to avoid giving the HMO an incentive or disincentive for enrolling certain categories of people who represent net revenue or net losses. The cost of administering such examinations must also be taken into account.

Thus it seems that a creative solution which would both redistribute health benefits toward the poor and preserve the market incentives of the HMO strategy is not beyond reach, although working out the details of such a scheme is a project worthy of another article.

4. Prospective Reimbursement

Pending the delayed development of an HMO system, incentives for cost containment can probably only be attained by changing methods of hospital reimbursement to a prospective basis.[137] Given the mixed history of efforts at detailed regulation, much can be said for setting such reimbursement levels by a sophisticated statistically based formula, rather than through detailed case-by-case and line-by-line examination.[138] Such complex bureaucratic processes are more likely to be subject to manipulation because of the hospitals' financial ability to employ expert lawyers and accountants as their advocates.[139] In the absence of competition, there is much to be said for basing such reimbursement on the number of admissions for illnesses of various types. While such a scheme could lead to diagnostic biases, at least it would provide reasons for the hospital to limit lengths of stay and the number of tests and services provided to each patient (although not first admissions).

There are other formulas which come closer to forcibly converting hospitals to HMO's (at least with respect to their government

---

137 K. BAUER, CONTAINING COSTS OF HEALTH SERVICES THROUGH INCENTIVE REIMBURSEMENT 73-74 (1973); Dowling, *Prospective Reimbursement of Hospitals*, 11 INQUIRY, Sept. 1974, at 164.

138 *E.g.*, Lave, Lave & Silverman, *A Proposal for Incentive Reimbursement for Hospitals*, 11 MED. CARE, Mar.-Apr. 1973, at 79. Some limits of the formula approach are discussed in Messier, *Caution: Prospective Reimbursement is No Panacea*, HOSP. FIN. MANAGEMENT, Sept. 1975, at 22.

139 *See* Dowling, *supra* note 137, at 174-79.

business) — for example, capitation payments or a fixed annual hospital budget, concepts which have been voluntarily adopted by some Blue Cross plans and state hospital associations.[140] Even these mechanisms would affect neither the continued fee-for-service delivery of physician services nor the reimbursement systems of Blue Cross and other hospitalization insurers who did not participate voluntarily, unless their role is eliminated by national health insurance.

B. *Research on Effectiveness and Efficiency*

1. Research and Development

Even if changes in the health care financing system are successful in encouraging more cost-conscious behavior by both providers and consumers, there remains a clear need for empirical research on both the efficacy of health services and the efficiency with which they are delivered to enable physicians, administrators, and policy makers to respond better to market incentives.

As noted earlier,[141] the government spends large sums on biomedical research to develop new medical techniques. It would seem to be rational to shift a portion of these research funds to sophisticated, experimental, outcome-oriented research[142] on the effectiveness of commonly used medical procedures and modes of treatment. The first task would be to develop a measurable definition of individual health status as a criterion of the benefits of particular medical practices. Integrating such research efforts with the data gathering and standard setting functions of PSRO's would facilitate systematic and ongoing evaluation of effectiveness and would also prevent PSRO's from becoming an anti-competitive device for institutionalizing and defending the unexamined status quo. The problem is that PSRO standards are required to reflect "typical patterns of practice" of an area.[143] If those patterns in-

140 *Id.* at 165-71.
141 See note 15 and accompanying text *supra.*
142 For the application of a technique known as the double-blind randomized controlled trial (RCT) in research on the effectiveness of medical care conducted in England see A. COCHRANE, EFFECTIVENESS AND EFFICIENCY (1972).
143 Social Security Act § 1156(a), 42 U.S.C. § 1320c-5(a) (Supp. III, 1973).

clude overutilization of hospitals or ancillary services, PSRO's will do nothing to discourage use of such ineffective but "typical" services and may even hinder the development of HMO's, whose physicians rely less on hospitalization in treating their patients. However, if research results are used continually to validate and revise the process criteria being employed, the PSRO system could potentially be oriented toward enhancing the positive health impact of medical care per dollar.

Although "effectiveness" research can begin to sort out beneficial procedures from those that are ineffective, there is also a need for critical evaluation of the efficiency or cost of current health care technology. Perhaps it would be sensible to employ the slightly monomaniacal funding scheme which has been so "successful" in some biomedical research projects and other major research and development efforts.[144] While many such projects have involved large expenditures for little benefit, in this case the cost would only be a tiny fraction of total health sector expenditures, and the information generated would have the potential of producing savings much greater than the amount invested.[145] Furthermore, this type of research is now severely underfunded. The purpose of the research would be to stimulate the development of new cost-minimizing technology (*e.g.*, new data systems) to compensate for the current emphasis on "spare no expense" medical gadgetry which often is valuable only to a few patients who happen to have exotic diseases. Another fruitful area of inquiry would be the evaluation of such regulatory schemes as restrictions on facility construction to determine if they are cost effective methods of achieving their nominal aims.

An important objective of such research activities should be to

---

144 *E.g.*, F. Chu, While You're Up, Get Me a Grant: A Study of Resource Allocation at the National Institute of Mental Health 109-19, Apr. 2, 1971 (unpublished undergraduate honors thesis in Harvard University Archives). This chapter of the paper is a case study of suicide prevention centers which were shown not to prevent suicides.

145 It would hardly be extravagant, for example, for the federal government to spend one tenth of one percent of total national health care expenditures (which would be in excess of $100 million annually) on research, if the information or technology generated has the potential of reducing that expenditure by even as little as one percent (which would save $10 for each dollar spent) without impairing anyone's health.

publicize the findings widely so that consumers can begin to base health care choices on more complete and reliable information. Such publicity might begin to reverse the prevailing "more-is-better" view of health care and incidentally to promote the growth of HMO's by justifying their relatively lower rate of hospitalization and other cost effectiveness measures. It also may enable consumers to make more intelligent choices about quality.

2. Use of Results

The American medical profession has always zealously guarded its independence from control by non-physician decision-makers. Prevailing patterns of physician behavior and practice have thus often proved highly resistant to change. If the empirical data from the "effectiveness" studies reveal a need for significant changes in the manner in which patients are treated, will it be possible to implement a new set of norms for medical practice without alienating the medical community? To avoid the appearance of unjustified imposition on the judgment of individual doctors, it would seem to be advisable, at least as an initial strategy, to attempt to induce voluntary modification of physician behavior from within the profession. The logical vehicle would be the PSRO system. The possibility that supplying information from research studies to all physicians and particularly to those involved in the peer review process would itself begin to change modes of practice should not be ruled out.[146] If in fact the research findings are used periodically to revise the PSRO standards, peer pressure and official regulatory sanctions will begin to play a role. It is at least conceivable that fundamental but gradual change in the practice of medicine could be legitimized in this fashion. Existing peer review data systems provide their committees with very little information on outcomes.[147] When the studies begin to generate out-

146 Jessee, Munier, Fielding & Goran, *supra* note 122.

147 The Department of Health, Education, and Welfare has financially supported the development of several regional hospital discharge data systems which agreed to collect 14 items of information known as the Uniform Hospital Abstract Data Set selected by the U.S. National Committee on Vital and Health Statistics. The primary purpose of such data systems is to provide information for medical peer review. The only item which in any way relates to outcomes is "disposition of patient," which includes whether the patient lived or died in the hospital. L.

come measures of the efficacy procedures, it will be difficult for the participating doctors to ignore results which suggest the ineffectiveness of a procedure they commonly use.

The behavior of the HMO or hospital manager is likely to be responsive more to economic incentives and less to professional norms than is true for physicians. However, it may not be easy for administrators to encourage the medical staff to make more cost effective utilization decisions without appearing to trench on physician prerogatives. The administration could be a useful conduit for assuring that research findings are fully presented to the medical staff. It also should not be difficult to induce health managers to introduce more efficient, cost-cutting technology into the hospital once competitive market incentives begin to influence their decision-making.

If market incentives and new information do succeed in lowering hospital occupancy rates, then effective programs to close or convert the underutilized beds are needed in order to realize savings. In an HMO, surplus capacity would be sold or put on standby or would provoke aggressive marketing efforts (perhaps including price cutting) for a period of time. On the other hand, in the cost-reimbursement regulatory context difficult political problems must be faced. As community institutions, hospitals tend to generate supportive constituencies which can create formidable obstacles to closing them down. In Massachusetts two hospitals that were denied certificates of need by the state Health Department persuaded the state legislature to pass legislation which mandated the Health Department to approve their projects,[148] and the legislation was sustained over a subsequent judicial challenge.[149] Thus, in the absence of HMO's, the effectiveness of efforts to regulate facility construction will depend upon the sophistication of state political

---

Kirsch, I. Altman, T. Frazier, J. Kavet & J. Mannis, PSRO Information and Consumer Choice: The Case for Public Disclosure of Health Services Data 156-57, 172, Feb. 1975 (available from the Harvard University Center for Community Health and Medical Care).

148 Reider, Mason & Glantz, *Certificate of Need: The Massachusetts Experience*, 1 AM. J.L. & MED. 13, 32 & n.82, 39 n.122 (1975).

149 Commissioner of Pub. Health v. Bessie M. Burke Memorial Hosp., Mass. 1975 Adv. Sh. 253, 323 N.E.2d 309 (1975).

systems and upon public education as to the costs associated with unnecessary facilities. The alternative would be a federal certificate of need process, which might be less subject to local pressures but also more rigid and indifferent to local problems.

C. *Health Manpower Policy*

As discussed above,[150] health manpower is seriously maldistributed not only geographically, but also with respect to medical specialty and level of skill (*i.e.*, physician v. non-physician). Both of the latter types of maldistribution contribute to the pressures for the delivery of marginally effective services or to inefficient delivery. A more rational, economical deployment of health manpower is therefore in order.

It was noted above that the amount of surgery relative to population performed in an area is positively correlated with the density of surgeons practicing in that area[151] and that where there is a relatively high number of surgeons relative to population, they tend to have higher net income per operation and a lower average workload per doctor.[152] Thus, it seems likely that limiting the production of surgeons and other specialists and thus increasing the average workload would have the effect of decreasing the financial pressures to perform surgical procedures of marginal effectiveness as well as reducing the tendency to maintain relatively higher fee schedules. Fewer surgeons should also result in fewer total operations and therefore fewer surgical deaths and injuries.

The two bills which may result in new health manpower legislation by the end of this Congress both would establish controls over the number of postgraduate training programs in each specialty in each region.[153] The rationale for such controls is that hospitals and their medical staffs have substantial incentives unrelated to health manpower needs to establish and expand residency programs, particularly those for specialists, because they provide cheap (to the staff doctor, free) labor to assist in managing

150 See text accompanying notes 59-78 *supra.*
151 See text accompanying notes 24, 103-06 *supra.*
152 See text accompanying notes 107-08 *supra.*
153 See text accompanying notes 64-67 *supra.*

hospitalized patients and they enhance the reputation and prestige of the hospital.[154]

Surely such direct regulation can achieve the desired effect of training fewer surgeons and other specialists and more primary care practitioners; however, there are numerous important questions regarding the detailed organization of such a regulatory system which are beyond the scope of this article. In the long run, it would be advisable to evaluate the content of a system of medical education which seems to channel the majority of students into research and the most sophisticated specialities and away from primary practice.[155] It also should be understood that controlling distribution of physicians by specialty alone is no assurance that there will be a substantial shift of medical manpower toward underserved areas, since it is not clear that the extent of geographic maldistribution is substantially greater for specialists than for primary care doctors.[156]

Federal support for the increased training and use of paramedical providers[157] seems well-advised, but the potential for the delivery system to achieve significant efficiency savings from the use of non-physician providers is still largely unrealized. One of the important tasks of the research efforts discussed above[158] should be to determine the extent to which physician's assistants and nurse practitioners can be employed for routine diagnosis and treatment and referral of complicated cases to a physician without unduly sacrificing quality.[159] Also, obstacles to the full utilization of paraprofessional providers such as restrictive state laws[160] must

---

154 H.R. REP., *supra* note 13, at 41.

155 For statistics on distribution of physicians by specialty see *id.* at 38-43.

156 While there is little or no geographic maldistribution among physicians in general practice, other specialties often counted as primary care, such as internal medicine, pediatrics, and obstetrics-gynecology, are severely maldistributed by geographic area. *See* J. Winsten, *supra* note 13, at 46-47.

157 See text accompanying notes 73-75 *supra.*

158 See text accompanying notes 141-46 *supra.*

159 For an overview of such possibilities see Lincoln Laboratory, Massachusetts Institute of Technology & Beth Israel Hospital, Harvard Medical School, Ambulatory Care Project, Final Contract Report 14A, Feb. 29, 1976, and sources cited therein at 59-64 (available from Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Mass.).

160 For a recent, comprehensive, well-documented study of state legislation and regulation of physician assistants and nurse practitioners see Kissam, *supra* note 11. Kissam documents the need for new legislation, *id.* at 11-13, catalogues recent state

be removed, perhaps by conditioning federal support on state level reform. To the extent the delivery system is converted to prepaid financing and efficiency becomes more important to health managers, the demand for such personnel should increase.

Greater utilization of physician's assistants and nurse practitioners and reduced utilization of medical specialists also will have a feed-back effect on the composition of medical providers being educated. There is already substantial federal funding for the training of paramedicals,[161] and the two bills now under consideration will have the effect of eventually reducing the number of surgeons and other specialists trained.[162] Again, it is unclear whether training more paramedicals will guarantee substantially more availability of medical services in shortage areas. They may respond to the same locational incentives which have caused the current geographic maldistribution of physicians.

Redistributing doctors toward rural and core urban areas will have only an indirect effect on the efficacy and efficiency of health care. Only to the extent that there is a corresponding decrease of physicians in suburban areas would there be any reduction in the delivery of marginally effective services or doctors' fees in such overserved areas.[163] Discussion of strategies for correcting the geographic maldistribution problem is beyond the scope of this article. However, it would appear that the solution lies in the direction of fuller utilization of paramedical personnel supervised by physicians and increased institutional and professional support for physicians in isolated areas rather than simply increased financial incentives directed at medical students.[164]

---

legislative action, *id.* at 20-29, and concludes that most of the new laws are unduly restrictive because they are administered by medical or nursing licensing boards, which tend to protect the professional interests of nurses and physicians by limiting the tasks which can be performed by non-nurse paramedicals. *Id.* at 52-65.

161 See text accompanying notes 73-75 *supra.*

162 Both bills would decrease the total number of residencies and increase the number of primary care residencies at the expense of residencies in surgery and the other more sophisticated specialties. H.R. 5546, 94th Cong., 1st Sess. § 801 (1975) (which would have added §§ 1701(a), 1703(a)-(b) to the Public Health Service Act); S. 3239, 94th Cong., 2d Sess. §§ 501, 802 (1976) (which would add §§ 798, 799(a), 771(b)(2), (7) to the Public Health Service Act).

163 See text accompanying notes 151-53 *supra.*

164 The National Health Service Corps holds promise for achieving such results, at least in rural areas. See note 61 *supra.*

Conclusion

Too often in the historical development of the American health care system, the system's responses to perceived problems have created more difficulties than they have solved. Thus the medical profession convinced the public that, because the citizenry was being deceived and endangered by incompetent practitioners, physicians should be given legal authority to restrict entry into the field to those who were imbued with the training and the commitment to deliver medical care of the highest possible quality. When it became apparent that such "quality" was too expensive for the poor, the answer was a societal commitment, motivated by the best of purposes, to make quality care available to all. Finally, it has been realized that the society cannot afford to pour unlimited resources into the seemingly insatiable health care machine. In response, some in government have publicly decried the spiraling costs and demanded that providers stop charging so much. Others have urged or undertaken well-publicized campaigns against fraud and cheating or threatened providers with tighter controls. Finally, the government has implemented programs which typically deal with only the symptoms of an open-ended financing system and its associated counter-productive incentives.

At the same time, both research studies and a malpractice crisis seem to indicate that quality in fact is less than optimal, certain categories of people and isolated geographical areas are desperately underserved while others are glutted with providers, and the producers' complex of incentives and traditional practices continues to accelerate the rate of cost increase and raise the proportion of GNP devoted to health care.

At last there is an awareness that more health care is not always a benefit, or at least not so beneficial as to outweigh the cost. The issue now is whether Congress can fashion responses which solve the cost, access, and quality problems without creating still others. The authors have argued that several programs ought to be considered before the final step on the access ladder — national health insurance — is taken. In a fee-for-service, cost reimbursement system, the effect of such a financing scheme could be to so exacerbate the cost problem that political pressures will mount to cut back on benefits or increase beneficiary contributions even as financing is being extended to a broader range of citizens. Thus, it is not clear that those most in need — the sick and the poor — will be better off. And more financing also fails to address the issues of geographic maldistribution and ineffectiveness.

The Congress must recognize that previous health programs have failed largely because they ignored the financial and personal incentives, such as income, status, influence, and conformance with norms, which motivate health professionals. Short of direct coercion, the government has little power, at least in the short run, to change the way physicians and hospital administrators behave. Instead, incentives must be re-directed so that they encourage the deliverers of care to place primary emphasis on improving the health of their patients at a minimum cost. Aggressively promoting a market approach which encourages the development of competing HMO's offers the best hope of achieving this goal.

However, it is also clear that the government must deal with several existing impediments to a better functioning market. First, both providers and consumers are still quite ignorant about which health services are actually beneficial to the patient's health, and to what extent. Second, consumers lack access to much of the already available evaluative data about the performance of providers which is essential to informed choices about their health care. Third, the current maldistribution of income (and health resources) in the country must be at least partially corrected before health services can be available to citizens on a less arbitrary, more socially acceptable basis than affluence or location.

The government must sponsor research and development of the more effective and efficient modes of delivery for which there would be a market in a reorganized system. However, reformers must also confront the problem of how to alter the professional behavior of the key decisionmaker in the system — the physician. His commitment to doing everything he can for his patient is in many ways appropriate and admirable. However, it is clear that society does not benefit from expending scarce resources on health services which have little or no demonstrable positive effect on the patient's health. Giving managers, physicians, and consumers financial incentives to economize and generating information about the effectiveness of the procedures doctors employ could have an impact on physician behavior. However, reformers must also examine the extent to which existing patterns of medical education create attitudes which inhibit acceptance by physicians of new technology and modes of treatment.

But even after the utilization of ineffective services is reduced or eliminated, the hospital manager must find ways to deliver the remaining services more efficiently. Otherwise, there will be a tendency for the savings from decreasing utilization to be absorbed by higher cost services. The hospital manager must decrease the level of inputs for each bed or each service, which can be accomplished by reducing both capital and labor, by decreasing one without increasing the other, or by substituting a small amount of capital for a greater amount of labor. Given the labor-intensive nature of health care,[165] substantial efficiencies in the short run can probably best be achieved by increasing the productivity of health workers, *e.g.*, by substituting cheap labor (paramedicals) for more expensive labor (physicians) in the performance of particular tasks.

Until such actions are initiated, the government should not attempt to solve the distributional problem through a national health insurance financing scheme which perpetuates the retrospective cost-reimbursement fee-for-service system and all the counterproductive incentives it entails. In short, society must stop rewarding health providers for using more scarce resources without any assurance that the goal of better health is being attained.

---

165 Iglehart, *supra* note 20, at 1322.

THE DIVESTITURE REMEDY IN SHERMAN ACT § 2 CASES

KEVIN J. O'CONNOR*

In recent years, Sherman Act § 2 has enjoyed a resurgence in the form of several monopolization actions brought by the Justice Department, the Federal Trade Commission, and private plaintiffs against some of the nation's leading firms including IBM, ATT, Xerox, Kodak, and the largest members of the cereal and tire manufacturing industries. Mr. O'Connor argues that the crucial issues in the successful application of the § 2 monopolization charge are not so much the liability questions but rather the guidelines for formulation and implementation of an effective remedy, once liability is found to exist. Noting the federal district courts' reluctance to grant significant divestiture over the past century despite their broad equity powers under the Sherman Act, Mr. O'Connor argues that the conduct-prohibitory and regulatory remedies upon which the courts have relied heavily, tend to be less effective and more costly both to the general public and to those with a direct financial interest in the target firms than direct structural remedies. In short, Mr. O'Connor argues that since divestiture has worked in the past and since the current industrial organization learning suggests that market structure conditions are the ultimate determinant of market performance, divestiture ought to be the presumed remedy in § 2 cases.

Table of Contents

Introduction 689

I. *The Need for More Vigorous Use of the Divestiture Remedy* 695

A. *Divestiture as Currently Applied* 695

*B.A., University of Wisconsin, 1972; J.D., Harvard, 1976. The author is grateful to Professor Stephen Breyer, Harvard Law School, and to Mr. J. David Wade, member of the Class of 1976 at Harvard Law School for their helpful comments and criticisms. The author is especially indebted to Mr. James C. Burling, member of the Class of 1976 at Harvard Law School, for his substantial contribution to the final version of this manuscript.

B. *The Call for Broadened Interpretation of Existing Monopolization Liability Standards and New Deconcentration Legislation* 698
C. *The Relation Between Liability Standards and Remedy Guidelines* 704

II. *The History and Use of the Divestiture Power* 707
A. *The Divestiture Power and the Courts' Reluctance to Utilize It* 707
B. *The Reasons for Judicial Timidity in Remedy Selection* 714

III. *Perspectives on Remedy Choice and Construction for Sherman Act Monopolization Violations* 717
A. *The Populist Perspective Versus the Economic Efficiency Perspective* 717
B. *The Economic Efficiency Approach: Structure, Conduct, Performance* 719
1. *The Structuralist School* 721
2. *The Conduct/Performance School* 726
3. *Market Conduct as a Determinant of Market Structure: A Hybrid of the Conduct Approach* 729

IV. *Comparative Analysis of Structural and Conduct Remedies* 730
A. *Effectiveness* 731
1. *Structural Remedies* 731
2. *Conduct Remedies* 731
B. *Technical Efficiencies* 734
C. *Effects on Third Parties* 741
D. *Judicial Supervision* 749

V. *Current Law and Its Implications for the Divestiture Remedy* 752
A. *Substantial Violation* 752
B. *Proximate Cause* 755
C. *Necessity* 762
D. *Practicality* 769

Conclusion 772

Introduction

Concentration[1] in the nation's leading nonfinancial manufacturing sectors has been facilitated by wide gaps in the scope of the antitrust laws[2] and the unavailability of effective remedies where violations of the current laws have been found to exist.[3] Such concentration, whether in the form of an oligopolized or dominant firm industry, both indicates the possession and facilitates the development of market power by the constituent firms, and tends to lead ultimately to undesirable conduct and performance results.[4]

---

1 The degree of seller concentration in an industry refers to the extent to which industry sales (or physical output, employment, value added, or assets) are contributed by a few firms in the industry. F. SCHERER, INDUSTRIAL MARKET STRUCTURE AND ECONOMIC PERFORMANCE 50-51 (1970). The most common measure of seller concentration, and the most operational measure in terms of industrial organization theory, is the four- or eight-firm concentration ratio. *Id.* This ratio is the sum of the market shares in terms of percentages of the total market sales of the four or eight firms with the greatest market shares in the given industry. This ratio suffers from several defects including, for example, the fact that it does not incorporate the skewness of the distribution of market shares throughout the industry nor does it include the total number of sellers in the market into its index. However, it has generally been found to be more operational than the alternative measures of market power such as the Lerner index, the Gini coefficient, and the Herfindahl index. *Id.* at 50-57.

2 *See* White House Task Force Report on Antitrust Policy (July 5, 1968), reprinted in BNA ANTITRUST & TRADE REG. REP., No. 411, Special Supplement Part II, May 27, 1969, and in 2 ANTITRUST L. & ECON. REV. 11 (Winter 1968-1969) (hereinafter the Neal Report, paginated to the ANTITRUST L. & ECON. REV. reprint). See also note 37 *infra.*

3 See text accompanying notes 32-40, *infra.*

4 The concept of market power is developed more fully at text accompanying notes 129-30 *infra.* In general, market power is characterized by the existence of high concentration and barriers to entry in an industry which enable individual firms to exert some degree of control over prices. *See* F. SCHERER, *supra* note 1, at 10-11. Market power is associated with certain structural patterns, modes of conduct or certain performance characteristics. Market structure elements which tend to indicate, although often not conclusively, the existence of market power include: fewness of sellers and buyers; the possession of high market shares; the assymmetry of seller and buyer sizes; substantial product differentiation; moderate to high barriers to entry; absolutely large size; similarity of cost structures among a few competitors (thus making collusion easier); significant vertical integration; and substantial levels of diversification. *See* F. SCHERER, *id.* at 5. There is a significant interrelation among the elements listed. For example, where the number of sellers in a market is few, they will tend to have high market shares. But, where a few firms in a market have high market shares, there may be a substantial number of firms composing a competitive fringe. In other words, each element listed adds something to the explanatory power of market structure or power on market conduct and hence market performance. However, the most significant explanatory elements of market structure in terms of their effect on market performance have

Frequently-observed phenomena, including planned obsolescence of automobiles by firms in the oligopolistic automobile production industry[5] and restrictive marketing practices of dominant firms such as International Business Machines in the integrated computer systems market[6] or Kodak in the market for the production of cameras,[7] readily testify to this concentrated structure-suboptimal conduct and performance connection.

Prevailing economic theory and a substantial quantity of empirical research lend rigorous support to such popular intuitions by asserting that significant causal links run from market structure to market performance via market conduct.[8] Some have argued that high concentration is desirable because of the alleged economies of scale or superiority in innovation.[9] However, those

---

been found to be number of sellers and market share. *See* note 120 and accompanying text *infra.*

Market conduct elements include: pricing behavior; product strategy; research and innovation; advertising; and legal tactics. *See* F. SCHERER, *supra* note 1, at 5. Specific types of pricing behavior include outright price-fixing by competitors, *id.* at 434-38, tacit collusion by oligopolists, *id.* at 135-36, 179-82, and price leadership by dominant firms, *id.* at 164-66, 216-19.

Market performance elements include: production and allocative efficiency; technical efficiency; progress; full employment; and equity. *Id.* at 5. Allocative efficiency, generally speaking, is the satisfaction of consumer wants with maximum effectiveness. *Id.* at 15. Monopoly power can lead to an increase in price and restriction in the quantity produced of a commodity away from competitive (and, hence, optimal) levels. *Id.* at 13-19. The existence of above normal profits in an industry is often taken as evidence of allocative inefficiency — although, this measure includes resources that are redistributed from consumers to sellers as well as unadulterated resources losses. *Id.* Technical inefficiency is distinguishable from allocative inefficiency in that it is a label for the resources lost due to the operation of an economic enterprise at greater cost than the minimum attainable and thus does not concern directly the market misallocation of resources due to allocative inefficiency. *See* Leibenstein, *Allocative Efficiency vs. "x-Efficiency,"* AMER. ECON. REV. 392 (June 1966). In fact technical inefficiency may actually eat into the excess profits that are taken to be a measure of allocative inefficiency. *See* F. SCHERER, *supra* note 1, at 403-04.

5 *See generally* L. WHITE, THE AUTOMOBILE INDUSTRY SINCE 1945 ch. 8, 11 (1971).

6 *See generally* G. BROCK, THE U.S. COMPUTER INDUSTRY, 1954-1973: A STUDY OF MARKET POWER ch. 10 (1975).

7 *See generally* Berkey Photo Co. v. Eastman Kodak Co., Civil No. 73-424 (S.D.N.Y., filed Jan. 29, 1973). Among other complaints, Kodak is accused of illegally excluding its competitors from joint research and development of the flash-cube and magic flash-bar with Sylvania. See Memorandum of Law, prepared by author on this question (on file at the HARVARD JOURNAL ON LEGISLATION).

8 See text accompanying notes 120-49 *infra.*

9 *See, e.g.,* McGee, *Efficiency and Economies of Size,* in INDUSTRIAL CONCENTRATION: THE NEW LEARNING 55-97, 101-04 (H. GOLDSCHMID, H. MANN, J. WESTON eds. 1974) [hereinafter GOLDSCHMID]. *But see* notes 193-227 *infra.*

For a general discussion of economies of scale, *see* F. SCHERER, *supra* note 1, at arguments are losing considerable force as industrial organization economists increasingly associate highly-concentrated industry structures with restrictive market conduct and the resultant poor market performance.[10]

The existence of these links has prompted calls for a more vigorous deconcentration policy either within the framework of the present antitrust liability standards[11] or through new legislation.[12] The judicially imposed requirement that anticompetitive conduct[13] accompany substantial market power in order to establish

---

72-103. The presumed realization of substantial economies by large concerns has been a major factor behind the judicial reluctance to adopt remedial measures directed at structural reform. But, considering plant level, firm level, and product specific economies of scale, the available data does not support the proposition that scale economies justify high concentration in most American industries. *See, e.g.*, Scherer, *Economies of Scale and Industrial Concentration*, in GOLDSCHMID, *supra*, at 16-54; W. MUELLER, A PRIMER ON MONOPOLY AND COMPETITION 49 (1970); G. STIGLER, THE THEORY OF PRICE 223 (3d ed. 1966); Esposito & Esposito, *Dissolution and Scale Economies*, 4 ANTITRUST L. & ECON. REV. 77 (Summer 1971).

Several authors have argued that firms which possess the attributes of large size and market power are better able to innovate and implement inventions than their smaller competitors. *See, e.g.*, J. GALBRAITH, THE NEW INDUSTRIAL STATE (1971); J. SCHUMPETER, CAPITALISM, SOCIALISM, DEMOCRACY 88, 102-03 (3d ed. 1950). *But see* F. SCHERER, *supra* note 1, at 372-76 (summarizing the ambiguous quantitative evidence on the effect of market power on innovativeness).

The courts have proved receptive to arguments based on technical efficiency put forth by defendants in Sherman Act § 2 cases, and have generally refused to find against firms able to establish that their market power grew solely out of "superior skill, foresight, and industry." *See* United States v. Aluminum Co. of America, 148 F.2d 416, 429-30 (2d Cir. 1945); United States v. Grinnell Corp., 384 U.S. 563, 570-71 (1966).

10 See text accompanying notes 120-51 *infra*.

11 See text accompanying notes 41-42 *infra*.

12 See text accompanying notes 45-48 *infra*.

13 *See, e.g.*, United States v. United States Steel Corp., 251 U.S. 417, 451 (1920); United States v. Swift & Co., 286 U.S. 106, 116 (1932) (mere size is not an offense). Many commentators have argued that the courts are, in fact, preoccupied with the conduct element of antitrust charges in general, *see* Irwin & Barrett, *Antitrust Enforcement in the United States: Market Structure Versus Market Conduct*, WASH. L.Q. 37, 46-47 (1974). The Antitrust Division of the Department of Justice (AT) as well as the Federal Trade Commission (FTC) are in part responsible for this emphasis on conduct. Shepherd estimates that the cases brought by the enforcement arms break down as follows:

FTC	AT	
60%	40%	conduct related.
20	20	mergers.
5	15	restorative treatment of regulated industry.
15	10	retroactive action in industry and trade.
100%	85%	total.
	15%	miscellaneous.

W. SHEPHERD, THE TREATMENT OF MARKET POWER 143, 145 (1975) *citing* Posner, a § 2 monopolizing violation under the Sherman Act[14] has left many non-competitive industries untouched by the antitrust laws.[15] For example courts have been reluctant to attack under the Sherman Act either oligopolistic market structures[16] or single firm dominance[17] obtained via internal expansion. As a result, several legislative proposals have been offered which call for a revised liability standard consisting solely of specific market-share trip-wires.[18]

The primary deficiency of current antitrust policy, however, does not result from its narrow and cumbersome standard of liability, but from the inability or unwillingness of the courts to perceive and apply an effective remedy once a violation is found to exist. Often when the courts have found the market power and abusive conduct required for a Sherman Act § 2 monopolization violation, they have unnecessarily restricted the remedy to relatively ineffective and unwieldy injunctions prohibiting certain

---

A Statistical Study of Antitrust Enforcement, J.L. & ECON. 365 (1970). There is considerable overlap between the two agencies although the FTC generally deals with consumer related industries whereas the Antitrust Division deals with heavy industries. However, friction does arise as in the *American Telegraph & Telephone* case where the FTC asserts that it has exclusive jurisdiction over 28 of the 30 conduct modes challenged by the Justice Department's complaint. BNA ANTITRUST & TRADE REG. REP., No. 746, A-14 (Jan. 6, 1976). Justice Department's complaint is at *id.*, No. 690, D-1. *See also* J. BAIN, INDUSTRIAL ORGANIZATION 562 (2d ed. 1968). *See* note 116 *infra.*

14 Section 2 of the Sherman Act reads as follows:

> Every person who shall *monopolize,* or *attempt to monopolize,* or *combine or conspire with any other person or persons, to monopolize* any part of the trade or commerce among the several States, or with foreign nations, shall be deemed guilty of a misdemeanor, and, on conviction thereof, shall be punished by fine not exceeding fifty thousand dollars, or by imprisonment not exceeding one year, or by both said punishments, in the discretion of the court (emphasis added).

15 U.S.C. § 2 (1970).

For background reading on the question of judicial treatment of suits brought under Sherman Act § 2, 15 U.S.C. § 2 (1970), see A. KAHN & J. DIRLAM, FAIR COMPETITION: THE LAW & ECONOMICS OF ANTITRUST POLICY (1954); A. NEALE, THE ANTITRUST LAWS OF THE UNITED STATES OF AMERICA 95-184 (1960); Brodner, *Monopolization and Attempts to Monopolize: Whatever Happened to Section 2?* 41 ABA ANTITRUST L.J. 589, 591 (1972); Cox, *Competition & Section 2 of the Sherman Act,* 27 ABA ANTITRUST L.J. 72, 76 (1965); Fortas, *Part II: Portents for New Antitrust Policy,* 10 ANTITRUST BULL. 41 (1965).

15 See examples cited in notes 6, 7, 8 *supra.*

16 See note 41 *infra.*

17 See note 42 *infra.*

18 See text accompanying notes 45-48 *infra.*

types of conduct.[19] The courts generally have been reluctant to grant significant structural relief such as divestiture.[20]

Above all else, this article seeks to emphasize that such structural relief is usually the most appropriate type of remedy under current Sherman Act § 2 standards, as well as under recent proposals for antitrust reform which would explicitly "extend" the liability standard to focus exclusively on structure. That is, whatever the resolution of the current debate as between the present dual structure-conduct standard and the proposed purely structural liability trigger, divestiture usually will be the proper form of relief, and its effective implementation the *sine qua non* of successful § 2 enforcement.

The current judicial myopia in the selection of Sherman Act § 2 remedies stems from three fundamental misperceptions. First, the courts often have attempted to denude market power of its deleterious effects on performance by enjoining the particular type of conduct which formed the basis for the "abusive conduct" predicate for § 2 liability.[21] The conduct-oriented approach generally derives from the mistaken belief that enjoining a particular mode of conduct is a more effective tool for remedying non-optimal performance than altering structural determinants. It ignores the reality that unsatisfactory market performance can result from any one of a wide variety of particular conduct modes and that a wide variety of conduct modes can result from a given market structure.[22] Thus conduct injunctions, by dealing with the symptoms rather than the underlying ailment, often have left in-

---

19 See text accompanying notes 65-87 *infra*.

20 See text accompanying notes 65-87 *infra*.

> Although distinctions may be made between the remedies of dissolution, divorcement, and divestiture, the term divestiture is broad enough to cover the other two. Divestiture refers to divesting a defendant of property, securities, or other assets; divorcement applies to the effect of a decree ordering particular types of divestiture and is especially applicable to vertically integrated organizations; dissolution refers to any situation where the dissolving of an illegal combination is involved, including the dissolution of such combinations by divestiture and divorcement.

Note, *Availability of Divestiture in Private Litigation as a Remedy for Violation of Section 7 of the Clayton Act*, 49 MINN. L. REV. 267, 270 n.21 (1964) *citing* S. OPPENHEIM, CASES ON FEDERAL ANTI-TRUST LAWS 885 (1948).

21 See text accompanying notes 89-91 *infra*.

22 See text accompanying notes 121-61 *infra*.

tact a market structure which continues to dictate undesirable performance results.[23]

Second, the courts have viewed divestiture as particularly "harsh" in terms of real economic costs as well as the anticipated effect on those with a direct financial interest in the defendant enterprise.[24] However, with respect to both types of costs, divestiture can often be less burdensome than the conduct remedies now granted so freely.[25]

Finally, the judiciary has taken an unnecessarily restrictive view of the § 2 precedent on remedy formulation. Of the several judicially enunciated guidelines for divestiture application, two are simply ill-conceived,[26] while the remaining two, if sensibly interpreted, would suggest that there is a more vigorous role for divestiture to play in monopolization cases.[27]

In short, this article contends that the use of divestiture as the presumed remedy under Sherman Act § 2 not only is compelled by the dictates of economic learning but finds significant support in existing judicial guidelines. If the defendant can affirmatively prove that the real economic costs of divestiture outweigh the expected economic benefit of divestiture, other structurally-oriented remedies ought to be considered.[28] The third-party costs of divesti-

---

23 An assessment of the relation between power [structure] and practices [conduct] is unavoidable in remedy proceedings. The choice between an injunction or a "cease and desist" order as against some form of divestiture depends on the anticipated consequences for competition of a prohibition of particular practices. And this, in turn, depends largely on the power of the relevant firms or the structure of the market. If the power is significantly great, injunction may be without effect. E. Mason, *Preface* to C. KAYSEN & D. TURNER, ANTITRUST POLICY: AN ECONOMIC AND LEGAL ANALYSIS xvi (1959).

24 The costs to these groups are, in a sense, costs to a segment of the "public." However, it will be shown *infra* at notes 212-39 that such factors should not enter into the initial calculations as to whether divestiture or some other approach represents the most effective remedy in monopolization cases in terms of increasing total consumer welfare. Rather, it will be argued, such costs ought to be the deciding factor only as among equally effective remedies.

25 See text accompanying notes 190-210 *infra*.

26 See text accompanying notes 251-73 *infra* concerning the "clear violation" and "proximate cause" criteria, respectively.

27 See text accompanying notes 299-349 *infra* concerning the "necessity" and "practicality" requirements, respectively.

28 *See, e.g.*, Baldwin, *The Feedback Effect of Business Conduct on Structure*, 12 J.L. & ECON. 169 (1969).

ture — the costs to those with a direct financial interest in the defendant divestee — are not justifiably included in the real economic costs of divestiture and thus ought not to be included in the balancing test.[29] However, once a remedy has been ordered, these transaction costs ought to be minimized within the framework of an effectively formulated remedy.

Although the courts could adopt such remedy guidelines under current law,[30] they are unlikely to do so. Thus, these standards ought to be incorporated in the various legislative deconcentration proposals.[31]

The first section of this article examines the need for a closer look at remedy selection procedures, both under current liability standards and under the pending deconcentration proposals. Next follows an exploration of the deficiencies of the conduct approach, together with a proffered explanation of the prevalence of this orientation in remedy selection. Then, the economic basis for more vigorous use of the divestiture remedy is presented. Finally, the article probes the scope of the federal judiciary's power to grant divestiture against the background of the explicit approaches hitherto taken by the courts in cases where divestiture was an issue.

I. The Need for More Vigorous Use of the Divestiture Remedy

A. *Divestiture as Currently Applied*

One of the parodoxes of current antitrust policy is that while those firms which are most commonly perceived as monopolies or

---

29 See text accompanying notes 212-39 *infra*.

30 See text accompanying notes 65-76 *infra*. Even before it was recognized that structural remedies were effective in attacking the conduct element of the "two-prong" Sherman Act § 2 liability test developed by the judiciary, the courts did not ignore totally the structural element of the offenses and the possibility of direct structural remedy. *See, e.g.*, the dicta of Judge Learned Hand in United States v. Aluminum Co. of America, 148 F.2d 416 (2d Cir. 1945) ("great industrial consolidations are inherently undesirable, regardless of the economic results").

31 See text accompanying notes 49-52 *infra*.

near-monopolies remain unmolested under current laws because they achieved their market shares through internal expansion, firms which may have considerably smaller market shares are prosecuted vigorously for their anticompetitive actions in procuring those shares through mergers. This curiosity results from a lower threshold of liability in the anti-merger provision, Clayton Act § 7,[32] than in the judicially derived two-prong test under § 2 of the Sherman Act.[33] Thus, firms of equal market share are differentially liable under the antitrust laws according to the route chosen to reach their current market position. This irony is accentuated by development of the "backward sweep doctrine" in Clayton Act § 7 cases, under which a merger can be prosecuted many years after its consummation as its anticompetitive effects become apparent.[34] However, even though § 7 covers both "new" and "old" consummated mergers, the prosecution of new mergers predominates.[35]

Consequently, while the use of divestiture has increased greatly since the mid-1950's, the upsurge is due almost solely to serious but selective enforcement of the anti-merger provisions of the

---

32 Section 7 of the Clayton Act (15 U.S.C. § 18, 1970) provides that:

> No corporation engaged in commerce shall acquire, directly or indirectly, the whole or any part of the stock or other share capital and no corporation subject to the jurisdiction of the Federal Trade Commission shall acquire the whole or any part of the assets of another corporation engaged also in commerce, where in any line of commerce in any section of the country, the effect of such acquisition may be substantially to lessen competition, or to tend to create a monopoly.

For examples of the vigorous judicial enforcement of this section, as contrasted with the more cautious approach taken to § 2 Sherman Act cases, *see, e.g.*, United States v. Brown Shoe Co., 370 U.S. 294 (1962); United States v. Philadelphia National Bank, 374 U.S. 321 (1963).

33 See text accompanying notes 13-17 *supra*.

34 Under the so-called "backward sweep doctrine" some courts have, in passing on the validity of a particular merger, looked to market facts existing at the time the merger occurred and not merely to conditions prevailing when suit is brought. *See* United States v. E. I. du Pont de Nemours, 353 U.S. 586, 588 (1957), *remedy granted*, 177 F. Supp. 1 (N.D. Ill. 1959), *remedy modified*, 366 U.S. 316, 328-31 (1961) (merger attacked in 1948 occurred in 1917-1919; passage of time alone does not immunize merger); United States v. Greater Buffalo Press, Inc., 402 U.S. 549, 556 (1971) (passage of time no barrier per se to divestiture); United States v. Jerrold Electronics Corp., 187 F. Supp. 545 (E.D. Pa. 1960) (backward sweep doctrine applied to invalidate mergers which had taken place three years before merger). *See also* Ward, *A Current Perspective on Enforcement of Laws Against Monopolization*, 41 A.B.A. ANTITRUST L.J., 601, 604 (1972); Bork, *Anticompetitive Enforcement Doctrines Under Section 7 of the Clayton Act*, 39 TEXAS L. REV. 832 (1961).

35 *See generally* M. GREEN, THE CLOSED ENTERPRISE SYSTEM (1972).

Clayton Act § 7.[36] Long-standing dominant firms, then, whether they have achieved their market positions through internal growth or past merger, enjoy a privileged position under the antitrust laws as presently applied.[37]

The explanation for the attack of concentration at the "new merger" level under § 7 of the Clayton Act rather than under § 2 of the Sherman Act, comprehends not only the lower threshold for liability of the former, but also the greater *apparent* availability of a "natural," effective, and less severe structural remedy in such cases.[38] As a result, enforcement of Sherman Act § 2 generally has been infrequent and erratic during the last two or three decades,[39] although the disposition of several monopolization cases in the 1940's might suggest that the Sherman Act can be and has been used more vigorously.[40]

---

36 *See* Comment, *Private Divestiture: Antitrust's Latest Problem Child,* 41 FORDHAM L. REV. 569, 580, 586 (1973); *Cf.* Bork, *supra* note 34, at 832.

37 Of course, challenges to the possession and pernicious use of dominant market power acquired through internal expansion must still be brought under the Sherman Act. *See generally* the Clayton Act's § 7 proscription that ". . . no corporation . . . shall *acquire* . . . another corporation." 15 U.S.C. § 18 (1970), *as amended,* 64 Stat. 1125 (1950) (emphasis added). Growth via internal expansion has generally been viewed more favorably than growth through mergers. *See* United States v. Ford Motor Co., 286 F. Supp. 407, 441 (E.D. Mich. 1968), *aff'd,* 405 U.S. 562, 568 (1972); ANTITRUST LAW DEVELOPMENTS, SECTION OF THE ABA 47 (1975); Keyes, *The Proposed Concentrated Industries Act,* 15 ANTITRUST BULL. 469, 484-85 (1970). Keyes notes that:

> . . . the difference [between internal expansion and merger cases] lies in the fact that the enlargement of a firm's share of sales by internal growth cannot in general be accomplished without outperforming a competitor, either by offering a superior product or a lower price or both. Hence it is highly unlikely that such an enlargement will represent a reduction in the number of alternative sources of supply open to buyers, and an abatement of competitive pressure on the enlarged firm, unaccompanied by any offsetting economic benefits. Where growth is by means of a merger, there appears to be no such presumption of benefit.

Id. at 485. *But see* Neal Report, *supra* note 2, at 25. The longer a merger goes unchallenged, that is, the greater the length of time between accomplishment of the merger and challenge in the courts, the more the merger becomes like an internal expansion case. *See generally* cases cited in note 82, *infra.*

38 *See, e.g.,* United States v. E. I. du Pont de Nemours & Co., 366 U.S. 316, 328-29 (1961) ("complete divestiture is peculiarly appropriate in . . . acquisitions which violate § 7. . . . The very words of § 7 suggest that an undoing of the acquisition is a natural remedy"). Ford Motor Co. v. United States, 405 U.S. 562, 573 (1972).

39 *See* W. SHEPHERD, *supra* note 13, at 195-96.

40 *See* C. KAYSEN & D. TURNER, *supra* note 23, at 21. Early restorative actions probably have reduced market power in certain industries such as the oil, cigarette, aluminum, and tin can industries. W. SHEPARD, *supra* note 13, at 183-215.

B. *The Call for Broadened Interpretation of Existing Monopolization Liability Standards and New Deconcentration Legislation*

With the growing perception that restrictive judicial interpretations of § 2 of the Sherman Act have rendered the provision ineffectual, calls for a more liberal reading have come from several quarters. Some commentators believe that the existing Sherman Act § 2 can be extended to include cases of persistent market power[41] and tacit collusion among oligopolists.[42] Both of these

---

41 *See, e.g.*, Williamson, *Dominant Firms and the Monopoly Problem: Market Failure Considerations*, 85 HARV. L. REV. 1512, 1520 (1972); Turner, *The Scope of Antitrust and Other Economic Regulatory Policies*, 82 HARV. L. REV. 1207, 1230 (1969). Williamson argues that the existence of persistent dominance ought to be sufficient for a presumption of a violation: provided only that the industry must have reached an advanced stage of development. *Id.* Implicit in this approach is the assumption that dominance is unlikely to be undone by market forces where it is found in a mature industry. See notes 146-51 *infra* for a discussion of the rate at which market power will dissipate in the absence of outside intervention. However, dissolution of dominance should not be dealt with as though it were punishment for success. Williamson, *id.* at 1522-25.

The question is not whether dominance attained through the exercise of "superior skill, foresight, and industry" is to be tolerated, but rather whether it is appropriate to put a time limit on continuing monopolistic power, even where that power is attributable to superior market performance during some prior period.

The arguments in favor of imposing such a time limit are persuasive, particularly in view of the fact that, practically speaking, it is impossible to disassociate market power from exclusionary conduct. For example, an original monopoly position will have inevitably played a role in enabling a firm continually to attract top-flight management and research personnel.

Moreover, one can never be confident that a firm has perpetuated its monopoly power solely by superior skill when it has no effective competitors against which to test its performance. *See* Turner, *id.* at 1219-20.

42 It is probable that this phenomenon could be attacked more vigorously under the Sherman Act § 2. There are those who once thought that § 2 was inappropriate in this regard, *see, e.g.*, C. KAYSEN & D. TURNER, *supra* note 23, at 21, 110-11 (1959), but who now argue that, indeed, Section 2 can be extended to cover cases of tacit collusion among oligopolists. *See, e.g.*, Turner, *supra* note 41 at 1216-17, 1225-31. Professor Turner points out that continued toleration of oligopoly power and the limited use of the divestiture remedy has led to pressure for overuse of the anti-merger laws. *Id.* at 1213-15. For a cogent, general presentation of the arguments in favor of applying Sherman Act § 2 against oligopolistic industries, *see* Sherman and Tollinson, *Public Policy Toward Oligopoly: Dissolution and Scale Economies*, 4 ANTITRUST L. & ECON. REV. 77, 82 (Summer, 1971).

Some have contended that, since framers of the Sherman Act could clearly not have had in mind an oligopoly theory which did not exist at that time, and since the problem of oligopoly was not even discussed in Supreme Court decisions prior to 1962, the necessary legal basis for application of § 2 to oligopolistic collusion is lacking. However, such a rigidly historical interpretation of the Sherman Act would tend to keep the analysis in antitrust cases within the traditional framework of a proposed changes in judicially interpreted liability standards presume a decreased emphasis on the intent or conduct element of monopolization offenses. In the former case, the conduct element is eliminated entirely; in the latter case, the requirement that collusion be explicit is removed. Up to the present, the courts have managed to avoid a careful analysis of such arguments.[43] However, the several Sherman Act § 2 cases pending in the courts will compel a confrontation of these expansive theories over the next few years.[44]

Because of the federal judiciary's reluctance to deal with the question of a broader interpretation of Sherman Act § 2 liability, several legislative proposals designed to deconcentrate American industry more effectively have been set forth over the past 20 years.[45] Most of these proposals have called for an explicitly struc-

---

simple dichotomy of monopoly and competition and render the Act ineffective. Consequently, this approach to § 2 has been largely rejected. *See* Bradley, *Oligopoly Power Under the Sherman and Clayton Acts — From Economic Theory to Legal Policy*, 19 STAN. L. REV. 285, 293-94 (1967).

43 The courts cannot be completely faulted for this state of affairs, since the enforcement arms of the government have been less than vigorous in their prosecution of significant structural cases. *See generally* M. GREEN, *supra* note 35; Blake, *Legislative Proposals for Industrial Concentration* in H. GOLDSCHMID, *supra* note 9, at 341.

44 For example, International Business Machines is being sued by the Justice Department for monopolization of the integrated computer systems industry and monopolization of the peripheral equipment product markets, BNA ANTITRUST TRADE REG. REP., No. 696, A-16 (Jan. 14, 1975). The Justice Department is also suing under § 2 to require divestiture by American Telephone and Telegraph of its wholly owned subsidiary, Western Electric, which produces most of the Bell System's equipment; AT&T's Long Lines Department, which provides long distance and overseas communications; and, possibly Bell Telephone Laboratories, a jointly-owned subsidiary of Western Electric and AT&T. BNA ANTITRUST TRADE REG. REP., No. 751, A-5 (Feb. 17, 1976); *id.*, No. 690, AA-1 (Nov. 26, 1974).

Finally, the Justice Department also has § 2 actions pending against IBM and two of the largest tire manufacturing companies. Four cereal manufacturers — General Foods, Kellogg, General Mills, and Quaker Oats — have been charged by the FTC with acting individually and collectively to maintain a highly concentrated, non-competitive market structure. BNA ANTITRUST & TRADE REG. REP., No. 729, A-11 (Sept. 9, 1975). Divestiture of the three largest firms — General Mills, Kellogg, and General Foods — has been requested as an appropriate remedy. The FTC also has § 2 suits pending against Xerox and Esso.

45 *See, e.g.*, S. 1167, 93d Cong. 1st Sess. (1973), originally introduced as S. 3832, 92d Cong., 2d Sess. (1972) (introduced by Senator Philip Hart (D-Mich.); hereinafter cited as the "Hart Bill"); C. KAYSEN & D. TURNER, *supra* note 23, at 265-72 (hereinafter the "Kaysen-Turner Bill"); *Proposed Concentrated Industries Act*, Appendix A, Neal Report, *supra* note 1 (hereinafter cited as the "Neal Bill"). (On September 30, 1971, Senator Harris (D-Ok.) introduced the Neal Bill, with very slight modification.

tural approach to liability. For example, most utilize a per se market share test as an appropriate trip-wire for liability or for a rebuttable presumption of liability.[46] In addition to these cross-

---

S. 2614, 92d Cong., 1st Sess. (1971), reprinted by Senator Harris in his Statement, 117 CONG. REC. S15442, S15447, col. 2 (daily ed. Sept. 30, 1971)). The Nader Study Group on Antitrust Enforcement has recommended a per se market share criterion similar to the Hart Bill's. See M. GREEN, *supra*, note 35 at 312. For a general discussion of legislative proposals in this area, see Blake, *supra* note 43, at 340-41.

46 Hart Bill § 101(b)(3); Kaysen-Turner Bill § 3; Neal Bill §§ 1(b), 1(c), 1(e), 4. Senator Philip Hart's (D-Mich.) Industrial Reorganization Act, perhaps the most progressive proposal, makes mere possession of "monopoly power" an offense except where it is attributable to valid patents or would result in the loss of substantial economies. *Id.* §§ 101(b), 101(c).

Senator Hart's bill would raise a rebuttable presumption of monopoly power wherever ". . . four or fewer corporations account for 50 percent or more of sales in any line of commerce in any section of the country in any year out of the most recent three." *Id.* § 101(b)(3).

In addition, the Hart Bill raises a rebuttable presumption of monopoly power on the basis of certain conduct, *id.* § 101(b)(2), and market performance, *id.* § 101(b)(1), liability trip-wires.

The White House Task Force on Antitrust Policy, under the chairmanship of Neal, proposed a deconcentration bill which was less far-reaching and more specific than the Hart Bill. See note 45, *supra*. *See generally* Keyes, *supra* note 37, for critique of this bill. The Neal Bill posits a more restricted structural liability predicate than the Hart Bill. It applies only to industries where:

> (1) any four or fewer firms have accounted for 70 percent or more of aggregate market shares (based on industry sales) during at least seven of the ten and four of the most recent five base years,
>
> (2) industry sales and the aggregate market shares of the oligopoly group are not in substantial decline,
>
> (3) the identity of the four largest firms has not substantially changed in recent years, and
>
> (4) aggregate sales have amounted to more than $500 million in at least four of the five most recent years.

Id. § 4. The Neal Bill has no conduct or performance trip-wires for liability such as those found in the Hart Bill.

The Kaysen-Turner proposal takes a less clear-cut approach to the liability standards question. *See generally* Blake, *supra* note 43, at 346. It makes illegal the "possession of market power" where market power is defined as:

> [T]he persistent ability of a person, or of a group of persons whether or not acting pursuant to agreement or conspiracy, to restrict output or determine prices without losing a substantial share of the market, or without losing substantial profits or incurring heavier losses, because of the increased output or lower prices of rivals."

Kaysen-Turner Bill § 2. A market thus defined can be evidenced by:

> (1) the persistent failure of prices to reflect substantial declines of demand or costs, or to reflect substantial excess capacity;
>
> (2) the persistence of profits that are abnormally high, taking into account such factors as risks and excess capacity; or
>
> (3) the failure of new rivals to enter the market during prolonged periods of abnormally high profits or of persistent or recurring rationing.

Id. § 2(a). Moreover, market power is conclusively presumed to exist under the industry bills dealing with market concentration, several deconcentration proposals have been aimed at specific industries.[47]

The proposals for judicial and legislative reorientation of liability standards have stimulated a good deal of academic discussion.[48] Yet, even if a broadened interpretation of the Sherman Act § 2 is ultimately adopted by the courts or new legislation is passed by Congress, the effect on market performance may be marginal unless adequate remedies accompany these broadened liability standards. It is crucial to note the important nexus between the standards for liability and guidelines employed in remedy formulation. Acceptance of the call for a structural approach to antitrust policy will necessitate greater reliance on the divestiture remedy, as it directly affects the structural characteristics which lead to liability. However, it is the thesis of this article that markedly increased usage of structural remedies[49] is dictated, whether the

---

Kaysen-Turner Bill if for five years, one company has a market share not less than 50 percent, or four or fewer firms have an aggregate market share not less than 80 percent of sales. *Id.* § 2(b). See text accompanying notes 123-27 *infra* for discussion of the problems of anticipatory degradation, which some argue accompany strict market share liability standards such as those cited above.

47 *See, e.g.,* Energy Industry Vertical Divestiture Bill, S. 2387, 94th Cong., 2d Sess. (1976). The Senate Antitrust Subcomm. of the Senate Judiciary Comm. cleared the bill for the entire Judiciary Committee by a 4 to 3 vote. BNA ANTITRUST & TRADE REG. REP., No. 758, A-1 (Apr. 6, 1976) (Text of bill *id.* at E-1). *See also,* the specific industries targeted for special attention in the Hart Bill § 203 including the following industries: chemicals and drugs; electrical machinery and equipment; electronic computing and communication equipment; energy; iron and steel; motor vehicles; and nonferrous metals.

48 *See* notes 41-42, *supra.*

49 Structural remedies are those remedies designed to affect the structure of an industry directly, by increasing the number of competitors, reducing the level of concentration or quickly lowering barriers to entry. Conduct remedies are those remedies directed at particular acts or practices of the defendant. See note 4 *supra.* The terminology employed here, "structure" and "conduct," are derived, as noted earlier, from the broadly descriptive paradigm used in industrial relations studies: basic conditions — market structure — conduct — performance. Performance depends upon conduct, which in turn depends upon market structure. Basic conditions influence both market structure and conduct. *See* F. SCHERER, *supra* note 1, at 3-6.

Even though conduct and structural remedies are thought of as alternative methods for eliminating the effects of monopolization, they are not mutually exclusive remedies. *See generally* Standard Oil Co. v. United States, 221 U.S. 1, 77-78 (1911). There is no explicit Sherman Act authorization for particular remedies. The choice of remedy is left to the discretion of the federal district court hearing the case. See text accompanying notes 65-78 *infra.* While performance remedies have been suggested for their immediate expediency, *see, e.g.,* Pollard, *Antitrust and Price Stabilization: Price Controls as a Short-run Substitute for Structural* liability standard continues to contain both structure and conduct elements or evolves into one focused exclusively on market structure. Indeed, a more liberal use of the divestiture remedy under current law could obviate the need for broader liability standards.

While the various deconcentration bills mentioned all suggest that structural remedies are appropriate,[50] the special courts or agencies[51] created by these bills would have wide discretion to formulate remedies.[52] Such wide discretion is not undesirable per

---

Reform, 7 ANTITRUST L. & ECON. REV. 97, 106 (1975), these types of remedies have "been consistently rejected as alien to the antitrust policy of preserving . . . free enterprise." Baldwin, *supra* note 28, at 126.

Conduct remedies may be designed to end specific anticompetitive acts by the defendant which may have constituted evidence for the intent element of a monopolizing charge. Alternatively, the conduct remedy may be designed to indirectly affect structural elements by, for example, reducing barriers to entry into the industry of the defendant. Over time, such a conduct remedy may moot the necessity for a structural remedy. For example, the court in the *United Shoe Machinery* case forbade a tying arrangement in order to lower the barriers to entry into the market. United States v. United Shoe Machinery Corp., 110 F. Supp. 295, 345 (D. Mass. 1953), *aff'd per curiam,* 347 U.S. 521 (1954). In fact, United Shoe Machinery's market share and total sales did decline substantially subsequent to the decree. *See* Baldwin, *supra* note 28, at 129. Judge Wyzanski thought this indicated adequate development of competition. He refused the government's petition to speed that development by dividing United into two companies, United States v. United Shoe Machinery Corp., 266 F. Supp. 328, 331-32 (D. Mass. 1967) since no "new conditions" frustrated achievement of the goals of the 1953 conduct remedy, *citing* United States v. Swift & Co., 286 U.S. 106 (1932). The Supreme Court overruled that interpretation of *Swift,* 391 U.S. 244, 249 (1968) and ordered the district court to consider a divestiture remedy if the original decree had not achieved its "principal objects," regardless of unchanged conditions. 286 U.S. at 151. See text accompanying notes 111-30 *infra.*

Structural remedies are designed to attack market power directly. Adding new competitors and reducing the market shares of existing competitors fosters competition.

50 See note 52 *infra.*

51 Hart Bill §§ 301-05; Kaysen-Turner Bill § 3; Neal Bill § 3.

52 The Hart Bill probably has the most definitive guidelines for a special enforcement court:

§ 1593. Restoration of Effective Competition.

(a) Any corporation or two or more corporations may, within sixty days from the entry of judgment pursuant to section 1592(b), file an alternative proposed order or orders of reorganization.

(b) Prior to the entry of an order of reorganization, the Industrial Reorganization Court shall conduct a proceeding to determine whether or not the proposed order or orders of reorganization would restore effective competition. In making its determination, the court may call witnesses in accordance with the provisions of sections 2652 and 2653 of this title.

(c) The court shall enter an order of reorganization appropriate to effectuate the purposes of this Act. The order or reorganization may require a corporation or two or more corporations to take such action as the court se, if accompanied by rigorous development of guidelines for remedy formulation to avoid repetition of the haphazard and ill-

---

shall find *necessary to restore effective competition.* The order may include—

(1) A requirement that a corporation modify any contract to which it is a party, terminate any agreement with another corporation, or modify its methods of distribution;

(2) A requirement that a corporation grant licenses (with or without provision for the payment of royalties) under any patent, copyright, or trademark owned by that corporation, share technical information with others, or dispose of any such patent, copyright, or trademark;

(3) A requirement that a corporation divest itself of particular assets, including tangible and intangible assets, cash, stock, securities, accounts receivable, and other obligations; and

(4) Such other requirements as the court may find necessary to restore effective competition.

(d) Any order entered under this section shall be subject to judicial review as provided in section 2114 of this title.

Hart Bill § 303(a) (emphasis added).

The Neal Bill does not provide the defendant with any of the affirmative defenses incorporated in the Hart Bill § 101. Yet, the Neal Bill § 1 also vests wide discretion in the special court:

(e) After such one-year period, further proceedings shall be conducted and a decree entered providing such further relief as may be appropriate, in light of steps taken or initiated during the one-year period, to achieve, within a reasonable period of time not in excess of four years, a reduction of concentration such that the market share of each oligopoly firm in such oligopoly industry does not exceed 12%. Such decree may include provisions requiring a party (i) to modify its contractual relationships and/or methods of distribution; (ii) to grant licenses (which may, in the discretion of the court, provide for payment of royalties) under and/or dispose of any patents, technical information, copyrights and/or trademarks; and (iii) to divest itself of assets, whether or not such assets are used in an oligopoly industry, including tangible assets, cash, stock or securities (including securities in existing firms or firms to be informed), accounts receivable and such other obligations as are appropriate for the conduct of business. The decree may also make such other provisions and require such other actions, not inconsistent with the purposes of this Act and the antitrust laws, as the court shall deem appropriate, including any provisions which would be appropriate in a decree pursuant to the antitrust laws. Such decree shall not require that a firm take any steps which such firm establishes would result in substantial loss of economies of scale.

(f) Any decree entered pursuant to subsection (e) may be appealed directly to the Supreme Court.

(g) Between four and five years after entry or affirmance of a decree pursuant to subsection (e) or a further decree pursuant to this subsection (g), proceedings shall be conducted to determine whether the decree has achieved the reduction of concentration referred to in subsection (e). If the court determines that it has not attained such end, it shall enter a further decree ordering additional steps to be taken. Such decree may be appealed directly to the Supreme Court.

(h) Any decree entered pursuant to this section 1 shall be subject to conceived remedy selection experience under Sherman Act § 2. Even if the current standards for liability remain wholly intact, the existence of the causal links between market structure and ultimate performance provides support for more extensive use of direct structural remedies.

C. *The Relation Between Liability Standards and Remedy Guidelines*

In order fully to comprehend the problems of shaping a remedy to complement particular liability standards, it is necessary to

---

> modification on the motion of any party according to the usual principles governing decrees in equity.

Neal Bill §§ 1(e), 1(f), 1(g).

The Kaysen-Turner Bill is more restrictive in the types of relief which may be granted:

> The Kaysen-Turner approach is more complex. Once "market power" is shown, it is deemed unreasonable unless shown by defendant or defendants to have been created and maintained entirely or almost entirely, by one or more of the following:
>
> (1) such economies as are dependent upon size in relation to the market;
>
> (2) ownership of valid patents, lawfully acquired and lawfully used; provided that, on a showing that market power has been created and maintained by patents, the government shall have the burden of showing invalidity, unlawful acquisition, or unlawful use;
>
> (3) low prices or superior products attributable to the introduction of new processes, product improvements or marketing methods, or to extraordinary efficiency of a single firm in comparison with that of other firms having a substantial share of the market.
>
> Even if none of these defenses is made, the Economic Court may only provide "feasible" relief, such as division or divestiture of assets, to eliminate unreasonable market power, subject to important limitations:
>
> (1) the assets of a single plant may not be divided;
>
> (2) any probable permanent loss of substantial economies intrinsic to the defendant may be taken into account in deciding the "feasibility" of division or divestiture;
>
> (3) defendant may defeat a divestiture order by showing that other relief suggested by it would provide materially equivalent competitive conditions; and
>
> (4) defendant may show that one or more companies resulting from a divestiture order would lack reasonable prospects for survival under the competitive conditions likely to prevail.

Blake, *supra* note 43, at 347-48 *citing* Kaysen-Turner Bill § 2 and discussing § 3.

Such defenses against divestiture and limitations upon its use incorporate some of the rather restrictive Sherman Act § 2 remedy selection rhetoric, discussed at text accompanying notes 250-349 *infra*.

Special agencies to investigate and report antitrust violations and issue reorganization orders have been proposed by Report of the Attorney General's National Committee to Study the Antitrust Laws 357-58 (1955) (minority report of Schwartz, et al.).

explore the relationship between the standards for liability and remedy selection that has developed under the Sherman Act § 2. The two major elements of a § 2 monopolizing charge, possession of market power and general intent to monopolize, have had a distinct impact on the form of relief granted. The need for a showing of conduct evidencing a general intent to monopolize has led the courts to grant injunctions prohibiting specific conduct modes rather than dealing with the underlying causative market structure elements.[53] Generally, divestiture and other structural remedies are considered only after remedies enjoining conduct have been judged inadequate.[54] The courts have continually opted for wide-ranging equitable restraints on conduct where divestiture arguably would have accomplished the same ends with less ongoing court supervision of the implementation of the remedy,[55] and where divestiture was probably the less severe and more effective remedy.[56]

Judicial focus on conduct is appropriate only given the following common assumptions: (1) conduct injunctions and structurally oriented relief are equally effective in preventing continuing or renewed violation of § 2; and (2) conduct remedies are somehow less harsh than structural remedies, and therefore use of the more onerous relief measure would involve punitive overtones. However, the assumption regarding equal effectiveness of remedies has been refuted by the literature of industrial organization economists indicating that market structure is largely determinative of conduct and performance. In nearly all cases, it is clear that only alteration of market structure can significantly affect market

---

53 See text accompanying notes 65-99 *infra*. Emphasis on the conduct element of § 2 offenses has not led to the imposition of criminal sanctions. That remedy is widely considered to be inappropriate when industry organization in violation of § 2 has resulted in good performance and is not morally derelict, by common law standards. *See* Turner, *supra* note 39, at 1221-22; W. SHEPHERD, *supra* note 13, at 185; P. AREEDA, ANTITRUST ANALYSIS 52-53 (2d ed. 1974).

54 See text accompanying notes 300-39 *infra* for a discussion of the "necessity" requirement. The courts have interpreted the Sherman Act as a mandate to develop a federal common law of antitrust requiring equitable restraints on conduct which is per se anti-competitive and unjustified, but not criminal. *See* Northern Pacific Ry. Co. v. United States, 356 U.S. 1 (1958); United States v. Aluminum Co. of America, 148 F.2d 416 (2d Cir. 1945).

55 See text accompanying notes 240-49 *infra* for a discussion of the relative harshness of divestiture.

56 See text accompanying notes 170-89 *infra*.

conduct and thereby remedy undesirable market performance.[57] The supposed greater harshness of structural remedies, therefore, is largely irrelevant. Sherman Act § 2 interdicts monopolization and requires effective counteraction. It would not be proper to balance the harshness of an effective structural remedy against the harshness of a conduct-regulating scheme which is ineffective and not, in fact, a remedy.[58] And, not insignificantly, years of experience have exhibited that a conduct-directed remedy may be substantially more burdensome and costly to those directly involved than its structurally-oriented alternative.[59]

In light of these serious challenges to the viability of the dual underpinnings which support the use of conduct remedies, it becomes increasingly difficult to advocate the utilization of specific conduct injunctions as § 2 remedies. Furthermore, it becomes increasingly apparent that, whether one focuses on the conduct element or the structural prerequisite of the current two-prong Sherman Act § 2 violation standard, a direct structural remedy generally will be the most appropriate and effective. Until the challenge of implementing such remedies is met, § 2 may remain the dead letter of antitrust law: for in most monopoly cases, it is the foreseeability, availability, and likelihood of a practical remedy that determine whether the case will be brought[60] and substantially influence resolution of the liability question.[61] The success or failure of the recent resurgence of § 2 cases[62] will depend to a great extent upon the ingenuity of the courts in their effort to design and implement effective remedies.

In summary, an examination of the merits and shortcomings of divestiture under the current interpretation of Sherman Act § 2 is relevant not only to the selection of cases that will be prosecuted thereunder, but also to remedy selection under the suggested broadened interpretation of § 2. Additionally, examination of this experience is useful given the broad discretion granted to

57 See text accompanying notes 191-210 *infra.*
58 See text accompanying notes 211-39 *infra.*
59 See text accompanying notes 211-39 *infra.*
60 See text accompanying note 13 *supra.*
61 See text accompanying note 53 *supra.*
62 See note 44 *supra.*

the special enforcement courts which the various deconcentration bills propose.[64] The policies expressed in Sherman Act § 2 can be successfully pursued, regardless of the liability standards employed, only with effective remedy utilization.

II. The History and Use of the Divestiture Power

A. *The Divestiture Power and the Courts' Reluctance to Utilize It*

In providing the setting for this exhortation toward increased use of divestiture under § 2 of the Sherman Act, it is appropriate to examine the present power of the courts to specify direct structural intervention, and the propensity toward the exercise of such power. The Sherman Act contains no explicit authorization of divestiture relief.[65] Rather, that section:

> invests United States district courts with authority to "prevent and restrain" violations of the act in equity proceedings instituted by the Attorney General of the United States. 26 Stat. 209 (1890), *as amended,* 15 U.S.C. § 4 (1970). An essential aspect of a court of equity is that it possesses a degree of discretion not present in a court of law, enabling it to adapt its remedy to the particular case.[66]

Thus, the district courts' power to grant divestiture arises out of long-standing equity powers of the federal courts to shape remedies to meet the facts of particular cases.[67]

---

64 See note 45 *supra.*

65 Wise, *Three D's of Antitrust Enforcement,* Hoffmann's Antitrust Law and Techniques (1963).

66 Fraidin, *Dissolution and Reconstitution: A Structural Remedy, and Alternatives,* 33 Geo. Wash. L. Rev. 899, 903 (1965), *citing* I. Pomeroy, Equity Jurisprudence 217 (Symons ed. 1941).

Similar assessment has been expressed by the Supreme Court:

> . . . [the district courts] . . . are clothed with large discretion to model their judgments to fit the exigencies of the particular case.

United States v. E. I. du Pont de Nemours & Co., 353 U.S. 586, 607-08 (1957) *citing* Int'l Salt Co. v. United States, 332 U.S. 392, 400-01 (1947). It is interesting to note that § 4 of the Sherman Act, 15 U.S.C. § 4 (1970), grants the same remedy powers as those provided for in the Wilson, Clayton, and Robinson-Patman Acts, 15 U.S.C. §§ 9, 25, 25 (1970), respectively.

67 *See, e.g.,* Timken Roller Bearing Co. v. United States, 341 U.S. 593, 600 (1951);

Theoretically, the equitable power to fashion relief in antitrust litigation is wideranging.[68] Once a violation of the Sherman Act is established the entire panoply of remedies, including divestiture, becomes available.[69] While it is generally true that the courts have usually chosen to employ only a narrow range of conduct remedies,[70] direct structural action has been prescribed more than occasionally. The *Paramount Pictures* case[71] evidences the latitude permitted the federal courts in fashioning § 2 remedies. The district court, in order to limit restrictive methods employed in granting film exhibition rights, compelled the defendant distributors to award exhibition rights on the basis of competitive bidding among the distributor-owned exhibitors and others.[72] The Su-

---

Appalachian Coals, Inc. v. United States, 288 U.S. 344 (1933); United States v. Crescent Amusement Co., 323 U.S. 173, 185 (1944); United States v. National Lead, 332 U.S. 319, 334-35 (1947); United States v. E. I. du Pont de Nemours, 366 U.S. 316, 322 (1961). *But see* Timken Roller Bearing Co. v. United States, 341 U.S. 593, 602 (1951) (dissent).

The legislative history behind this section and the parallel section of the Clayton Act indicates that divestiture was a contemplated remedy for antitrust violations. *See generally* Standard Oil Co. v. United States, 221 U.S. 1, 63-70 (1911); Comment, *Private Divestiture: Antitrust's Latest Problem Child,* 41 FORDHAM L. REV. 569, 588-90 (1973); Adams, *Dissolution, Divorcement, Divestiture: The Pyrrhic Victories of Antitrust,* 27 IND. L.J. 1 (1951); Van Cise, *Limitations Upon Divestiture,* 19 GEO. WASH. L. REV. 147 (1950); Comment, *Aspects of Divestiture as an Antitrust Remedy,* 32 FORDHAM L. REV. 135 (1963); Note, *Availability of Divestiture in Private Litigation as a Remedy for Violation of Section 7 of the Clayton Act,* 49 MINN. L. REV. 267 (1964). For example, in a floor debate Senator Sherman was asked what sort of relief would be available to the Government in a Sherman Act case. He replied that "it may be a judgment of ouster of the corporation; it may be a judgment for damages." 21 CONG. REC. 1768 (1890). Later, he elaborated saying:

> . . . [in] a civil proceeding commenced by the people of the United States against these corporations, . . . a judgment may be, as in ordinary cases, an ouster of the power of a corporation; it may be for damages; there may be an injunction; there may be proceedings in *quo warranto,* and so of the other ordinary civil proceedings which are fixed by the judiciary act of the United States.

Id. at 2564. Sherman later referred to "enjoining", "restraining", "preventing", or "tying up" the combinations which were in violation of § 2. *Id.* at 2569.

68 *See* A. NEALE, THE ANTITRUST LAWS OF THE UNITED STATES 403 (2d ed. 1970). *See* United States v. American Can Co., 234 F. 1019 (D.C. Md. 1916); United States v. Crescent Amusement Co., 323 U.S. 173, 185 (1944); International Salt Co. v. United States, 332 U.S. 392 (1947); Associated Press v. United States, 326 U.S. 1 (1945).

69 *See* A. NEALE, *id.* at 404-05; Celler, *The Trial Court's Competence to Pass Upon Divestiture Relief,* 10 ANTITRUST BULL. 693, 695-96 (1965).

70 See text accompanying notes 76-87 *infra.*

71 United States v. Paramount Pictures, Inc., 334 U.S. 131 (1948).

72 *Id.* at 174-75.

preme Court found this conduct-regulating relief insufficient, and ordered the district court to require the distributors to divest completely their interest in some of the exhibitors.[73] The Court rejected the conduct-related relief not because such relief exceeded the district court's power, but because divestiture was a more effective remedy.[74]

The courts' flexibility in remedy formulation allows for the complementary use of conduct injunctions and divestiture within the same decree. In *United States v. American Tobacco Co.*, divestiture involved the unscrambling, creation, and dissolution of dozens of corporations.[75] Additionally, the defendants were subjected to a range of injunctions, including prohibition against participation in any similar conspiracy and against any form of common ownership or control of the 14 companies to which the assets of the combination were conveyed as a result of the divestiture order.[76]

Despite the Supreme Court's active role in the *Paramount Pictures* case, the general rule provides that the scope of an antitrust decree is peculiarly within the responsibility of the trial court.[77] Yet it is noteworthy that on those occasions where the Supreme Court has played a direct role in the formulation of relief, it has been considerably more aggressive in granting divestiture than have the district courts.[78]

Despite this apparently wide-ranging authority the courts generally have refused to prescribe divestiture. Professor Richard Posner has noted the infrequent application of dissolution or sub-

---

73 *Id.* at 166-76.

74 *Id. See* A. NEALE, *supra* note 88, at 407-15.

75 221 U.S. 106, 167-73 (1911).

76 221 U.S. at 184-88. The decree appears in 191 F. 371, 417-31 (C.C.S.D.N.Y. 1911). The *Standard Oil* case, Standard Oil Co. v. United States, 221 U.S. 1 (1911), was similar except that the offense in that case consisted of stock holding company transfers rather than asset acquisitions as in *American Tobacco.*

77 *See, e.g.,* United States v. E. I. du Pont de Nemours & Co., 366 U.S. 316, 322 (1961); United States v. United States Gypsum Co., 340 U.S. 76, 88-89 (1950); United States v. Crescent Amusement Co., 323 U.S. 173, 185 (1944); United States v. National Lead, 332 U.S. 319, 334-35 (1947). *See also* Celler, *The Trial Court's Competence to Pass Upon Divestiture Relief,* 10 ANTITRUST BULL. 693, 702 (1965). Note that a direct appeal lies to the Supreme Court in civil cases instituted by the United States. 15 U.S.C. § 29 (1970).

78 Celler, *supra* note 69, at 703-04.

stantial divestiture over nearly 70 years of Sherman Act § 2 enforcement:[79]

TABLE A
The Use of Dissolution or Divestiture Decrees in Department of Justice Monopolization Cases

Period in Which Case Was Instituted	Number of Cases in Which Significant Divestiture or Dissolution Was Carried Out	National or Large Regional Monopolist	Local or Small Regional Monopolist
1890-1904	0	0	0
1905-1909	3	3	0
1910-1914	5	5	0
1915-1919	0	0	0
1920-1924	2	2	0
1925-1929	1	0	1
1930-1934	2	0	2
1935-1939	4	2	2
1940-1944	1	1	0
1945-1949	7	6	1
1950-1954	5	3	2
1955-1959	0	0	0
1960-1964	2	2	0
1965-1969	0	0	0
Total	32	24	8

The Report of the Attorney General on the Antitrust Laws in 1955 found that in over 60 years of Sherman Act enforcement, divestiture had been granted in only 24 litigated cases.[80] Yet even these statistics may overstate the frequency with which substantial

79 Posner, *A Statistical Study of Antitrust Enforcement*, 13 J.L. & ECON. 365, 404 Table 29, 406 (1970). Table computed from the "Bluebook," note 80 *infra*. *See also* Wise, *supra* note 65, at 409.

80 Report of the Attorney General's National Committee to Study the Antitrust Laws 354 n.13 (1955). The discrepancy between Posner's statistics and those in the 1955 Report is understandable, given the reliance of the Report on the "Bluebook," *supra* note 79, at 366-67, which includes every complaint, indictment and information filed by the Department of Justice. The Bluebook, as cited by Posner, includes the CCH, THE FEDERAL ANTITRUST LAWS, WITH SUMMARY OF CASES INSTITUTED BY THE UNITED STATES 1890-1951 (1952); 1952-56 Supp. (1957); TRADE REG. REP., 10th ed., Transfer Binder, New U.S. Antitrust Cases — Complaints, Indictments, Developments 1957-1961; and 5 TRADE REG. REP. ¶ 45003-45069 (current). The Bluebook tends to overstate the number of cases litigated. For example, indictments often are dismissed and refiled thus leading to double counting.

divestiture is utilized. In many instances where divestiture was ordered, the combination of corporations had been consummated primarily through explicit collusive arrangements or common stock control,[81] and frequently with little or no integration of the

81 United States v. Grinnell Corp., 384 U.S. 563 (1966) (Sherman Act § 2 monopolization violation found where defendant, through four subsidiaries, controlled 87 percent of the national accredited central station fire and burglar alarm service business, *id.* at 571, and that this monopoly power was achieved by restrictive agreements, anti-competitive pricing decisions, and the acquisition of the subsidiaries themselves, *id.* at 576); International Boxing Club of New York, Inc. v. United States, 358 U.S. 242 (1959) (two boxing clubs found guilty of monopolizing and conspiring to restrain competition in the promotion of world championship boxing matches, *id.* at 245-52); Schine Chain Theaters, Inc. v. United States, 344 U.S. 110 (1948) (parent corporation and subsidiaries which used combined buying power to exclude other theaters from competition in selected cities over six state area found to be §§ 1 and 2 Sherman Act violations, *id.* at 116-17); United States v. National Lead Co., 332 U.S. 319 (1947) (formation of "international cartel" for control of titanium industry found to be violation of § 1 of Sherman Act; *id.* at 325-27); Hartford Empire Co. v. United States, 323 U.S. 386 (1945) (glass companies' formation of pool of patents which in effect controlled industry found to be violation of §§ 1 and 2 of Sherman Act, *id.* at 400-403); United States v. Crescent Amusement Co., 323 U.S. 173 (1944) (nine affiliated companies under common ownership which owned movie theaters in 70 towns in five states found in violation of § 2 of Sherman Act, *id.* at 185-91); United States v. Lehigh Valley R.R., 254 U.S. 255 (1920) (railroad company ownership of significant coal producing lands found to be § 1 and § 2 violation, *id.* at 269); United States v. Reading Co., 253 U.S. 26 (1919) (holding company control of two railroads and two coal companies which employed those railroads' shipping facilities found to be § 2 violation, *id.* at 57); United States v. Union Pacific R.R., 226 U.S. 63 (1912) (purchase of controlling portion of shares of one railroad by competing railroad found to be violation of § 2 Sherman Act, *id.* at 96); Standard Oil Co. of New Jersey v. United States, 221 U.S. 1 (1911) (oil holding company ordered dissolved pursuant to § 1 and § 2 of Sherman Act, *id.* at 77-82); United States v. American Tobacco Company, 221 U.S. 106 (1911) (horizontal conspiracy of tobacco company ordered dissolved pursuant to Sherman Act §§ 1 and 2, *id.* at 184-85); Northern Securities Co. v. United States, 193 U.S. 197 (1904) (combination of two competing railroads via holding company found to be violation of § 2 Sherman Act, *id.* at 325); United States v. Imperial Chemical Industries, 100 F. Supp. 504 (S.D.N.Y. 1951), *order on decree,* 105 F. Supp. 215 (S.D.N.Y. 1952) (patent pooling and joint acquisitions by chemical companies found to be violation of § 1 Sherman Act, *id.* at 593-94); United States v. Besser Mfg. Co., 96 F. Supp. 304 (E.D. Mich. 1951), *aff'd,* 343 U.S. 444 (1952) (violation of §§ 1 and 2 of Sherman Act found by company dominating concrete block manufacturing industry, *id.* at 314-15); United States v. Minnesota Mining and Manufacturing Co., 92 F. Supp. 947 (D. Mass. 1950) (combination of chemical companies to control export trade of coated abrasives found to be violation of § 1 of Sherman Act, *id.* at 964); United States v. New England Fish Exchange, 258 F. 732 (D. Mass. 1919) (creation of exclusive membership fish exchange found to be Sherman Act violation, *id.* at 746-48); United States v. Corn Products Refining Co., 234 F. 964 (S.D.N.Y. 1916), *appeal dismissed,* 249 U.S. 621 (1919), *decree modified,* Oct. 18, 1921 (illegal combination of plants into Corn Products Refining Co. found to be violation of §§ 1

productive operations by the combining firms even though the suits frequently were brought long after the alleged violation.[82]

---

and 2 of Sherman Act, *id.* at 1011-18); United States v. Eastman Kodak Co., 226 F. 62 (W.D.N.Y. 1915), *decree entered,* 230 F. 522 (W.D.N.Y. 1916), *cert. denied,* 255 U.S. 578 (1921) (Kodak forced to divest itself of property wrongfully obtained from competitors pursuant to §§ 1 and 2 of Sherman Act, *id.* at 80-81); United States v. International Harvester Co., 214 F. 987 (D. Minn. 1914), *petition for additional relief denied,* 10 F.2d 827 (D. Minn. 1926), *aff'd,* 274 U.S. 693 (1927) (consolidated ownership of harvester companies found to be §§ 1 and 2 Sherman Act violations, *id.* at 998-1000); United States v. Lake Shore & M.S. Ry. Co., 203 F. 295 (S.D. Ohio 1912) (combination of coal carrying railroads and coal mines under common ownership found to be violation of §§ 1 and 2 of Sherman Act, *id.* at 318); United States v. E. I. du Pont de Nemours, 188 F. 127 (C.C. Del. 1911) (violation of §§ 1 and 2 of Sherman Act found among gunpowder producers who allocated markets and acquired shares of competing companies, *id.* at 134-52); United States v. Reading Co., 183 F. 427 (C.C.E.D. Penn. 1910), *aff'd,* 226 U.S. 324 (1912), *modified,* 228 U.S. 158 (1913) (no violation of Sherman Act found in general agreement between producers and shippers of anthracite coal, *id.* at 456, 457).

82 Schine Theaters v. United States, 344 U.S. 110 (1948) (acquisition of monopolistic theater chain begins in 1920, 22 years before initiation of suit, *id.* at 113); United States v. National Lead Co., 332 U.S. 319 (1947) (division of sales and allocation of markets by titanium producers occurred for 24 years before initiation of suit, *id.* at 323); Hartford Empire Co. v. United States, 323 U.S. 386 (1945) (22 years elapse between first patent pooling agreements in glass industry and initiation of temporary National Economic Committee investigation, *id.* at 400); United States v. Crescent Amusement Co., 323 U.S. 173 (1944) (major growth of monopolistic theater chain takes place in five years prior to initiation of suit, *id.* at 178); United States v. Lehigh Valley R.R. Co., 254 U.S. 255 (1920) (railroad company policy of acquiring coal producing land began in 1868, 52 years before completion of suit, *id.* at 259); United States v. Reading Co., 253 U.S. 26 (1919) (holding company railroads employed by some company's mines for shipping of coal for fourteen years before completion of ICC investigation of business completed, *id.* at 53); United States v. Union Pacific R.R. Co., 226 U.S. 63 (1912) (purchase by first railroad of second railroad's stock occurs eleven years prior to Supreme Court decree); Standard Oil Co. of New Jersey v. United States, 221 U.S. 1 (1911) (formation of holding company began 36 years before bringing of suit, *id.* at 31); United States v. American Tobacco Co., 221 U.S. 106 (1911) (combinations of stock ownership by tobacco companies existed for five years before bill brought, *id.* at 149); United States v. Imperial Chemical Industries, 100 F. Supp. 504 (S.D.N.Y. 1951), *order on decree,* 105 F. Supp. 215 (S.D.N.Y. 1952) (combinations concerning patents and acquisitions existed for 24 years before bringing of suit, *id.* at 508 and 593); United States v. Minnesota Mining and Manufacturing Co., 92 F. Supp. 947 (D. Mass. 1950) (formation of holding company to acquire and construct foreign plants occurs 21 years before suit brought, *id.* at 950 and 951); United States v. New England Fish Exchange, 258 F. 732 (D.C. Mass. 1919) (no new members allowed in fish exchange for eight years before bringing of suit, *id.* at 737); United States v. Corn Products Refining Co., 234 F. 964 (S.D.N.Y. 1916), *appeal dismissed,* 249 U.S. 621 (1919), *decree modified,* Oct. 18, 1921 (consolidation and acquisitions of corn products plants began in 1900, twelve years before final *decree, id.* at 966); United States v. Eastman Kodak Co., 226 F. 62 (W.D.N.Y. 1915), *decree entered,* 230 F. 522 (W.D.N.Y. 1916), *cert. denied,* 255 U.S. 578 (1921) (wrongful acquisition of competing photographic forms occurs over period of fifteen years,

Thus, the "structural" measures taken were often merely "paper transactions" rather than full-blown divisions of firm assets.

Divestiture has been ordered in three dominant-firm monopolization cases, yet in only one has the remedy been fully employed.[83] In fact, in only two instances has the corporate structure of an industry ever been seriously altered under the Sherman Act.[84] A partial explanation for this infrequency may lie in the fact that approximately 80 percent of the monopolization cases brought under Sherman Act § 2 have involved aggressive cartels rather than single-firm monopolists or oligopolistic tacit collusion.[85] Thus, in those instances where the industry is already relatively deconcentrated, it is seemingly more appropriate to interdict the collusive conduct, rather than to order further deconcentration.[86]

---

id. at 65-66); United States v. International Harvester Co., 214 F. 987 (D. Minn. 1914), *petition for additional relief denied,* 10 F.2d 827 (D. Minn. 1926), *aff'd,* 274 U.S. 693 (1927) (consolidation of harvester companies begins ten years before suit brought; integrated plants constructed only within year before suit brought, *id.* at 993); United States v. Lake Shore & M.S. Ry. Co., 203 F. 295 (S.D. Ohio 1912) (joint ownership of competing railroads and coal mines occurred for ten years before bringing complaint, *id.* at 300); United States v. E.I. du Pont de Nemours & Co., 188 F. 127 (C.C. Del. 1911) (acquisition of shares agreements concerning markets by competing companies begun 39 years before final decree, *id.* at 135); United States v. Reading Ry. Co., 183 F. 427 (C.C.E.D. Pa. 1910), *aff'd,* 226 U.S. 324 (1912), *modified,* 228 U.S. 158 (1913) (major violations of Sherman Act alleged to begin fourteen years before suit brought, *id.* at 430 and 432).

But see United States v. Grinnell Corp., 384 U.S. 563 (1963) (Grinnell and its subsidiaries involved in a variety of market sharing agreements from 1906 through to time of decision although there was little evidence of much actual integration of operation, *id.* at 566-70); International Boxing Club of New York, Inc. v. United States, 358 U.S. 242 (1959) (virtually no real economies from joint operation and control of Madison Square Garden which defendants controlled, *id.* at 245-52); United States v. Besser Manufacturing Co., 96 F. Supp. 304 (E.D. Mich. 1951), *aff'd,* 343 U.S. 444 (1952) (major acquisition by Besser of competitor's stock occurred five years before suit brought. No integration occurred, *id.* at 307); Northern Securities Co. v. United States, 193 U.S. 197 (1904) (holding company formed for two years before Supreme Court decree).

83 *See* United States v. Pullman Co., 50 F. Supp. 123 (E.D. Pa. 1943), *aff'd per curiam,* 330 U.S. 806 (1947). In the *Alcoa* and *United Shoe* cases the courts shied away from significant structural relief. *See* United States v. Aluminum Co. of America, 148 F.2d 416 (2d Cir. 1945); United States v. United Shoe Machinery Corp., 110 F. Supp. 295 (D. Mass. 1953), *aff'd per curiam,* 347 U.S. 521 (1954).

84 D. DEWEY, MONOPOLY IN LAW & ECONOMICS 257-63 (1959).

85 *See* cases cited at note 82 *supra.*

86 Although this paper is focused upon dominant firm cases, a necessary implication of the recommendation presumptively to employ divestiture in that area is that structural measures are also necessary to control collusive cartels. This results from the contention that conduct injunctions are ineffective, and that, thus, the

But even after taking into account this further breakdown of § 2 cases, it is apparent that the courts have been extremely reluctant to grant divestiture under the Sherman Act.[87]

B. *The Reasons for Judicial Timidity in Remedy Selection*

Examination of the judicial propensity to avoid the issuance of structural relief reveals that such a bias is not *happenstantial.* The reasons for this judicial timidity in granting the divestiture remedy are best categorized according to three tenets apparently held by the judiciary: (1) that conduct-related relief is sufficiently effective in dealing with monopolizing offenses; (2) that divestiture is an extremely "harsh" remedy vis-à-vis alternative remedies; and (3) that the district courts are precedentially circumscribed in their ability to order divestiture relief.[88]

Evidence of judicial acceptance of these premises is readily apparent. The notable rarity with which divestiture has been deemed appropriate testifies to the first of these propositions.[89] On those infrequent occasions where mere possession of market power is viewed as the target of a remedy, divestiture is presumptively more effective than alternative forms of relief.[90] But where the

---

structural situation, which provides the impetus for such collusion, must be altered. *See* text accompanying notes 170-89 *infra.*

87 *See* Hart, *Restructuring the Oligopoly Sector: The Case for a New Industrial Reorganization Act,* 5 ANTITRUST L. & ECON. REV., 35, 42 (Summer 1972); E. Mason, *Preface* to C. KAYSEN & D. TURNER, ANTITRUST POLICY xxi (1959); Baker, *Section 2 Enforcement — The View from the Trench,* 41 ABA ANTITRUST J. 613, 618 (1972).

88 Perhaps an additional reason for the reluctance of the courts to break up large companies is the oft-cited judicial bias in favor of established property rights. *See, e.g.,* D. DEWEY, MONOPOLY IN LAW AND ECONOMICS 247 (1959). It is clear that judges are drawn from particular social/educational strata. They are schooled in free enterprise and private property rights and are somewhat awed by business expertise and the dangers of disturbing the existing order. W. SHEPHERD, THE TREATMENT OF MARKET POWER 69 (1975). Additionally, there is a feeling that "mistakes" ought to be avoided at all costs. *See generally,* M. FLEMING, THE PRICE OF PERFECT JUSTICE (1973). Consequently, the margin of error is taken on the side of the seemingly less drastic and less permanent conduct oriented measures, and since prediction as to the effects on any alteration of industrial activity is difficult, divestiture is rarely employed.

89 *See* text accompanying notes 76-87 *supra.*

90 *See* text accompanying notes 300-32 and 170-89 *infra* for a discussion of the broadest interpretation of the necessity requirement and the relative effectiveness of structural relief, respectively.

target of the remedy, rightly or wrongly, is perceived to be the particular exercise of market power in a monopolization charge, the courts generally view conduct remedies as sufficient.[91]

The second of these judicial perceptions is the basic assumption, shared by the general public, that even on those occasions where divestiture might be more effective than conduct-related relief, the result of such dismantling might be in some way extraordinarily costly.[92] Frequently, as noted above, the courts confuse the analysis of such costs by commingling two discrete categories of costs when discussing the relative harshness of divestiture vis-à-vis other forms of relief: real economic costs to the public, and immediate impact on those with a direct financial interest in the defendant corporation.[93] But regardless of proper classification, it is apparent that potential adverse economic effects weigh heavily, perhaps too heavily, in the remedy selection balance. As Milton Handler has noted in discussing the defendants' victory in *United States Steel*:[94] "Indeed, one cannot read the opinion in the *Steel* case without obtaining the firm impression that the Court's apprehension of the adverse economic effects of dissolution were in part responsible for its doubtful construction and application of the law."[95] And as for the detrimental effects of divestiture on directly financially interested parties it is noteworthy that even in Clayton Act § 7 cases, where divestiture has been considered a "natural" remedy, the district courts frequently refused to grant divestiture because of this factor.[96]

An additional perception, often considered under the harshness rubric, is that divestiture will entail greater court supervision of an industry over a longer period of time than will a conduct-

91 *See* text accompanying notes 300-39 *infra*.

92 A. NEALE, *supra* note 88, at 95-96 (2d ed. 1970); Baldwin, *supra* note 28, at 139. *See also* Hart, *supra* note 87, at 42.

93 *See* cases cited at note 211 *infra*.

94 United States v. United States Steel Corp., 251 U.S. 417 (1920).

95 M. HANDLER, A STUDY OF ENFORCEMENT OF THE FEDERAL ANTITRUST LAWS 84 (TNEC Monograph No. 38, 1941). *See also* United States v. American Can Co., 230 F. 859, 903 (D. Md. 1916) ("I am frankly reluctant to destroy so finely adjusted an industrial machine as the record shows defendant to be. . . ." *Id.*).

96 United States v. E.I. du Pont de Nemours, 177 F. Supp. 1 (N.D. Ill. 1959), *modified*, 366 U.S. 316, 328-33 (1961) (Supreme Court says divestiture is preferred remedy in Clayton Act cases).

oriented remedy.[97] Such continuing regulation is deemed to be an inappropriate role for the courts. However, it is far from clear that the judicial monitoring required for a divestiture remedy would be more onerous than the post-trial supervision required for conduct prohibiting or regulating injunctions.[98]

The last of the three assumptions which have led to timidity in granting divestiture is that divestiture is disfavored under current law, and that its frequent and obvious utilization will trigger a legislative response rebuking such judicial activism.[99] Yet while divestiture is indeed a fairly visible remedy,[100] there has not been a case where a court which granted divestiture in either a Sherman or Clayton Act case has been legislatively rebuked. Nevertheless, the history of § 2 remedy selection does little to encourage utilization of the divestiture remedy and provides scant precedent for courts currently inclined toward implementation of structural measures.

The following sections of this paper will analyze the economic presuppositions which to some extent determine remedy selection under the Sherman Act and then explore the validity of these

---

97 Breit & Elzinga, *Antitrust Penalties and Attitudes Toward Risk: An Economic Analysis*, 86 HARV. L. REV. 693, 695-96; M. GOLDBERG, THE CONSENT DECREE; ITS FORMULATION AND USE 49-61 (1962); Adams, *Dissolution, Divorcement, Divestiture: The Pyrrhic Victories of Antitrust*, 27 IND. L.J. 1, 33 (1951).

98 *See* text accompanying notes 241-49 *infra*.

99 *See, e.g.*, United States v. United Shoe Machinery Corp., 110 F. Supp. 295, 348 (D. Mass. 1953), *aff'd per curiam*, 347 U.S. 521 (1954) (Wyzanski, J.).

100 *See* Baldwin, *supra* note 28, at 139.

> There are reasons why dissolution and divestiture of meaningful proportions might be regarded as drastic remedial devices and as solutions to be avoided whenever possible. Courts have properly been concerned that dissolution might destroy genuine and significant economies of scale. Dissolution or drastic divestiture is a dramatic and highly visible remedy. A corporate person is dismembered; and those who have had long association with the company, who are imbued with its traditions and who identify with it in the sense of "company loyalty" are prone to symbolize such dismemberment and view dissolution as a corporate analogue to the death penalty. The effects of both dissolution and divestiture are immediately obvious to customers, employees and investors; and the legal and financial mechanics of separation may be complex. Final compliance with the decree may not be achieved for a substantial length of time, and there may be disputes among claimants in the settlement. As a result, expenses of compliance may be substantial. Further, dissolution and divestiture decrees and their aftermaths are likely to receive broader and more protracted press coverage than are cases settled by restraints on conduct.

Id.

judicial perceptions which have led to the disfavored status of divestiture.

III. Perspectives on Remedy Choice and Construction for Sherman Act Monopolization Violations

A. *The Populist Perspective Versus the Economic Efficiency Perspective*

Reasoned selection of the appropriate Sherman Act § 2 remedy necessarily is informed with a particular perspective regarding the proper role of antitrust enforcement in the American economy. Debate over that role has proceeded at two levels. First, should antitrust policy be limited to economic goals or should it also encompass non-economic, social and political goals?[101] Second, assuming that economic goals should play a significant, if not determinative, role in the formulation of antitrust policy, does current economic literature in the field of industrial organization provide consistent guidance for the process of remedy selection?[102]

The first question reflects the perennial debate between those who believe that antitrust ought to deal exclusively with economic objectives, and those who believe that antitrust policy ought to include the pursuit of various populist goals. A deep mistrust of any large concentrations of economic power forms the basis of such a populist philosophy.[103] The populist mistrust of "bigness" is based not simply on the effects of firm size on market performance, but also on a belief that the possession of unfettered discretion somehow allows large enterprises to abuse the people who work for them, buy from them, or own part of them.[104]

---

101 *See generally* Comanor, *The Nader Report and the Goals and Limits of Antitrust*, 20 Pub. Pol. 397, 398-99 (1972). *Compare* M. Green, *supra* note 35, *with* Turner, *supra* note 41.

102 *See* text accompanying notes 116 *et seq., infra. See also* Irwin & Barrett, *Antitrust Enforcement in the United States: Market Structure Versus Market Conduct,* 1974 Wash. U.L.Q. 37, 38-42.

103 *See* Comanor, *supra* note 101, at 398.

104 *See, e.g.,* the dicta of Judge Learned Hand in United States v. Aluminum Co. of America, 148 F.2d 416, 428 (2d Cir. 1945): ". . . great industrial consolidations are inherently undesirable, regardless of the economic results." Professor Galbraith, however, emphasizes the positive results stemming from the possession of substantial

It has been urged that the populist philosophy should define both antitrust liability standards[105] and the enforcement strategy of the relevant federal antitrust enforcers.[106] In addition, the populist argument can easily be extended to the remedy selection problem: whether liability is found to exist under current Sherman Act § 2 standards or under liability provisions in any of the recently proposed § 2 revisions, the remedy selected ought to be the one designed to reduce most effectively the absolute size of the violating enterprise. It will be argued later that this prescription is unnecessarily severe in that it is oblivious to much of the current learning of industrial organization economics. This literature suggests that factors other than absolute size contribute to the firm's discretionary powers and emphasizes that considerable advantages may result from appropriately scaled firms.[107] However, because the populist tradition runs deep in this country,[108] the remedy selection decision must be predicated on the assumption that, given no substantial diseconomies, absolutely smaller industrial enterprises are socially and politically desirable.[109]

The second strain of antitrust policy, that enforcement strategies and standards of liability ought to be designed primarily with economic objectives in mind, has sounded equally throughout the history of America.[110] At the extreme, this theme has been the basis of a claim that the exclusive objective of antitrust policy is the attainment of the highest degree of allocative efficiency.[111] Something approaching this pure economic efficiency model informed the administration of the Antitrust Division of the Justice Department during the tenure of Professor Donald Turner from 1965 to 1969.[112]

The economic efficiency view suggests that antitrust policy should not be directed at the absolute size of firms, or at any specific business practices, except insofar as consumer welfare loss is

---

discretion. He argues that large enterprises are needed to plan effectively for the future. J. GALBRAITH, THE NEW INDUSTRIAL STATE 40 n.6 (1969).

105 *See* M. GREEN, *supra* note 43, at 249, 287.

106 *Id.* at 84.

107 *See* text accompanying note 9 *infra.*

108 *See* Comanor, *supra* note 101, at 398.

109 *But see* J. GALBRAITH, *supra* note 104.

110 *See* Comanor, *supra* note 101, at 399.

111 *Id. citing* Bork, 58 AM. ECON. REV. 242, 242-53 (May 1966).

112 Comanor, *supra* note 101, at 398-99. *See generally* Turner, *supra* note 41.

thought to result. Under this view, mere size ought not be a violation of the Sherman Act § 2[113] since there are a host of other structural and conduct elements which may affect performance.[114]

Applying the economic efficiency perspective to remedy selection would lead to the rule that, where a particular element of market structure or conduct is determined to be the source of poor economic performance, remedial action should be directed against the specific causal element. In choosing the remedy, the court would apply a rigorous cost-benefit test to each alternative, selecting the appropriate remedy on a purely economic basis. However, the costs and benefits of a particular remedy often will be uncertain at best and simply inestimable at worst.

Thus, the economic efficiency approach is not as clearly dispositive in remedy selection — even on the general plane between structural and conduct remedies — as is the populist philosophy. The economic efficiency approach necessitates careful scrutiny of the basic paradigm of industrial organization — structure, conduct, performance — to gain insights into the proper choice of a remedy.

In sum, the pure populist and pure economic efficiency perspectives have defined the boundaries of the antitrust policy debate. Since both traditions are deep-rooted, neither should be overlooked in the ultimate formulation of Sherman Act § 2 remedies. A policy which strikes blindly at mere size in formulating a remedy is intolerable. But neither can the noneconomic costs of large absolute size be ignored. Consequently, in remedy selection, the optimal welding of the two perspectives may result in the presumption that a remedy which leads directly to smaller economic entities ought to be favored over one that does not, so long as economic efficiency considerations do not clearly dictate otherwise.

B. *The Economic Efficiency Approach: Structure, Conduct, Performance*

Whether one accepts the economic efficiency theme with or without its populist overtones, it is prudent to examine industrial organization theory in pursuit of guidance toward remedy formu-

---

113 *See* United States v. United States Steel Corp., 251 U.S. 417, 451 (1920); United States v. Swift & Co., 286 U.S. 106, 116 (1932).

114 *See* text accompanying note 4 *supra*.

lation under § 2 of the Sherman Act. For even if it is desirable to overlay the economic dictates with populist trimmings, a well-reasoned theoretical model for antitrust policy is a necessary core. Given a general societal orientation against direct performance remedies,[115] the selection process demands an initial choice between only two general classes of remedies: those aimed directly at the structure of the industry in question; and those which attempt to limit the specific anticompetitive conduct which led to the § 2 monopolizing violation. In an *a priori* approach, either remedy would seem acceptable, as each is directed at one of the two elements which constitute the Sherman Act § 2 offense.[116] However, a general examination of the economic literature of

---

115 *See* note 49 *supra.*

116 There are essentially two elements to a monopolization charge under 15 U.S.C. § 2 (1970): the existence of market power, and the intent to increase that power as evidenced usually by certain elements of market conduct. *See generally* United States v. Grinnell Corp., 384 U.S. 563, 570-71 (1966); Kansas City Star Co. v. United States, 240 F.2d 643, 660 (8th Cir. 1957), *cert. denied,* 354 U.S. 923 (1957); United States v. E.I. du Pont de Nemours & Co., 351 U.S. 377, 389-94 (1956); Cal Distributing Co. v. Bay Distributors, Inc., 337 F. Supp. 1154, 1147 (M.D. Fla. 1971); Keco Industries, Inc. v. Borg-Warner Corp., 334 F. Supp. 1240, 1245 (M.D. Pa. 1971).

High market shares possessed by the defendant have been held to be sufficient evidence of monopoly power. United States v. Grinnell Corp., 384 U.S. 563 (1966) (87% market share was held to be a monopoly); American Tobacco Co. v. United States, 328 U.S. 781, 797 (1946) ("over two-thirds of the entire domestic field of cigarettes, and . . . over 80% of the field of comparable cigarettes" constituted a monopoly); United States v. Aluminum Co. of America, 148 F.2d 416, 424 (2d Cir. 1945) (90% share was sufficient; court had indicated that 60% or 64% might not be sufficient, and 33% "certainly" not). *But see* Stran Auto Sales Corp. v. World Wide Automobile Corp., 166 F. Supp. 313, 314 (S.D.N.Y. 1958) (56% share in 1955 that declined to 31% in 1957 sufficient to constitute a cause of action).

A finding of attempt to monopolize, although not directly at issue here, requires a "dangerous probability" of success, as well as proof of overt acts committed with the specific intent to destroy competition or to achieve monopoly power. Times Picayune Publishing Co. v. United States, 345 U.S. 594, 626 (1953); Lorain Journal Co. v. United States, 342 U.S. 143, 153 (1951); George R. Whitten, Jr. Inc. v. Paddock Pool Builders, Inc., 508 F.2d 547, 550 n.2 (1st Cir. 1974); Bernard Food Industries, Inc. v. Dietene Co., 415 F.2d 1279, 1284 (7th Cir. 1969), *cert. denied,* 397 U.S. 912 (1970); H.F. & S. Co., Inc. v. American Standard, Inc., 336 F. Supp. 110, 124 (D. Kan. 1972); Huron Valley Publishing Co. v. Booth Newspapers, Inc., 336 F. Supp. 659, 662 (E.D. Mich. 1972) (preliminary injunction denied); Cal Distributing Co. v. Bay Distributors, Inc., 337 F. Supp. 1154, 1157 (M.D. Fla. 1971); Keco Industries, Inc. v. Borg-Warner Corp., 334 F. Supp. 1240, 1245 (M.D. Pa. 1971). *But see* Hallmark Industry v. Reynolds Metals Co., 489 F.2d 8, 12 and n.3 (9th Cir. 1973) (dangerous probability or substantial market power need not be established where specific intent to set prices or exclude competition without legitimate business purpose existed).

industrial organization reveals the wisdom of presumptive use of structural remedies.

Disposition of the remedy selection problem depends, first, on an examination of the elements of the traditional industrial organization paradigm — structure, conduct, performance — and the links among these elements;[117] and, second, on a determination of the relative costs of the alternative approaches when applied in specific contexts. The structure-conduct-performance paradigm has framed the analytic and semantic debate concerning both standards for liability under the Sherman Act and guidelines for remedy selection and formulation.[118] The two dominant schools of thought that have emerged from this debate can be somewhat loosely categorized as the structuralist school and the conduct/performance school.[119]

1. The Structuralist School

The distinguishing feature of the structuralist school is its proposition that strong and unambiguous links exist between elements of market structure and ultimate economic performance.[120] Important elements of market structure which suggest the possibility of suboptimal economic performance include: fewness of sellers and buyers; large market shares in the possession of a few sellers or buyers; asymmetrical distribution of market shares among the largest firms; widespread vertical integration; substantial product differentiation; and significant barriers to entry.[121] Barriers to entry are those conditions in a particular industry which disadvantage potential entrants vis-à-vis established firms. They include the traditional economies of scale, absolute

---

117 See note 4 *supra.*

118 *See, e.g.,* Irwin & Barrett, *supra* note 102, at 38-39.

119 *Id.* at 39.

120 *See generally* J. BAIN, BARRIERS TO NEW COMPETITION (1956) [hereinafter cited as BARRIERS]; R. CAVES, AMERICAN INDUSTRY: STRUCTURE, CONDUCT, AND PERFORMANCE (1972); A. PHILLIPS, MARKET STRUCTURE, ORGANIZATION, AND PERFORMANCE (1962); F. SCHERER, *supra* note 5, at 400-11 (1970) (though Scherer tends to emphasize conduct elements somewhat more than the others, *see, e.g., id.,* at chs. 6-17); Mason, *The Current Status of the Monopoly Problem in the United States,* 62 HARV. L. REV. 1265 (1949); Mueller, *The New Antitrust: A "Structural" Approach,* 1 ANTITRUST L. & ECON. REV. 87 (Winter 1967).

121 *See* F. SCHERER, *supra* note 1, at 5, Figure 1.1.

cost advantages, and product differentiation advantages,[122] as well as high-intensity advertising[123] and large absolute size of existing firms in the industry.[124]

Adherents of both the structuralist and conduct/performance schools would probably accept this check-list of market structure elements as at least defining the relevant parameters of market structure. However, the structuralist argument continues by asserting that theoretical and empirically-supported links exist between these market structure elements and the various measures of economic performance — profit rates, prices, and rates of innovation.[125] Theoretically, in a market dominated by several firms, the few large sellers will be more inclined to raise prices, restrict output and thus reap monopoly profits than would their smaller counterparts in a market with a large number of sellers, each with an insignificant share.[126] This theoretical connection between market concentration and profits has been demonstrated empirically, especially in markets dominated by a single firm.[127] While

---

122 *See, e.g.*, BARRIERS, *supra* note 120, at 14.

123 Advertising has been discussed as a barrier to entry in numerous places in the literature. *See, e.g.*, Comanor & Wilson, *Advertising Market Structure and Performance,* 49 REV. ECON. & STATS. 423, 425-27 (1967); BARRIERS, *supra* note 120, at 65; Bok, *Section 7 of the Clayton Act and the Merging of Law and Economics,* 74 HARV. L. REV. 226, 276 (1960). *Contra*, Telser, *Advertising and Competition,* 72 J. POL. ECON. 537, 556-58 (1964). Telser's conclusion that the degree of market concentration is directly correlated with competitiveness is unsupportable because of his exclusion of several relevant market structure elements. *See* Disner, *Barrier Analysis in Antitrust Law,* 58 CORNELL L. REV. 862, 878 (1973); Brozen, *No "Scarlet Letters,"* 52 BARRON'S, Feb. 28, 1972, at 7, 8. Brozen views advertising as a means of entry despite the pervasiveness and scale of most consumer-oriented advertising. *See, e.g.*, Disner, *supra*, at 878-79, n.125; Porter, *Consumer Behavior, Retailer Power and Market Performance in Consumer Goods Industries,* 56 REV. ECON. & STATS. 419, 425-26, 435 (1974). Brozen's empirical results which show a negative correlation between firm size and advertising outlays measured as a fraction of sales are probably due to pecuniary economies of scale in advertising, rather than the vigor of present-day capitalism as Brozen suggests. Lecture by R. Caves in Economics 1550, Business Organization and Control, Harvard College, April 14, 1976.

124 *See, e.g.*, Edwards, *Conglomerate Bigness as a Source of Power,* in BUSINESS CONCENTRATION AND PRICE POLICY 331, 334-35 (1955).

125 *See* Irwin & Barrett, *supra* note 102, at 39-40.

126 *See* United States v. Philadelphia National Bank, 374 U.S. 321, 363 (1963) for a succinct judicial statement of this hypothesis: "Competition is likely to be greatest when there are many sellers, none of which has any significant market share." *See generally* E. CHAMBERLIN, THE THEORY OF MONOPOLISTIC COMPETITION 30-55 (8th ed. 1962); F. SCHERER, *supra* note 1, at 131-57. *Cf.* J. MCGEE, IN DEFENSE OF INDUSTRIAL CONCENTRATION 54-74 (1971).

127 *See* F. SCHERER, *supra* note 1, at 183-86; W. MUELLER, *supra* note 9, at 97-101.

the data from oligopolistic industries lends more qualified support to this concentration-profit linkage,[128] some experts believe that monopoly profits are virtually inevitable in any type of highly-concentrated market.[129] In sum, most empirical studies show a substantial correlation between market concentration and profits.[130]

It should be noted, however, that criticism of these studies has been offered on two general grounds. First, it has been suggested that methodological problems exist in the techniques employed to measure profits.[131] Second, an alternative explanation of the high profit levels found in the concentrated industries has been offered, *viz.*, returns on innovation.[132] Yet it can be said in response that, although none of these studies can be said to reveal an incontestable correlation between profitability and concentration,[133] many economists have vouched for the accuracy of the studies.[134]

---

128 *See, e.g.*, Posner, *Oligopoly and the Antitrust Laws: A Suggested Approach*, 21 STAN. L. REV. 1562, 1578 (1969); G. STIGLER, THE THEORY OF PRICE 230-38 (3d ed. 1966). *See* text accompanying note 42 *supra* for discussion of debate over liability standards concerning the oligopoly problems.

129 *See, e.g.*, Neal Report, *supra* note 1, at 28; Turner, *The Definition of Agreement Under the Sherman Act: Conscious Parallelism and Refusals to Deal*, 75 HARV. L. REV. 655, 660-66 (1962).

130 The most important studies showing a positive correlation between concentration and profits include: Weiss, *The Concentration-Profits Relationship and Antitrust*, in H. GOLDSCHMID, *supra* note 9, at 184-233; N. COLLINS & L. PRESTON, CONCENTRATION AND PRICE-COST MARGINS IN MANUFACTURING INDUSTRIES (1968); G. STIGLER, CAPITAL AND RATES OF RETURN IN MANUFACTURING INDUSTRIES 54-70 (1963); BAIN, RELATION OF PROFIT RATE TO INDUSTRY CONCENTRATION: AMERICAN MANUFACTURING, 1936-1940, 65 Q. J. OF ECON. 293 (1951); Comanor & Wilson, *supra* note 123; Shepherd, *The Yields from Abating Market Power*, 1 INDUSTRIAL ORG. REV. 47 (1973). Weiss attempts a crude test of the dominant firm theory, which has not, as yet, been the subject of a systematic statistical examination. He finds support for the theory, thereby giving credence to the Hart, Neal, and Kaysen-Turner market share standards for connecting one company to market power. *See* note 45, *supra*.

131 *See* F. SCHERER, *supra* note 1, at 185, col. 2; G. STIGLER, *supra* note 130, at 67-68. *Cf.* Kilpatrick, *Stigler on the Relationship Between Industry Profit Rates and Market Concentration*, 76 J. POL. 470, 480, 486 (1968). *See generally* Brozen, *Significance of Profit Data for Antitrust Policy*, 14 ANTITRUST BULL. 119, 124-30 (1969).

132 *See, e.g.*, Brozen, *The Antitrust Task Force Deconcentration Recommendation*, 13 J.L. & ECON. 279 (1970). *But see* Wenders, *Deconcentration Reconsidered*, 14 J.L. & ECON. 485 (1971); MacAvoy, McKie & Preston, *High and Stable Concentration Levels, Profitability, and Public Policy: A Response*, 14 J.L. & ECON. 493 (1971).

133 *See, e.g.*, sources cited by F. SCHERER, *supra* note 1, at 186, col. 1.

134 *Id.*

Arguably, one may explain the lack of a perfect correlation between concentration and profits by the presence of other structural elements which affect market performance.[135] The most important of these are barriers to entry. While the level of present competition reflected in market share analysis is indeed significant in explaining market performance, the negative effects of market concentration can be offset to some degree by the threat of potential competition.[136] Profit levels may be held down so as not to attract new entrants whose presence would decrease market shares, stimulate competition, and very probably reduce profit margins industrywide. By definition, barriers to entry reduce the threat by disadvantaging the potential entrants vis-à-vis the established firms.[137] Thus, where high levels of seller concentration exist without significant barriers to entry, it is unlikely that the sellers will behave anticompetitively to a significant degree.[138] Where seller concentration coexists with high or even moderate barriers to entry, it is more likely that anticompetitive market behavior, as manifested in supernormal profit, will be observed.[139]

However, this theoretical quandary over the relative importance of seller concentration versus barriers to entry as an explanatory variable is mooted to a great extent by the existence of a high positive correlation between substantiality of market shares and the height of entry barriers.[140] In most cases, it would seem that

---

135 Among these other factors are product heterogeneity, variations of cost from firm to firm, fluctuation in the size of buyers' orders, secrecy of transactions, and industry social structure. *Id.* at 186-98, 206-12. Mueller discussed the effects of barriers to entry, including advertising expenditures. *See* W. MUELLER, *supra* note 9, at 99-100.

136 *See, e.g.*, Greenhut, *An Economic Theory for Use in Antitrust Cases*, 7 HOUSTON L. REV. 318, 323-35 (1970) (Professor Greenhut's methods are somewhat exotic: *i.e.*, he uses spatial economics to show that the prevention of collusive agreements and the elimination of entry barriers are the most important factors inducing competitive behavior in oligopolistic industries).

137 Report of the Attorney General's National Committee to Study the Antitrust Laws 324 (1955); *cf.* F. SCHERER, *supra* note 1, at 10.

138 *See* Disner, *supra* note 123, at 864-65. The situation where anticompetitive behavior occurs in the absence of high levels of concentration with or without high entry barriers is not at issue here; where a defendant is not shown to possess a large market share a Sherman Act § 2 monopolization violation would be difficult to establish.

139 *See* Disner, *supra* note 123, at 864.

140 *See* J. BAIN, *supra* note 120, at 299. Examples of high seller concentration existing with low barriers to entry are the tire, gypsum products, and metal con-

reliance on high market shares as the principal structural variable determining poor market performance is justified.

Thus, market share, as evidenced by its high correlation with profits levels, is the chief indicator of market power.[141] Since such profit characteristics signal restricted output and higher than optimal prices,[142] it would seem prudent for antitrust policy to attack directly the root of the socially undesirable result. For purposes of remedy selection after a monopolization violation has been established, such a prescription would order, in the first instance, full consideration of divestiture as an appropriate corrective measure.

Yet employment of the divestiture remedy cannot be totally automatic. At least two factors must be considered further. First, the benefits of reduced market power to consumers resulting from divestiture must be balanced against the accompanying potential economic costs.[143] Further, the costs and benefits of alternative remedies, in particular conduct-prohibitory and regulatory injunctions, must be evaluated in light of the potential net beneficial effect of divestiture. In assessing these considerations in past § 2 cases, the courts generally have decided against divestiture.[144] Reexamination of the underpinnings of those decisions, however, reveals that this anti-divestiture bias is ill-founded.[145]

A second important factor to consider when estimating the benefits of reduced market power is the extent to which market power might dissipate without direct intervention. For example, if market concentration decays rapidly over time, divestiture will have only a superfluous positive effect on market performance.[146] However, where market power dynamics indicate relative stability of market positions, divestiture would realize its maximum beneficial impact.[147] It has been argued that the mere persistence of

---

tainers industries. Mann, *Seller Concentration, Barriers to Entry, and Rates of Return in Thirty Industries, 1950-1960*, 48 REV. ECON. & STATS. 296, 299-300 (1966).

141 *See* W. SHEPHERD, THE TREATMENT OF MARKET POWER 96 (1975); studies cited at note 130 *supra*.

142 See text accompanying note 130 *supra*.

143 See text accompanying note 9 *supra*. Note that there is a distinction between economic costs (adverse effects on the firm's economic performance) and costs to shareholders, creditors and others.

144 See text accompanying notes 79-87 *supra*.

145 See text accompanying notes 170-249 *infra*.

146 *See* W. SHEPHERD, *supra* note 141, at 66-67, Figure 3.2.

147 *Id.*

market power ought to provide a sufficient basis for a monopolization violation.[148] Generally, market power in the United States appears to be a relatively stable phenomenon, with the already low rate of decline decreasing in recent years.[149] Dominance does not appear to dissipate naturally in most American industries.[150] This empirical finding is consistent with the prior observation of high correlation between market concentration and barriers to entry.[151] Where substantial concentration is protected by barriers to entry, disintegration of market power is highly unlikely. Because of the stability of market power, then, it is manifest that, if structural alteration is the prescribed remedy under Sherman Act § 2, direct intervention must occur to effect the desired alteration in the market structure.

2. The Conduct/Performance School

The conduct school, in contrast to the structuralist position, deemphasizes the general explanatory power of market structure elements on market performance and tends to concentrate on the specific types of conduct of particular business entities. This approach implies that, since understanding of the links between market structure and market performance is not perfect, appraisal of the effects of behavior of particular enterprises is necessary.

---

148 See note 41 *supra* for discussion of this proposed reinterpretation of the Sherman Act § 2 liability standards.

149 *See* W. SHEPHERD, *supra* note 141, at 113-29. Professor Shepherd notes that the "natural" decline of the market power of leading dominant firms was much more rapid in 1910-1935 than in 1948-1973. *Id.* Shepherd offers the following explanations for this change: (1) the increased transferability of banker oligopoly market power to the industrial sector, *id.* at 118; (2) antimerger policy which prevents the rise of substantial competitors in concentrated industries, *id.* at 118-19; and (3) the possibility of economies of scale, *id.* at 119-20. (*But see* sources cited at note 9 *supra.*)

Professor Shepherd has devised a model which predicts variations over time in market share and profitability. The model suggests that firms may have control over the rate at which they trade market share for profits. *Id.* at 50-57.

150 *See* Williamson, *Dominant Firms and the Monopoly Problem: Market Failure Considerations,* 85 HARV. L. REV. 1512, 1514 n.11 (1972). *See also* MacAvoy, McKie, & Preston, *supra* note 132. *See also* W. SHEPHERD, *supra* note 141, at 116-18, where he notes that American firms have appeared less susceptible to decay of market power from both natural causes and public policy measures than have comparable British firms. *But see* Brozen, *supra* note 132, at 286; Posner, *supra* note 128, at 1597.

151 *See* Irwin & Barrett, *supra* note 102, at 40.

These conduct-oriented economists view specific firm behavior as an independent variable, rather than as a mere conduit through which structural determinants specify performance characteristics.

The influence of such a perception on remedy formulation under § 2 of the Sherman Act is apparent. Under present practice, the conduct in question is generally that conduct which formed the basis for the monopolization violation.[152] Where the conduct is found to be clearly anticompetitive, it is generally prohibited.[153] Where the conduct is found questionable but not clearly reprehensible, the court may simply regulate the conduct in question.[154] Only where it is obvious that the conduct in question cannot be enjoined effectively, will the courts seriously consider the imposition of a structural remedy.[155]

The crucial distinguishing feature of the conduct approach to remedy selection is its conviction that the particular conduct mode drawn into question by the monopolization charge should be examined in minute detail.[156] Although the conduct enthusiasts are well aware of the substantial theoretical and empirical support for the structuralist approach,[157] they argue that only by examining each conduct mode as it appears in a particular industry context can its value vis-à-vis economic performance be ascertained. Such an approach views each monopolization situation as a probable exception to the general rule that market structure directly affects performance through relatively indeterminable and unimportant conduct modes. Additionally, it has been observed that most existing market structures provide at least modestly competitive frameworks, so that remedies designed to eliminate anticompetitive conduct can assure satisfactory performance.[158]

---

152 *See, e.g.,* cases cited in notes 260-98 and text accompanying, *infra,* in which the courts have argued that the use of divestiture ought to be limited to the property involved in the particular monopolization offense, since the abusive use of that property provided the conduct element necessary for the Section 2 violation.

153 *See generally* Brown, *Injunctions and Divestiture,* in THE ABA SECTION ON ANTITRUST LAW HANDBOOK 540-43 (1958).

154 *Id.* at 540-41.

155 See text accompanying notes 300-39 *infra,* discussing necessity requirements.

156 *See* Irwin & Barrett, *supra* note 102, at 40.

157 *See* J. DIRLAM & A. KAHN, FAIR COMPETITION: THE LAW AND ECONOMICS OF ANTITRUST POLICY 34 (1954).

158 *See id.* at 31, 33, 49.

The economic rationale of the law rests on two assumptions. The first

Yet dealing with specific conduct modes per se, and neglecting the accompanying structural determinants, involves a serious functional shortcoming. For in any particular instance, a conduct practice which is generally reprehensible may have desirable effects. It must be noted that most of the evidence upon which the conduct school relies is drawn from that inexhaustible bundle of industrial organization literature called "case studies."[159] In the instances cited, the attacked behavior indeed may be anticompetitive. But analysis of such conduct in industries with varying structural settings may reveal that the specific practice under examination has differential effects on ultimate performance, and that such variations can be explained only by the particular structural configurations.[160]

In sum, the conduct school views the behavior which links market structure to ultimate performance as a significant independent determinant rather than a simple conduit. Hence, where conduct which led to the Sherman Act § 2 violation is found to be justifiable on general welfare principles, controllable through specially tailored equitable remedies, or destructible through blanket injunctive relief, advocates of the conduct approach will not presume to tamper with market structure. They view such

---

> is that the will to "get ahead," to outdo others, in short, to compete, is so strong and so widespread that it needs only to be channelized by negative prohibitions. The second is that functions and optimum business size in most industries are such that . . . the number of sellers (and buyers) emerging will not be so small as to weaken seriously the force of competition in the market. These assumptions have often been questioned but seldom refuted on the basis of concrete examinations of the structural pattern and performance of specific industries. It follows that the law need only prevent the deliberate impairment, misdirection, or suppression of competition to protect both the public interest and the legitimate interests of business competitors.

Id. at 28.

159 *See, e.g.*, J. BACKMAN, CHEMICAL PRICES, PRODUCTIVITY, WAGES, AND PROFITS (1964); J. MARKHAM, COMPETITION IN THE RAYON INDUSTRY (1952); A. PHILLIPS, TECHNOLOGY AND MARKET STRUCTURE — A STUDY OF THE AIRCRAFT INDUSTRY (1971); *Hearings on Competitive Problems in the Drug Industry Before the Subcommittee on Monopoly of the Senate Select Comm. on Small Business*, 90th Cong., 1st Sess., pt. 5 (1967).

160 See text accompanying notes 120-51 *supra*.

action as excessive and severe, appropriate only in instances where efforts to properly circumscribe conduct have failed.

3. Market Conduct as a Determinant of Market Structure: A Hybrid of the Conduct Approach

A variant of the conduct approach has been espoused by those who see a causal link between some forms of market conduct, especially barriers to entry, and various structural elements.[161] In general, the direction of causality has been viewed as running from structure, through conduct, to performance. However, the existence of a feedback loop between market conduct and barriers to entry is widely accepted.[162] Product differentiation strategies, the building of excess capacity,[163] mergers, advertising, and predatory conduct can raise barriers to entry. Also, such conduct may constitute the necessary intent or conduct element of a monopolization charge.[164] This "notion is inherent in the antitrust laws of the United States, particularly in § 2 of the Sherman Act and amended § 7 of the Clayton Act."[165] Adherents of the conduct-to-structure approach to antitrust policy bemoan the framing of much of the debate over remedies in terms of conduct-oriented injunctions versus direct structural attacks.[166] As a result, they argue, the traditional industrial organization literature has unnecessarily segregated the two general classes of relief and has ignored the impact of conduct adjustments on market structure.[167]

However, there are certain deficiencies with this modified conduct approach. For the most part, such shortcomings reflect the general objections raised to the basic conduct approach, and therefore will be considered in the following section which directly

161 *See* Baldwin, *supra* note 49, at 126-27. *Cf.* Disner, *supra* note 123, at 863-70.

162 *See* Baldwin, *supra* note 49, at 126; R. CAVES, *supra* note 120, at 49-50.

163 *See* United States v. Aluminum Co. of America, 148 F.2d 416 (2d Cir. 1945). Where barriers to entry are based upon economies of scale, few would argue that economies ought to be traded off to encourage new entry into an industry. *See, e.g.,* C. KAYSEN & D. TURNER, *supra* note 23, at 78, 114; F. SCHERER, *supra* note 1, at 233.

164 *See* J. DIRLAM & A. KAHN, *supra* note 157, at 28.

165 *See* Baldwin, *supra* note 49, at 126.

166 *Id.* at 127.

167 *Id. But see* F. SCHERER, *supra* note 1, at 324-45.

compares the relative merits of the structural and conduct approaches.

IV. Comparative Analysis of Structural and Conduct Remedies

The idea that a remedy can be formulated which will avoid the supposed harshness of divestiture but affect indirectly the structure of an industry has definite appeal.[168] Yet there are serious problems with excessive reliance on conduct remedies in general, and with that genre of conduct remedies designed to affect market structure in particular. Although not completely without shortcomings, structural relief, especially divestiture, offers antitrust policy-makers an effective and feasible mechanism with which to attack undesirable conduct thereby improving market performance.

Examination and comparison of the two general classes of remedies along four vital dimensions — effectiveness in improving overall market performance, impact on technical efficiencies of the individual firm(s), consequences to third parties with a direct pecuniary interest in the target firm(s), and extent of judicial resources required to enforce the remedy — reveal that structural relief is generally superior and should be the presumed remedy in Sherman Act § 2 cases. The second and third categories, and possibly the fourth, are often considered together under the general rubric of "harshness." Even though certain facets of this harshness grouping are only of secondary importance, the analysis which follows suggests how divestiture has been much maligned but often misunderstood when evaluated from this perspective.[169]

---

168 See text accompanying notes 161-65 *supra*.

169 Celler, *The Trial Court's Competence to Pass Upon Divestiture Relief*, 10 ANTITRUST BULL. 693, 695 n.16 (1965). *See also*, J. DIRLAM & A. KAHN, *supra* note 157, at 170.

There remains a strong prejudice against dissolution of integrated firms, even when, like United Shoe Machinery Company, they are almost complete monopolies. In *United Shoe*, Judge Wyzanski was unwilling to dissolve the company into three independent manufacturing entities because he could not see how the physical difficulties of independent operation could be overcome. The company's operations were concentrated in one plant (although not in one building) in Beverly, Massa-

A. *Effectiveness*

The effectiveness of the relief employed ought to be measured by the expected success of the remedy in achieving the goal of § 2 enforcement, namely, improved market performance.

1. Structural Remedies

Since even the conduct school admits to the causal effect of structural elements on market performance, there is little dispute as to the ability of structural remedies to govern market performance. Reduced market concentration generally leads to dissipation of market power, and hence to dilution or termination of the ability of the target firm or firms to steer prices and output from the competitive norm. Although "fine-tuning" of divestiture may be difficult, this difficulty should not be overemphasized: since the competitive norm envisions an unlimited number of firms within a market, it is difficult to deconcentrate excessively.

Reducing constituent firm size does have various technical efficiency effects and consequences for third parties, but these considerations are of secondary importance. In the subsequent exploration of these secondary criteria, it will become apparent that deconcentration may be sufficiently effective long before it reduces technical efficiencies or unfairly affects third parties.

2. Conduct Remedies

Conduct remedies, whether directed primarily at performance results or indirectly at market structure changes, tend to be ineffective. Both the theory and empirical work aimed at assessing the

---

chusetts; it had one set of jigs and tools, one foundry and one laboratory. However, from the facts stated in the opinion, it would appear that the court in this case was overawed by the difficulties that supposedly attend any change in business organization. The same timidity found expression in the first *U.S. Steel* and the *Hartford-Empire* and *Timken* decisions. Although private business management in the United States has been extraordinarily flexible in meeting emergencies and devising novel methods of operation and organization, the typical judicial attitude in dissolution cases is that reversal or change of present methods of operation would cause incalculable loss. See text accompanying notes 88-96 *supra* for a discussion of the reasons for the timidity of the federal judiciary in granting divestiture. See text accompanying notes 300-48 *infra* for a discussion of the courts' use of the "necessity" and "practicality" requirements for granting divestiture.

effect of this type of remedy, as well as practical problems inherent in their implementation, suggest that conduct remedies in general are, at best, unpredictable and, at worst, detrimental to market performance. While a particular type of behavior can be enjoined in response to a finding of monopolization, often a firm with monopoly power — which, as has been noted, is an element of a monopolization violation — can adopt alternative patterns of behavior to effectuate its market power.[170] While this behavioral pattern applies with brutal force in the dominant firm case,[171] it also emerges in closely-knit oligopolistic markets. For example, where members of a tightly knit ologopoly are found in violation of the Sherman Act § 2 on the basis of their shared monopoly power[172] and exchange of price or other data,[173] or tacit collusion,[174] a court operating under the conduct/performance framework would simply enjoin the conduct involved, allowing the market power of the oligopolists to remain intact. However, it is clear that the firms involved would have a bevy of alternative conduct modes available to continue the exercise of their market power. They could engage in restrictive product and know-how licensing,[175] use indirectly interlocked directorates to coordinate pricing decisions,[176] engage in persistent price discrimination,[177] or engage in purchase and sale, or other restrictive trade arrangements.[178] Theoretically, the problem described has its basis in the

170 *See generally* F. SCHERER, *supra* note 1, chs. 6-17.

171 *Id.* at 164-66, 216-38.

172 Such a theory has formed the basis for at most one case in the history of § 2. *See* United States v. American Tobacco Co., 221 U.S. 106 (1911). But tacit collusion could form such a basis for a violation of § 2. *See* Cox, *Competition and Section 2 of the Sherman Act,* 27 A.B.A. ANTITRUST L.J. 72, 82 (1965). *See also,* Turner, *supra* note 41, at 1228-31; Ward *supra* note 41, at 606 n.20. See note 42 *supra.* See also note 44 *supra* where the FTC's suit against the four leading cereal manufacturers is discussed.

173 *Cf.* United States v. Container Corp. of America, 393 U.S. 333 (1969) (case involved Sherman Act § 1, 15 U.S.C. § 1 (1970), violation rather than § 2). *See generally* Campbell & Emanuel, *A Proposal for a Revised Price Discrimination and Predatory Pricing Statute,* 13 HARV. J. LEGIS. 125, 142-48 (1975).

174 *See* United States v. American Tobacco Co., 221 U.S. 106 (1911).

175 *See* Walker Process Equip., Inc. v. Food Machinery and Chem. Corp., 382 U.S. 172 (1965).

176 *See* REPORT OF THE F.T.C. ON INTERLOCKING DIRECTORATES 17-27 (1951).

177 *See* Puerto Rican American Tobacco Co. v. American Tobacco Co., 30 F.2d 234 (2d Cir. 1929).

178 *See* FTC v. Consolidated Foods Corp., 380 U.S. 592 (1965). *See also,* Ammer,

lack of recognition of the clear-cut links between particular elements of market structure and particular modes of conduct effectuating that market power. For market power — the ability to raise prices, restrict output, and in general, depart from the competitive ideal[179] — is the source of the problem, not the specific form of conduct used to transform the market power into a tangible reward for those possessing such power.[180] This criticism of conduct-related remedies applies not only in paradigmatic oligopoly cases, but, with added weight, in dominant firm cases.[181]

Conduct remedies are likely to be ineffective not only because of ease of evasion, but also because the consequences of such conduct-oriented relief are basically unpredictable.[182] This unpredictability results because the effects of various types of conduct on market performance are themselves indeterminate.

A good example of this indeterminacy is advertising.[183] It is widely recognized that advertising may raise substantial barriers to entry by forcing a potential entrant to expend considerable resources on the development of "brand names."[184] Advertising may also produce allocative inefficiency by misleading consumers as to the "true value" of the product advertised.[185] Yet advertising has been viewed by some as an acceptable form of rivalry among firms in an industry where there are relatively few competitors.[186]

Realistic Reciprocity, Purchasing Power and Competition, 48 MICH. L. REV. 523 (1964).

179 *See* Baldwin, *supra* note 49, at 145 for discussion of practical advantages of this definition of market power.

180 Bain's emphasis on the links between market structure and performance and on the intervening structure-conduct and conduct-performance links is based on three factors. The most important is the great ambiguity of prediction of the causal links between structure and conduct, and conduct and performance. J. BAIN, INDUSTRIAL ORGANIZATION 295-301 (2d ed. 1968). The other two factors are: (1) the recognition that the inclusion of conduct elements is not necessary to an operational theory of industrial organization, *id.* at 36-38; and (2) the serious empirical problems one would encounter in testing hypotheses including such elements, *id.* at 310-15.

181 See text accompanying notes 171-78 *supra.*

182 *See* Baldwin, *supra* note 49, at 146-50.

183 *See generally* F. SCHERER, *supra* note 120, at 324-45.

184 In a sense, advertising raises both absolute cost and product differentiation barriers to entry. *Id.*

185 *See* Disner, *supra* note 123, at 916-17.

186 *See generally* Doyle, *Economic Aspects of Advertising: A Survey,* 78 ECON. J. 570-602 (1968).

A court which attempted to assess the potential impact of an injunction prohibiting certain types of advertising by a monopolizer would face the unenviable task of attempting to distinguish between the informational (and hence procompetitive) and the persuasive (and hence misallocative and barrier-creating) content of specific advertising practices. Such a line is not easy to draw.[187] Since advertising to the "optimal" level increases rivalry among the dominant firms in an industry (as well as among their potential competitors), the enforcement of an industry-wide ceiling below that level would actually enhance the market power of the defendant.[188] Thus, while anti-advertising remedies must be included in the court's tool-kit of available relief, and in fact may be the indicated measure in particular cases,[189] they well illustrate the proposition that, in general, conduct-oriented relief has unpredictable and possibly harmful effects. Thus, conduct remedies are generally ineffective not only because they do not attack the ultimate source of poor performance, but also because they differentially impact those conduct elements at which they are directed.

B. *Technical Efficiencies*

The differential effects of divestiture and specific conduct injunctions on individual firm efficiencies are widely mischaracterized. The misperception that divestiture is inherently destructive of existing production advantages and the assessment that direct structural intervention results in unnecessary harm to third parties, leads to divestiture's reputation for "harshness."

Chief among criticisms leveled at divestiture for its anti-efficiency tendencies is that deconcentration will adversely affect real econ-

---

187 *See* Disner, *supra* note 123, at 917. *Cf.* Bork, *Contrasts in Antitrust Theory: I*, 65 COLUM. L. REV. 401, 411 n.11 (1956).

188 See note 186 *supra*.

189 This is especially the case in merger suits. *See, e.g.*, FTC v. Procter & Gamble Co., 386 U.S. 568, 573 (1967) (existence of economies in advertising was one factor inducing the Court to find a merger illegal). Arguably, restrictions on advertising could be applied against firms in violation of the Sherman Act § 2. *Cf.* United States v. United Shoe Machinery Corp., 110 F. Supp. 295, 350 (D. Mass. 1953). *See also* Disner, *supra* note 123, at 917-18, for a discussion of this possibility. See text accompanying notes 65-75 *supra* for discussion of the broad discretion district courts have in formulating remedies in Sherman Act cases.

omies of scale achieved by the target firm.[190] Many courts have viewed the presence of any economies of scale as a complete bar to divestiture.[191] This problem has been compounded by the misperception that from the large absolute size of a defendant corporation one can infer the presence of substantial economies.[192] Yet, in prescribing divestiture for large firms under § 2, fear of resultant diseconomies generally is unwarranted, since only in exceptional cases will diminution of existing scale lead to inefficiencies. Professor Scherer has demonstrated that economies of scale are not as pervasive in the American economy as were once supposed.[193] His studies have revealed that high market concentration in American manufacturing industries cannot be justified in general on the basis of plant economies,[194] product specific econ-

---

190 *See generally* Baldwin, *supra* note 49, at 138-41. See note 9 *supra.*

191 *See, e.g.,* United States v. United Shoe Machinery Corp., 110 F. Supp. 295 (D. Mass. 1953), *aff'd per curiam,* 347 U.S. 521 (1954).

192 *See generally* W. SHEPHARD, MARKET POWER AND ECONOMIC WELFARE 37 (1970); McGee, *Efficiency and Economies of Size,* in GOLDSCHMID, *supra* note 130 at 55. *But see* Scherer, *Economies of Scale and Industrial Concentration,* in GOLDSCHMID, *supra* note 130, at 28.

193 *See generally* Scherer in GOLDSCHMID, *supra* note 130, at 22-54. Scherer casts his discussion of economies in terms of the minimum optimal size (MOS) of the economic entity as a percentage of the market demand for the good in question. *Id.* at 19-22. MOS is "conventionally defined as that scale at which scale economy opportunities are first fully exhausted and the unit cost curve becomes horizontal" or non-declining. *Id.* at 19.

Multiplant economies of scale tend to be present for two reasons. First, where there is no competitive market for the buying and selling of supplies, a multiplant operation may enable a firm to simulate the competitive market operation on an intrafirm basis. *Id.* at 39. However, this benefit is dependent upon the breakdown of the competitive market situation and is really only a second-best solution, *see generally* Lipsey & Lancaster, *The General Theory of Second Best,* 24 REV. OF ECON. STUD. 11 (1956), and thus would be non-existent where antitrust policy was effective in restoring and maintaining competition.

The second explanation for multiplant economies, "the possibility of holding smaller capacity reserves against random regional demand fluctuations when one plant confronted by booming demand can be helped out by another with excess capacity," Scherer, *supra* at 39-40, was found not to be very significant at all. *Id.* at 40.

194 *See, e.g., id.* at 22-33. Scherer compares the results of several studies concerning plant scale economies and concludes that ". . . nationwide oligopoly and high seller concentration cannot be viewed primarily as the inevitable consequence of production scale economies at the plant level." *Id.* at 28. The studies he compares include, J. BAIN, BARRIERS TO NEW COMPETITION 72, 84 (1956); C. PRATTEN, ECONOMIES OF SCALE IN MANUFACTURING INDUSTRY 269-77 (1971); F. SCHERER, A. BECKENSTEIN, E. KAUFER, & R. MURPHY, THE ECONOMICS OF MULTI-PLANT OPERATION: AN INTERNATIONAL COMPARISONS STUDY ch. 3 (1975).

omies,[195] or other economies of multiplant firm scale.[196] But perhaps more importantly, Scherer's work implies that courts can measure the importance of economies of scale as an explanatory variable of market concentration.[197]

Given this ability, courts, in devising § 2 remedies, can include such technical efficiency losses as economies of scale in balancing the costs of divestiture against the expected benefits of greater competition within the industry in question.[198] Such a balancing

---

195 Product specific economies are defined as the per unit cost savings realized from longer production runs — due ultimately to the high fixed set-up costs of a production line and the cumulative growth of operative skills and production engineering know-how as the length of the production run increases. Scherer, in GOLDSCHMID, *supra* note 130, at 33. Scherer concludes that product-specific economies do not offer an explanation (or justification) for high levels of market concentration. In those industries where such economies would be expected, small and moderate-sized firms tend to have production runs as long as those of the larger firms. The difference in size was attributable to the wider product range carried by larger firms. *Id.* at 36-37.

196 The other types of economies investigated by Scherer and found to be lacking in explanatory value vis-à-vis the high concentration ratios in American industry were capital-raising economies, *id.* at 41-42, procurement economics, *id.* at 42-43, sales promotional and market access advantages, *id.* at 43-46, technological innovation economies, *id.* at 46-50, and management economies, *id.* at 50-51. *See generally id.* at 52, Table 6 for a summary of the overall effect of these types of economies.

197 *See, e.g.,* Scherer, in GOLDSCHMID, *supra* note 9, at 18-19. Professor Scherer describes the three main methods for estimating economies of scale: statistical cost studies, the survivor method, and the engineering method. Statistical cost studies involve questioning businessmen about the per unit costs they are experiencing. Data on costs, outputs, and other characteristics of plants varying widely in size are assembled and then analyzed. *Id. See also* studies cited in *id.* at 18 n.2. The survivor technique relies on the abundant data obtainable from the *Census of Manufactures* to analyze the ebb and flow of activity at different entity class sizes. Sizes where activity is declining are deemed suboptimal or of excessive scale whereas sizes with increasing activity are of optimal scale. *Id. See also* Shepherd, *What Does the Survivor Technique Show About Economies of Scale?,* 34 SOUTHERN ECON. J. 113 (1967). Finally, the engineering method relies on the expert opinions of those responsible for the actual choice of plant size to estimate the average cost of production at each plant size. Scherer, *supra,* at 18.

198 A similar balancing approach is utilized in Clayton Act § 7 cases. *See, e.g.,* United States v. E.I. du Pont de Nemours Co., 366 U.S. 316, 327-333 (1960). The "unmerger" cases, however, often provide only collateral precedent for the real economic costs involved in a divestiture order since generally integration of actual physical plant and firm operation may not take place at least until several years after the initial merger. *See generally,* Fraidin, *supra* note 66, at 906-11. It is interesting to note that some statutorily imposed schemes of divestiture have gone even further in that they simply order divestiture if it appears that some public purpose will be served — unless the defendant affirmatively proves that there will be substantial costs to the public as a result of divestiture. *See, e.g.,* the litigation under

procedure is appropriate for the courts under the Sherman Act[199] as well as under the major legislative deconcentration proposals.[200]

In implementing this cost-benefit approach, the courts ought to

---

the Public Utility Holding Company Act of 1935, 15 U.S.C. § 79j (1970). Divestiture is the contemplated remedy under this act. 15 U.S.C. § 79k (1970). Divestiture will not be enforced when the defendant can prove that divestiture: (A) would cause loss of substantial economies; (B) that all of the holdings of the holding company are in one state, or in adjoining states; and (C) that the continued combination of the additional systems under the holding company's control is not so large as to impair the advantages of localized management, efficient operation, or the effectiveness of regulation. *See* SEC v. New England Electric System (NEES), 390 U.S. 207 (1968). A complete history of this case follows: 38 S.E.C. 193 (1958) (SEC brings suit against NEES); 41 S.E.C. 888 (1964) (SEC orders gas system divided; no loss of substantial economies), *rev'd*, 346 F.2d 399, 406 (1st Cir. 1965) (economies decision ought to be based on business judgment of defendant), *rev'd*, 384 U.S. 176 (1966) (Court adopted SEC test), 376 F.2d 107 (1st Cir. 1967) (there is not sufficient evidence for SEC finding under SEC test), *rev'd*, 390 U.S. 207 (1968) (supporting SEC findings).

199 Note that the Court's balancing process is complicated by unavoidable value judgments regarding the social desirability of the particular function in which such economies exist. A good example of this non-economic evaluation process is the necessary assessment of the value or harm of brand-name advertising. See notes 183-88 *supra*. *Compare Foreword*, 4 ANTITRUST L. & ECON. REV. 1-2 (Spring 1971), with Posner, *Advertising and Product Differentiation* in REPORT OF THE TASK FORCE ON PRODUCTIVITY AND COMPETITION (1969), *reprinted in* 115 CONG. REC. 15933, 15940 and in 2 ANTITRUST L. & ECON. REV. 13, 47 (Spring 1969). If a defendant attempted to oppose divestiture relief on the basis that such relief would prevent the realization of economies of scale in advertising, the court (or any special industrial court created under the various legislative deconcentration proposals, *see* Comment, *The Industrial Reorganization Act: An Antitrust Proposal to Restructure the American Economy*, 73 COL. L. REV. 635, 669 (1973)) would have to assess the social value of the advertising as well as the economies that might be present. In forbidding the merger of Clorox with Procter & Gamble, the Supreme Court sacrificed economies in advertising. FTC v. Procter & Gamble Co. (Clorox), 386 U.S. 568 (1967). *Compare* Bork, in Blake, Bork, Bowman & Jones, *The Goals of Antitrust: A Dialogue on Policy*, 65 COLUM. L. REV. 363, 411 n.11 (1965) *with* Blake & Jones, in *id.*, at 460 n.144.

200 *See* Blake's discussion of the comparative treatment of economies in the three major deconcentration proposals mentioned at note 45 *supra*. He argues that the Kaysen-Turner approach — which

> requir[es] the court to undertake the apparently more complex (or impossible) task of balancing claimed economies against other factors — may be more amenable to judical process than approaches which require that it define, more or less in the abstract, relevant market boundaries or the substantiality of economies. Judges are accustomed to balancing "equities" to arrive at a workable result, even though they would be unable accurately to quantify any of them. . . . A good case can thus be made for providing both the commission and the court with more flexible rather than unduly rigid criteria for illegality and for remedy; a good case can also be made that the two should be dealt with simultaneously.

Blake, *Legislative Proposals for Industrial Deconcentration*, in GOLDSCHMID, *supra* note 130, at 358.

maximize the anticipated benefits from increased competition realizable by creating entities through the divestiture order which are both independent and viable. The analysis of these benefits depends upon the choice of the specific assets involved in the divestiture decree. The courts must strive to design remedies that effectively guarantee the independence of the severed portions of the defendant, assure the viability of the new competitors or buyers, and operate quickly.[201]

Admittedly there is some divergence between the independence and viability criteria. The goals of the independent operation of the divested portion of the defendant's business may conflict with the objective of assuring that a successful firm will operate the divested assets.[202] The most desirable strategy toward the institution of an independent entity is the creation of a wholly new competitor in the industry via some sort of "spin-off" plan, or the sale of the divested assets to a non-competitor.[203] Yet it may be impossible to create a viable new competitor — due either to the lack of divisible assets[204] or to the unavailability of sufficient managerial talent.[205] In such instances, the optimal strategy may be to order the defendant to sell a portion of its business to existing rivals or potential entrants in related lines of business. Finally, whether construction of a new enterprise is possible or whether sale to an existing competitor is necessary, the divestiture order ought to be and has in the past been designed to occur as expeditiously as possible.[206]

---

201 *See generally* Elzinga, *The Antimerger Law: Pyrrhic Victories,* 12 J.L. & ECON. 43, 44-46, 52-53 (1969). On the independence requirement, Elzinga notes that:

> Only this sort of relief [divestiture] strikes at the very structure of the markets involved. Injunctive relief, that is, some form of order directing the acquiring firm to behave as if it did not gain this market power, is clearly unacceptable.

Id. at 45. All three requirements — independence, viability, and quickness — are deemed necessary by Elzinga for effective relief. *Id.*

202 *Id.* at 61-66.

203 *See* Baldwin, *Selective Divestiture by Spin-Off and Lottery: A Modest Proposal,* 6 ANTITRUST L. & ECON. REV. 107, 110 (Winter 1972-1973) [hereinafter cited as Baldwin (II)].

204 *See, e.g.,* United States v. National Lead Co., 332 U.S. 319, 351-53 (1947) (Court found that viable competitors could not be created out of National Lead's and DuPont's plants).

205 *See* Williamson, *supra* note 150, at 1516-18.

206 The courts have in fact developed criteria for the formulation of divestiture

Even with these additional concerns as to the independence and viability of the enterprise controlling the divested assets, and with

remedies — whether implemented through spin-off or mere sale of assets in Clayton Act § 7 cases. In formulating a spin-off type divestiture decree the courts, according to Baldwin, have generally followed the following steps: first, identify and remove from the defendant's control those assets that give it anti-competitive power in the market (note that the purpose of this step is to remedy the anticompetitive market situation rather than to punish the defendant); second, supplement those assets segregated under step one with additional assets which assure that the new competitor will be viable; third, equalize the relative bargaining strength of the two parties; and fourth, establish a mechanism that will assure reasonable promptness in actually implementing the divestiture. These guidelines have provided the courts with an adequate methodology in Clayton Act cases where the merged firms have operated as a single economic entity for a considerable period of time. *See generally* Elzinga, *supra* note 97, at 44-46.

Additionally, where a spin-off has been infeasible and sale to an already existing firm has been necessary, the courts have developed supplementary guidelines under § 7 of the Clayton Act. Those cases in which the government action was not finalized until after considerable intermingling of the assets of the merger parties prove particularly instructive to the Sherman Act situation. For the solution is no longer a simple unwinding, but rather the severing of a fully integrated enterprise. The discussion under § 7 has been framed in terms of two interrelated choices: first, should the courts require partial divestiture (the relinquishment of that portion of the merger addition which led to a reduction of competition in a particular line of commerce) or total divestiture (the separation of the existing approximation of the original merger addition); second, should the assets divested included "after-acquired" property (post-merger improvements added to the assets obtained through merger). The lesson derived from the most recent resolutions of these issues is that a sufficient portion of assets must be sold in order to insure that the purchasing firm is to be functionally competitive in that line of commerce which was subject to the divestiture order. Thus, the questions posed above are misleading to the extent they ignore the viability criterion. For earlier examples of "partial divestitures" which proved insufficient, *see* Hooker Chemical, 59 F.T.C. 254 (1961). (Hooker purchased two competitors. The court ordered the divestiture of one — which promptly failed due to insufficient backing.) *See* Elzinga, *supra* note 97, at 54-59. *See also* Gulf Oil Corp., 56 F.T.C. 260 (1956); Brillo Mfg. Co., 64 F.T.C. 245 (1964). The "box score" of the partial divestiture versus total divestiture contest is well-summarized in Brock, Mergers and Markets, Studies in Business Economics No. 105 (National Industrial Conference Board, 1969). Brock notes that from 1951 to 1958 there were 108 merger cases (20 involving total divestiture and 59 involving partial divestiture). From 1968 to 1971, there were 44 merger decrees with twelve involving total divestiture and nineteen partial divestiture.

However, the current resolution holds that if a choice between the two must be made, complete divestiture is desirable. *See* Utah Public Service Co. v. El Paso Natural Gas Co., 395 U.S. 464, 471-472 (1969). For a discussion of the merits of the two approaches, *see generally* Elzinga, *supra* note 97; Martin, *The Brown Shoe Case and the Antimerger Policy*, 53 AMER. ECON. REV. 340 (1963).

For cases specifying that after acquired property should not automatically be excluded from divestiture decrees, *see* United States v. Combustion Eng'r. Inc., 1971 Trade Cas. ¶ 73.648 (D. Conn. 1971) (consent decree requiring divestiture plus sale of all improvements made by acquiring company); Gates Rubber Co., CCH [1970-1973 Transfer Binder] TRADE REG. REP. 19.657 (F.T.C. 1971) (consent order re-the possibility of efficiency losses due to reduced scale, the balancing process will specify divestiture as the appropriate remedy in the vast majority of cases. Hence, a presumption in favor of this structural remedy is indicated. If the defendant can affirmatively prove that the real economic cost of divestiture outweighs the expected benefit of divestiture, other structurally-oriented remedies ought to be considered.

The propriety of this presumption is supported by the clear superiority of divestiture in terms of effectiveness and by the rarity of injurious efficiency results due to its utilization. Moreover, conduct remedies may involve anti-efficiency effects more frequently than do direct structural attacks. Conduct regulatory injunctions can severely limit the ability of a business to function efficiently for years to come. A good example is the case of a conduct injunction which requires a firm to grant royalty-free licenses of valuable patents.[207] Baldwin notes that:

> The most widely known example of broad and severe consequences resulting from restraints on conduct is, undoubtedly, the later impact of the meatpacking decree of 1920 [citation omitted], in particular the effects of the provisions enjoining the consenting firms from entry into the fields of groceries, fluid milk and other non-meat food products. The continuing burdensome nature of these restraints is made clear by the subsequent efforts of the meatpackers to have the decree modified or set aside and by the Government's unaltered and so far successful opposition to any change. One of the primary objectives of the decree, the Government argued in 1958, "was to bar the defendants permanently from using the great size which they had achieved in the meat industry and which they were permitted to retain as an aid in obtaining competitive advantages in any invasion of the grocery and food fields."[208]

---

quired divestiture of acquired assets plus subsequent additions and improvements). *See generally* Elzinga, *supra* note 97, at 59-61. *But see,* Reynolds Metal Co. v. FTC, 309 F.2d 223 (D.C. Cir. 1962).

207 *See* Baldwin, *supra* note 28, at 140. *See also* Brookshire and Carroll, *Patents and Vertical Integration as a Source of Monopoly Power: The Photographic Industry,* 7 ANTITRUST L. & ECON. REV. 49 (1974).

208 Baldwin, *supra* note 28, at 140. Baldwin notes that an additional problem results from such conduct injunctions in that the assets of a restrained firm may be worth much less to it than to others not bound by the conduct injunction concerning the particular property. Given this disparity in the perceived value of the property, the defendant's likely reaction is to sell the property, thus freeing it

In short, loss of substantial control over an aspect of its business operation may be more detrimental to the efficient functioning of a firm than the structural modification required by a divestiture decree.[209] Conduct regulatory decrees tend to be long-lived and of unpredictable effect,[210] because the period of adjustment to such decrees is often prolonged. This contrasts sharply with structural relief which impacts in a fairly brief time frame. In short, these perpetual constraints on business activity prevent optimal reaction to any necessary efficiency alteration and thereby aggravate the efficiency loss resulting from the conduct measure.

C. *Effects on Third Parties*

Third among the criteria for comparison of structurally oriented and conduct directed remedies, and the second major component of the harshness charge frequently leveled against divestiture, is the tendency toward penalization of innocent shareholders, employees, and creditors. Although harm to such individuals is immaterial to the remedy selection process under a strict interpretation of the law, judges frequently have hesitated to grant

---

from the conduct injunction and enabling its return to anti-competitive performance. *See, e.g.,* Baldwin's discussion of the 1948 Gamewell Co. consent decree, *id.* at 148-49. Baldwin's solution to this problem is to suggest that the conduct injunction ought to be made applicable to all subsequent purchasers of the property. *Id.* at 149. However, this adjustment is not satisfactory since it permanently institutes the restrictions which delimit the firm's range of activities and may thereby prevent the pursuit of certain efficiencies.

209 *See generally* Baldwin, *supra* note 28, at 138-49. Baldwin recognizes this possibility, and therefore argues that the enforcement agencies ought to gear their requests toward conduct relief designed to affect structure, since the courts are more prone to accept conduct relief rather than unmitigated structural relief. "Thus, the enforcement agencies may be more successful in obtaining meaningful and 'harsh' conduct relief than they would be in seeking equally [or more] effective structural relief." *Id.* at 141. Thus implicit in Baldwin's analysis is the perception that divestiture might in many situations be both more effective and less harsh to efficient operation of the firm than conduct-regulatory injunctions.

210 A complex regulatory decree entered by a court may place the latter in the position of being the chief regulator of an industry or of an important aspect of it for years to come, as in the *ASCAP* case, 341 F.2d 1005 (2d Cir. 1964) *cert. denied,* 382 U.S. 877 (1965). *See* Sam Fox Publishing Co. v. United States, 366 U.S. 683 (1961).

Under a judgment in a private antitrust case Judge Igoe of the Federal District Court in Chicago found himself passing upon the merits of particular motion pictures from time to time. Celler, *supra* note 169, at 696 n.19.

divestiture on this basis.[211] Arguably, shareholders are harmed by the decrease in the value of the firm affected by divestiture.[212] Employees may be affected by employment dislocations, or by loss of pension benefits due to the decline in the value of the business enterprise.[213] Creditors may be adversely affected by the reduction of present earnings, and by the supposedly increased riskiness of the business.[214]

However, the negative impact of divestiture on such third parties is most probably overstated, and in general structural relief may be little more severe than a conduct-oriented remedy. In divestiture, real assets are neither dismembered nor destroyed.[215] As Baldwin has noted:

> In 1912, for example, E.I. duPont de Nemours was subjected to divestiture amounting to virtual dissolution, with Atlas Powder Company and Hercules Powder Company being created out of the divested properties. United States v. E.I. duPont de Nemours and Co., 188 Fed. 127 (C.C. Del. 1911). The divestiture, as well as the very large earnings made during the First World War, undoubtedly played an important role in DuPont's decision to move into new fields and its resulting emergence as a great diversified chemical firm. Indeed, it could be argued that DuPont's loss of its monopolistic position in explosives turned out to be a stimulus of great value to the company. In any event, it would be very difficult to show that the divestiture did any serious long-run harm to DuPont's investors [or creditors or employees].[216]

In a subsequent *DuPont* case brought under § 7 of the Clayton Act, the Supreme Court ordered the district court to require DuPont to divest itself of 63 million shares of General Motors common stock worth almost 10 billion dollars.[217] The district

211 *See*, Dewey, *Romance and Realism in Antitrust Policy*, 63 J. POL. ECON. 93, 95-99 (1955).

212 *See Hearings on S. 1167 Before the Subcomm. on Antitrust and Monopoly of the Senate Comm. on the Judiciary*, 94th Cong., 1st Sess. 6152-53 (1975).

213 *Id.* at 5916-17, 6003-05.

214 *Id.* at 6193-95.

215 Baldwin, *supra* note 28, at 139. Note that to some extent the real costs of divestiture overlap with the costs to third parties in that the reduced viability of a business may harm the public as well as third parties.

216 *Id.*

217 United States v. E.I. du Pont de Nemours & Co., 366 U.S. 316, 328-31 (1961),

court had granted a complicated remedy which, in essence, allowed DuPont to retain ownership of the stock but prevented it from utilizing the voting power of the stock.[218] The district court selected this remedy on the basis of a balancing test: the expected advantages of the remedy in reducing restraints on competition set against the remedy's expected harm to shareholders. Such a test indicated greater benefit from use of the voting restriction than from implementation of divestiture.[219] The Supreme Court rejected the test and held that the harm to shareholders ought to enter the remedy selection calculus only where the remedies were equally effective.[220] If a particular remedy was most effective in its procompetitive impact, it should be chosen regardless of its potential cost to shareholders.[221] The Supreme Court found the divestiture remedy to be clearly more effective than the voting-control option.[222]

Many commentators on the antitrust laws have long agreed with this Supreme Court ruling that one must distinguish real economic costs to the public from specific costs to third parties.[223] The 1955 *Report of the Attorney General's National Committee to Study the Antitrust Laws* includes among its four general criteria[224] for assessing the value of divestiture relief the following two:

> 3. It is important to consider the effect of a possible resultant disruption upon the industry involved, its cognate markets, and the public needs in peace and war.
>
> 4. Though we recognize it is not feasible to prepare before trial a final plan for effectuating such relief, *we feel that once divestiture has been ordered*, the Division must take account,

---

rev'g, 177 F. Supp. 1 (N.D. Ill. 1959). Earlier the Supreme Court, 353 U.S. 586 (1957), had reversed the district court's dismissal of the Government's complaint, 126 F. Supp. 235 (N.D. Ill. 1954).

218 *DuPont*, 177 F. Supp. 1, 39-43 (N.D. Ill. 1959).

219 *Id.*

220 *DuPont*, 366 U.S. 316, 327 (1961).

221 *Id.* at 328.

222 *Id.* at 331-32.

223 See Comment, *Aspects of Divestiture as an Antitrust Remedy*, 32 FORDHAM L. REV. 135, 142 (1963).

224 REPORT 355-56. The first and second requirements are:

> 1. It should not be decreed as a penalty;
>
> 2. It should not be invoked where less drastic remedies will accomplish the purpose of the litigation; . . .

Id. at 355.

> in submitting a plan to effectuate the order, of its effect on the public as well as on the defendant and persons interested in it, as investors, customers, and employees. Such appraisal, we emphasize, seems a prime responsibility of any antitrust enforcement agency. [Emphasis added.][225]

The first of these statements suggests that the real economic costs[226] ought to be weighed by the Antitrust Division in determining whether to seek divestiture. The second guideline suggests that only after divestiture has been ordered ought the costs of divestiture to third parties enter into the remedy formulation procedure.

The reasons for the distinction between public and third-party costs are manifold. First, the dominant purpose of the Sherman Act monopolization prohibition is to achieve the real economic benefits that are assumed to flow from competitive markets. Clearly, these benefits ought to be balanced against the real economic costs of various formulations of both the standards for liability and remedy selection. But costs to third parties, to the extent that they consist of a mere redistribution of benefits from the third parties to the general public rather than a "dead-weight" social loss, are simply beyond the general concern of antitrust law. Nowhere in the legislative history of the Sherman Act does there appear much solicitude for the owners of the great trusts.[227]

Second, the costs of divestiture to third parties ought not enter into the remedy selection calculus because such costs generally are not considered in the selection process vis-à-vis other types of relief, especially conduct-oriented relief.[228] Where such costs are considered with regard to divestiture but not with regard to other forms of relief, remedy selection is biased against divestiture. Costs to third parties are probably ignored in the consideration of conduct injunctions because such remedies are not perceived to affect the shareholders' interest directly.[229]

A third reason for this exclusion is that the level of such costs

---

225 *Id.* at 356.

226 See text accompanying notes 190-210 *supra.*

227 See text accompanying notes 101-14 *supra* for discussion of populist overtones behind antitrust laws.

228 *See, e.g.,* United States v. United Shoe Machinery Corp., 110 F. Supp. 295, 346-351 (D. Mass. 1953). For subsequent case history see note 82 *supra.*

229 *Id.*

is to a great extent discretionary with the defendant.[230] The costs which a particular divestiture order imposes on third parties can vary in accordance with the method chosen to implement the divestiture order. Further, since the defendant controls access to most of the detailed information needed to assess the tax and financial[231] consequences of various modes of effectuating a divestiture order, it may be difficult to assess accurately the impact of such considerations because the defendant may be inclined to overstate the costs of a given procedure.

In the end, the better procedure excludes third-party costs from consideration in the remedy selection proceeding but, once the remedy is chosen, allows the defendant to introduce a plan which effectuates the order, estimates the costs to all relevant parties, and predicts the plan's effectiveness in meeting the purpose of the order.[232]

Thus, any divestiture remedy which is effective in curtailing monopoly power will concomitantly reduce the expected stream of future monopoly profits for the target firm. Such a reduction in monopoly profits will be reflected in lower stock prices. Current shareholders who purchased stock prior to the time when the company attained monopoly power will see the value of their stock settle to that level which it would have maintained had their firm

---

230 *See* W. SHEPHERD, THE TREATMENT OF MARKET POWER 183-215 (1975).

231 *Id.*

232 *See* 1 LOSS, SEC. REG. 138 (1969) for discussion of the "voluntary compliance" features of the PUHCA:

> Compliance with § 11 became, in most cases, a two-step affair. First the Commission would hold a hearing and enter an order declaring *what* action the particular holding company or subsidiary had to take under the 11(b) standards, without specifying *how* it was to be accomplished. This action might vary from a reclassification of securities to a divestment of certain holdings to outright liquidation. After the entry of an order under § 11(b) — or, so far as the statute is concerned, without waiting for such an order — any holding company or subsidiary may submit a plan under § 11(e) designed to bring it into compliance with the § 11(b) standards. The accent is thus on voluntary compliance. Section 11(d) authorizes the Commission to apply to the appropriate District Court to enforce an 11(b) order, on the basis of a reorganization plan proposed either by the Commission itself or by any interested person, if the holding company or subsidiary does not comply with the 11(b) order within a year from the date of its entry; but the very presence of the 11(d) power made it unnecessary to resort to it.

Id.

not achieved and exercised market power in violation of the Sherman Act. Consequently, they lose only the unpurchased expectation of a supernormal dividend flow and resale value; more plainly, their anticipated windfall vanishes. Such equity participants have no real complaint, as they enjoyed the fruits of monopoly for a period, and are now being returned to their *ex ante* status. However, shareholders who purchased stock subsequent to the firm's attainment of monopoly power probably did so at a price which reflected the capitalized stream of future monopoly profits — thus, the return on their initial investment was "normalized." Divestiture, by dissipating monopoly power, reduces the future profit stream to a level which will yield below normal return to the investors who purchased when the monopoly was intact. Hence, divestiture may be said to penalize them unjustly.

The harm befalling such shareholders as a result of divestiture is no different, however, from the harm which follows from conduct remedies. Any conduct-limiting decree which is effective in reducing monopoly power similarly will decrease the stream of future profits. Thus, only if real efficiencies are differentially impacted by the two classes of remedies will the consequent alteration of earnings differ; and this is an effect to be considered under general costs to the public rather than under its special result to shareholders.[233]

Further, the priority ordering of various classes of equity need not be affected by divestiture. The process by which such interests are to be kept intact lends itself to simple description if somewhat more difficult implementation — although the courts have successfully employed such methods before. These methods value the target firm as a going concern, calculate the present proportionate interests of the equity holders, and distribute the new securities so that the different classes will be allowed to participate to that same extent in the new and old firms.[234] Utilization of conduct injunc-

---

233 See text accompanying note 211 *supra*.

234 The "going concern value" of a business is the value of the firm's assets calculated by capitalizing and discounting the expected future stream of the income generated by those assets. It is to be contrasted with the forced sale value of the business' assets as though the business were in bankruptcy. Professor Louis Loss has drawn the distinction between these two methods of valuation with reference

tions, however, arguably imposes disproportionate loss on the holders of simple common stock, as other forms of equity retain

---

to the experience of the Public Utility Holding Company Act of 1935, 15 U.S.C. § 79a (1971), as follows:

> Many of the rules developed under traditional reorganization procedures were found to be equally applicable under § 11(e). For example, not infrequently there was found to be such wrongdoing on the part of a parent company as to require subordination of its claims to those of other security holders under the "Deep Rock" doctrine. But there are substantial differences between Chapter X of the Bankruptcy Act and § 11(e). These are due primarily to the fact that § 11(e) has to do normally with going businesses. In the application of the traditional standard of "absolute priority" to § 11(e), the Commission evolved the so-called "investment value" doctrine. This holds, in substance, that the measure of equitable equivalence for the purpose of § 11(e) is the value of the securities on a going concern basis and not as though a liquidation were taking place. Thus, it has been held that fairness and equity require that the common stock be permitted to participate in a reorganization under § 11(e) when it has a legitimate investment interest in the holding company and would be in a position to receive earnings in the future except for the necessity of winding the company up; in other words, § 11 does not have the effect of maturing the liquidation claims of the preferred stock. On the other hand, since the liquidation is compelled by the statute, the doctrine of "equitable equivalence" does not limit the preferred stock to the *involuntary* liquidation value stated in the charter.

Loss, 1 SECURITIES REGULATION 140-41 (1961) (emphasis added, footnotes omitted). *See also* V. BRUDNEY & M. CHIRELSTEIN, CORPORATE FINANCE 4-32, 66-72 (1972) on the concept of "going concern value."

The specified methodology has been employed previously in Sherman Act § 2 cases. *See, e.g.*, Harriman v. Northern Securities Co., 197 U.S. 244 (1905) (on dissolution of a corporation as a remedy for violating the antitrust laws, where title to the stocks of the component enterprises involved in the illegal corporation has been passed, the former owners are not entitled to the return of their specific shares, but only to a ratable proportion of the assets of the illegal corporation, *id.* at 294-97; and, where termination of unlawful situation may be effected by a sale and distribution in cash or by distribution in kind, the corporation is justified in using either method, *id.* at 299); United States v. Reading Co., 273 F. 848 (E.D. Pa. 1921), *aff'd sub nom.*, 259 U.S. 156 (1922) (equity requires that no distinction be made between holders of common and preferred stock).

Perhaps the most comparable forerunner of the pro rata approach in antitrust divestiture remedies is found in the history of the implementation of the Public Utility Holding Company Act of 1935, 15 U.S.C. § 79a (1971). The Act was passed to correct abuses in the use of the holding company in the nation's gas and electric industry, North American Co. v. SEC, 327 U.S. 686 (1946); and to protect consumers through the elimination of restraints of free and independent competition, SEC v. New England Electric Service Corp., 390 U.S. 207, 210 (1968). Professor Loss has noted that:

> Section 11 was designed to strengthen the capital structures of holding company systems and to return control over the Nation's utilities to local management and local regulation. *This portion of the act is essentially a specialized type of antitrust and corporate reorganization law.* However,

their prior right to the now reduced earnings pool. Hence, divestiture would appear to distribute more fairly any initial loss to shareholders caused by realigning the violating firm with the competitive norm.

In general, the arguments which are relevant to consideration of the impact on shareholders apply similarly in evaluating the effect on creditors and employees. Creditors who extend resources to a firm before monopoly power is realized are in no way harmed when the value of their credit returns to "normal" as the value of the firm drops from its monopoly level due to divestiture.[235] Creditors who loan money on the basis of a valuation which includes the monopoly power component clearly suffer a loss in security upon the institution of divestiture. Yet it is not unprecedented that the value of their security may actually increase as the

---

> because the electric and gas utility industries are to some extent natural monopolies, the statute is a peculiar sort of antimonopoly law in that it is designed to restore the effectiveness of state and federal regulation rather than the effectiveness of competition in a free market. The aim of § 11 is to restrict each holding company to a single integrated electric or gas utility system having a simple capital structure, with provision for the retention of additional utility systems and related incidental business under specific standards.

Loss, *supra*, at 135 (emphasis added).

With regard to shareholders' rights, it was the purpose of Congress in passing the PUHCA to preserve values of securities, not to diminish them, and Congress did not intend that maturities of securities should be accelerated to give any creditor or shareholder a special position to the detriment of any other party in interest in the corporate enterprise. *In re* SEC, 142 F.2d 411, 419 (3d Cir. 1944), *aff'd sub nom.*, Otis & Co. v. SEC, 323 U.S. 624 (1945); *In re* Laclede Gas Light Co., 57 F. Supp. 997 (D.C.E.D. Mo. 1944), *aff'd sub nom.*, Mass. Mut. Life Ins. Co. v. SEC, 151 F.2d 424 (8th Cir. Mo. 1945), *cert. denied*, 327 U.S. 795 (1945). Again, Professor Loss notes:

> The corporate simplification provision is contained in § 11(b)(2), which requires generally the elimination of undue and unnecessary corporate complexities, such as "great-grandfather" holding companies, and the redistribution of voting power on a fair and equitable basis among security holders of the entire holding company system.

Loss, *supra*, at 136-37.

See, e.g., Electric Bond & Share Co., 11 S.E.C. 1146 (1942), *aff'd sub nom.*, American Power & Light Co. v. SEC, 141 F.2d 606 (1st Cir. 1944), *aff'd*, 329 U.S. 90 (1946); Leary, *"Fair and Equitable" Distribution of Voting Power under the Public Utility Holding Company Act of 1935*, 52 MICH. L. REV. 71 (1953).

Finally, this prescribed view has received consideration under § 7 of the Clayton Act. *See generally* United States v. E.I. du Pont de Nemours & Co., 366 U.S. 316 (1961), discussed at text accompanying notes 217-22 *supra*.

235 *Cf.* sources discussed in text accompanying note 216 *supra*.

newly created competitors flourish.[236] Creditors may fear that the spun-off entities will be ultimately incapable of meeting their assigned debt obligations. Yet this eventuality will be carefully scrutinized in the courts' assessment of the viability of the new enterprises, so that where divestiture is ordered, it will be upon the courts' determination that such financial responsibilities will be carried out.[237]

Finally, the conduct-injunction alternative to a divestiture remedy also may lead to a reduction in the value of the affected property as the business activities are curtailed.[238]

The last contingent of third parties to be considered is the employees of the target firm. Theoretically, divestiture would not adversely affect employment of those previously associated with the target firm since the reduction in monopoly power will generally lead to greater output at lower prices and hence higher employment by the more competitive successor firms.[239] Of course there will be certain dislocation effects, and particular employment categories may be negatively impacted. Yet once again, the same types of readjustment ripples may follow altered firm behavior due to the institution of conduct-affecting decrees.

In sum, not only does divestiture fare rather well in comparison with conduct injunctions when considering real economic consequences, but even where the two types of remedies are deemed equally effective and it becomes appropriate to assess the impact on third parties, divestiture is frequently no more detrimental to such interests than is its conduct-oriented counterpart.

D. *Judicial Supervision*

A fourth dimension along which divestiture has often been adversely compared with conduct-regulatory remedies is the amount of judicial supervision required. Divestiture relief often necessitates the district court's supervision to insure that a suitable purchaser of the stock or assets to be divested has been designated, that the proper items have been transferred, and that the decree

236 See text accompanying note 216 *supra*.
237 See text accompanying notes 201-34 *supra*, and notes 241-49 *infra*.
238 See text accompanying note 216 *supra*.
239 *See* Baldwin, *supra* note 28, at 140.

is not somehow circumvented.[240] That such supervision can continue for several years is exemplified by the *DuPont* Clayton Act § 7 case.[241]

However, a decree designed to regulate conduct often will involve a court even more substantially in the operation of a business and over a longer period of time than even the most far-reaching divestiture decree. For example, in *DuPont* the district court refused divestiture and granted only a limited conduct injunction.[242] The Supreme Court reversed and ordered divestiture even though the challenged stock transaction had taken place some 30 years before the institution of the suit, using the following language:

> ... But the public interest should not in this case be required to depend upon the often cumbersome and time-consuming injunctive remedy. Should a violation of one of the prohibitions be thought to occur, the Government would have the burden of initiating contempt proceedings and of proving by a preponderance of the evidence that a violation had indeed been committed. Such a remedy would, judging from the history of this litigation, take years to obtain. Moreover, an injunction can hardly be detailed enough to cover in advance all the many fashions in which improper influence might manifest itself. And the policing of an injunction would probably involve the courts and the Government in regulation of private affairs more deeply than the administration of a simple order of divestiture. We think the public is entitled to the surer, cleaner remedy of divestiture. The same result would follow even if we were in doubt. For it is well settled that once the Government has successfully borne the considerable

240 See text accompanying notes 201-06 *supra*.

241 See notes 217-22 and accompanying text *supra*. The *Paramount Pictures* case, although it involved substantial divestiture, had a comparatively brief history. United States v. Paramount Pictures, Inc., 70 F. Supp. 53 (S.D.N.Y. 1946) (A. Hand, J. opinion finding various film distributors in violation of the Sherman Act §§ 1 and 2 and ordering them to make their films available on a competitive bidding basis), *aff'd*, 75 F. Supp. 1002 (S.D.N.Y. 1948) (intervenor denied standing to sue defendants for violation of conduct injunction, *id.* at 1004), *remanded in part*, 334 U.S. 131 (1948) (liability finding upheld, but the Court ordered the lower court to reconsider the relative merits of divestiture vis-à-vis the competitive bidding requirement, *id.* at 166), *aff'd*, 85 F. Supp. 881, 895-96 (S.D.N.Y. 1949) (divestiture deemed necessary and thus ordered, *id.*), *aff'd per curiam*, 339 U.S. 974 (1950).

242 *See* United States v. E.I. du Pont de Nemours, Inc., 366 U.S. 316, 333-34 (1961).

> burden of establishing of law, all doubts as to the remedy are to be resolved in its favor.[243]

The Court in *Corn Products Refining* also opted for a divestiture decree and rejected a conduct injunction with the following language:

> The difficulties of proof, the delay, the cumbersome inquiry necessary to ascertain again whether the defendants shall have actually discontinued, all make [sic] against such limitation. It may be safely assumed that evidence such as was by chance available here of the actual purpose of those in charge will never again exist.[244]

Further, in the *Paramount Pictures* case, the Supreme Court again reversed the trial court's determination that it should supervise a large-scale bidding scheme designed to restore competition in the distribution and exhibition of motion pictures. The Court held that divestiture was the preferred remedy in this case.[245]

Arguably, the long delays in many antitrust actions result in large part from extensive scrutiny of the conduct predicates of such offenses.[246] Requiring a court to supervise the application of a conduct regulatory or prohibitory decree necessitates similar complex hearings in regard to compliance for years afterward.[247] The length of court supervision of such conduct relief is compounded by the fact that often conduct remedies require fairly long periods of time to be effective — if they are ever effective — in the terms set out in the decree.[248] Thus again, divestiture compares well with conduct-oriented measures.

To conclude, in terms of effectiveness, harshness, and the required court supervision of decrees, the divestiture remedy, in general, is superior to the conduct-regulatory or prohibitory injunction. Only where the harm due to a loss of economies of

---

243 *Id.*

244 *Id.* United States v. Corn Products Refining Co., 234 F. 964, 1018 (S.D.N.Y. 1916).

245 United States v. Paramount Pictures, Inc., 334 U.S. 131, 161-66 (1948). *See* note 241 *supra.*

246 *See* J. BAIN, *supra* note 180, at 562.

247 *See, e.g.,* United States v. United Shoe Machinery, 110 F. Supp. 295 (D. Mass. 1953). United Shoe's subsequent history is discussed at note 49 *supra.*

248 See text accompanying notes 223-32 *supra.*

scale outweighs the benefit from reduced concentration and increased competition should the divestiture remedy be ruled out. A trial court selecting the appropriate remedy in a monopolization case ought to look first at divestiture rather than conduct relief. Also, given this general conclusion, the special courts proposed by the various legislative deconcentration proposals ought to be given a more specific mandate to rely on divestiture relief as a matter of course.[249] But even under the present § 2, such a presumption in favor of a structural remedy is possible. The following section examines various justifications offered by the courts for particular remedy selections, and shows how proper application of well-founded, judicially-enunciated principles requires the divestiture presumption.

V. Current Law and Its Implications for the Divestiture Remedy

Given the emergence of economic theory which specifies that market structure largely determines market performance, one might expect to observe a concomitant movement toward structural measures in remedy selection under § 2 of the Sherman Act. Yet the ordering of divestiture continues to be a rare occurrence.[250] In part, this is explained by a perceived circumscription of action resulting from remedy selection guidelines adopted by the courts during the course of nearly nine decades of Sherman Act application. In adjusting the remedy formulation process in accordance with new developments in industrial organization economics, the judiciary must contend with a body of precedent which systematically limits the application of structural relief. Four more or less independent criteria have developed, and each will be considered below.

A. *Substantial Violation*

Several courts have indicated that divestiture ought to be reserved for those cases where a "substantial" violation of the Sher-

249 See notes 50-52 and accompanying text *supra*.
250 See text accompanying notes 77-87 *supra*.

man Act § 2 standards is shown.[251] It is difficult to determine whether this doctrine means that only when the currently accepted standards for § 2 liability are clearly violated will divestiture be granted, or whether it is, in fact, an argument for adopting standards of liability which are more favorable to defendants. Either interpretation of the rule, however, has apparent flaws.

In the former case, the court ignores the shift in the focus of the proceeding which can and ought to occur once a violation is found. Such a perspective unnecessarily constrains the remedy formulation proceeding. In the latter case, the debate concerning the proper liability standards[252] is muddled unnecessarily by the introduction of the remedy question. Both interpretations ignore the salient distinction between liability-determining and remedy-formulating proceedings: the former looks to the past to determine whether a violation has occurred, whereas the latter has as its object the prevention of future anticompetitive acts.[253]

The failure of the courts to distinguish between these different

---

251 *See* United States v. Aluminum Co. of America, Inc., 91 F. Supp. 333, 340-41 (S.D.N.Y. 1950), *remedy rev'd,* 153 F. Supp. 132 (S.D.N.Y. 1957). *See also* Van Cise, *Limitations Upon Divestiture,* 19 GEO. WASH. L. REV. 147, 148-49 (1950). *Cf.* Geddes v. Anaconda Copper Mining Co., 254 U.S. 590 (1921). With regard to the Sherman Act § 2 charge, the court found that the private plaintiffs could not maintain such an action. *Id.* at 593. But with regard to the Clayton Act § 16, 38 Stat. 730, 737 (1914), *as amended (later),* 15 U.S.C. § 16 (1970) the Court said:

> Upon the case here made by the evidence it is impossible to conclude that the defendants constituted in 1911 such a combination, within the terms of the Anti-Trust Act, as would justify the granting of an injunction to the plaintiffs even under the provisions of § 16 of the Clayton Act, which we have quoted.

Id. at 595. *But see* A. NEALE, *supra* note 68, at 419-20.

252 See notes 41-42 *supra* for a discussion of one element of this debate. For example, some have deemed persistence of monopoly power over time a sufficient basis for a finding of § 2 liability, Williamson, *supra* note 150, at 1512-22, whereas others have deemed its absence a sufficient condition for denying divestiture relief. P. AREEDA, ANTITRUST ANALYSIS: CASES, TEXT AND PROBLEMS 54-55 (2d ed. 1974). In the former interpretation, the debate has turned on the natural rate of decline of market power over time and the justice of finding liability where no conduct element is involved. See note 41 *supra.* In the latter formula concerning remedies, the confusion reflects the possibility that remedy selection may turn on the narrowness of the definition of one or more of the elements of the offense.

See, e.g., A. NEALE, *supra* note 68, at 406-07; Comment, *Private Divestiture: Antitrust's Latest Problem Child,* 41 FORDHAM L. REV. 569, 579 (1973). *See also* discussion of the relation between liability standards and remedies at note 53-64 *supra.*

253 *See* United States v. Aluminum Co. of America, 91 F. Supp. 333, 346 (S.D.N.Y. 1950), *remedy reviewed,* 153 F. Supp. 132 (S.D.N.Y. 1957). *See also* cases cited at note 80, *infra.*

perspectives is readily illustrated. Mere size alone has never been held to constitute monopolization without an additional showing of abusive acquisition or use of that size.[254] The usual justifications for this controversial doctrine[255] are that it is unfair to penalize defendants merely for characteristics which manifest no detrimental effect and that any penalty based solely upon per se market share standards may lead to "anticipatory degradation."[256] However, once the violation has been established and the remedy consideration process is underway, these two reasons for the refusal to focus exclusively on structure are largely vitiated. At that point the inquiry should shift from "what happened" to "how do we prevent future abuses of the defendant's monopoly power." And if industrial organization theory instructs that the superior manner of preventing future anticompetitive performance is to alter directly the structural context, such a focus can in no way be considered unfair.[257] Additionally, "anticipatory degradation" plays no role

---

254 *See, e.g.*, United States v. Swift & Co., 286 U.S. 106, 116 (1932) (Cardozo, J.); United States v. United States Steel Corp., 251 U.S. 417, 451 (1920); United States v. Int'l Harvester Co., 274 U.S. 693, 708 (1927); United States v. Aluminum Co. of America, 91 F. Supp. 333, 340-41 (S.D.N.Y. 1950) (court notes that although mere size is not enough for liability alone, the abusive conduct element was becoming less important).

255 *See generally* M. GREEN, THE CLOSED ENTERPRISE SYSTEM (1972).

256 Anticipatory degradation is a broad title given to various actions a firm may take to avoid crossing a per se market share barrier. For example, a firm may sell off its least productive plants, reduce advertising, or diversify rather than expand in its current markets so as to avoid acquiring a certain market share. *But see* Turner, *The Scope of Antitrust and other Economic Regulatory Policies*, 82 HARV. L. REV. 1207, 1221 (1969) for a rebuttal of the likelihood of anticipatory degradation.

257 As noted in text accompanying notes 170-78 *supra*, a given degree of market power can manifest itself in a myriad of market conduct modes so that attacking a limited set of modes without correcting the underlying structural context is ultimately ineffective. *See generally* F. SCHERER, *supra* note 1, chs. 5-17 (1970). While the argument in this paper is framed in terms of an attack on the conduct portion of the two prong § 2 liability test (because this has been the traditional focus), it is equally valid to prevent recurrence of suboptimal performance and monopolization under the Sherman Act by directly confronting the structural prong. This would avoid the economic debate concerning the exact nature of the structure-conduct-performance links. See text accompanying notes 120-51 *supra*. Courts have occasionally expressed openness to this approach:

> It is not necessary that a purpose or intent to exercise such power also be found to exist in order to justify a court in granting relief to the Government. In this respect, the doctrine now governing the application of remedies imposes a different standard than would be requisite in determining illegality in the first instance. On the practical side, the mere existence of monopoly power is instinct with the threat of future violations.

in the construction of appropriate relief measures, as the firm must comply with a court order which is developed specifically to have more than a cosmetic effect.

Further, not only should the focus of the remedy proceeding differ from that of the liability proceeding, but once the government or private plaintiff has established a monopolization charge, the burden should fall on the defendant to show why the plaintiff's proposed relief ought not be granted.[258] Such a shift is appealing both in terms of fairness and evidentiary policy. Once the plaintiff has established an abusive use or acquisition of monopoly power according to a standard which arguably favors the defendant,[259] it is not inequitable to require the defendant to indicate why a particular remedy is unacceptable. Nor is it intuitively unfair to resolve doubts as to the proper remedy against an adjudged monopolizer.

Thus, misdirection of the judicially developed "substantial violation" guideline renders that formulation inappropriate for the § 2 remedy selection process. The courts must recognize the shift of inquiry required after liability has been determined. A guideline which focuses on the prevention of the future abuse of monopoly power is dictated.

B. *Proximate Cause*

The "proximate cause" approach proscribes divestiture except where the property in question is directly involved in the violation of the antitrust laws.[260] Under this view, there must be a "clear

---

> On the theoretical side, it seems proper to consider monopoly power the fruit of an unlawful monopolization and, as such, provide that it be rendered impotent. However, when dealing with an integrated monopolizer, the "fruits" rule cannot literally be applied. A court cannot blindly divest particular ill-gotten gains without viewing constructively the creation of competitive units. Therefore, the power itself, and not the specific elements thereof, must sometimes be viewed as the "fruit."

United States v. Aluminum Co. of America, 91 F. Supp. 333, 346 (S.D.N.Y. 1950).

258 *Cf.* Hartford-Empire Co. v. United States, 323 U.S. 386, 409 (1945); United States v. Bausch & Lomb Optical Co., 321 U.S. 707, 726 (1944); United States v. Aluminum Co. of America, 91 F. Supp. 333, 346 (S.D.N.Y. 1950). *But cf.* United States v. National Lead Co., 332 U.S. 319, 351-52 (1947).

259 *See generally*, W. SHEPHERD, THE TREATMENT OF MONOPOLY POWER 143-46 (1975).

260 *See* Van Cise, *supra* note 251, at 150-51 especially cases cited at 150, nn.6-9.

and proximate causal relationship between any antitrust crime and such [a] drastic remedy."[261] Once again, the formulation of the remedy hinges directly on the past occurrence which led to the liability, rather than on the prevention of future violations.[262]

Often the extent to which the courts have relied upon this doctrine to deny divestiture or to limit its scope has been unclear. This confusion has emerged largely because the rhetoric employed by the courts in discussing remedy selection often suggests several reasons for refusing divestiture.[263] For example, in *United States v. National Lead Co.*,[264] the defendants, National Lead and DuPont, were found to have combined and conspired via patent licensing pools and cartel argreements to restrain the interstate and foreign commerce in titanium, in violation of the Sherman Act. Although the lower court ordered the defendants to dispose of their holdings in the foreign companies involved in the trade restraints, it refused the Government's request that each of the defendants be further required to dispose of one of its two domestic plants.[265] The Supreme Court upheld the lower court, stating:

> There is neither allegation in the complaint nor finding of fact by the District Court that the physical properties of either National Lead or DuPont have been acquired or used in a manner violative of the Sherman Act, except as such acquisition or use may have been incidental or related to the agreements above mentioned. The cancellation of such agreements and the injunction against the performance of them by the appellant companies eliminate them.[266]

---

261 Van Cise, *supra* note 251, at 150. *See generally* United States v. National Lead Co., 332 U.S. 319, 352 (1947) (industrial plants merely incidental to antitrust violations need not be divested); United States v. Crescent Amusement Co., 323 U.S. 173, 189 (1944) (divestiture of corporations affiliated with the corporate defendants justified in order to avoid future combination among conspirators); United States v. Great Lakes Towing Co., 217 F. 656 (N.D. Ohio 1914) (continuing operation of defendant company under stringent injunctive regulations held preferable to divestiture of company properties).

262 See text accompanying notes 51-59 *supra*.

263 *See, e.g.*, Hartford-Empire Co. v. United States, 323 U.S. 386, 409-10 (1945) (divestiture denied because remedies must be proscriptive, not punitive and must not endanger the defendant's business — courts are not to take legislative role); United States v. Imperial Chemical Industries, Ltd., 105 F. Supp. 215, 244-46 (S.D.N.Y. 1952) (need for immediate divestiture not shown in antitrust violations by international munitions companies).

264 United States v. National Lead Co., 63 F. Supp. 513 (S.D.N.Y. 1945).

265 *Id.* at 534.

266 United States v. National Lead Co., 332 U.S. 319, 351 (1947). *Accord* Timken

The Supreme Court went on to say that requiring each of the defendants to sell one of its two plants, thus creating two new competitors in the titanium market, was unnecessary because the smaller competitors of National Lead and DuPont had been strengthened sufficiently by the conduct injunctions directed against the defendants.[267] The Court also noted the possible impracticality of implementing the divestiture of the defendant's plants.[268] Thus, while requiring divestiture of only those properties directly involved in the violation played some role in the Court's reasoning, the exact nature of its influence was clouded by the presence of additional considerations.[269]

Despite its uncertain impact, the proximate cause analysis has been employed frequently and has been used in support of numerous limited divestiture orders. For example, in the many railroad stock acquisition and consolidation cases, the unlawful acquisition and manipulation of the stock in question often served as the basis for divestiture[270] or divestiture-like[271] orders. Further, where a defendant has used property directly to destroy competition and has appeared likely to misuse that property in the future, divestiture of such property has been deemed appropriate.[272]

Roller Bearing Co. v. United States, 341 U.S. 593, 600-01 (1951) (divestiture relief denied in § 1 Sherman Act case).

267 *Id.* at 352.

268 *Id.* at 353. *See* A. NEALE, *supra* note 68, at 417.

269 For a full discussion of these additional factors see text accompanying notes 300-49 *infra.*

270 *See generally,* United States v. Southern Pacific Co., 259 U.S. 214 (1922) (acquisition by railroad of controlling stock of competing railroad held unlawful and divestiture ordered); United States v. Reading Co., 253 U.S. 26 (1920) (holding company dominating two major railroads and two major coal companies divested); United States v. Union Pacific R.R., 226 U.S. 470 (1913) (decree requiring transfer of stock owned by competing railroad held insufficient where transferee company dominated by some directors).

271 *See* Northern Securities Co. v. United States, 193 U.S. 197, 355-56 (1904) (while actual divestiture was not ordered here, the Court imposed such severe restraints on the use of the stock in question by the current owners that the net effect of the decree was divestiture). *See also* Comment, *Private Divestiture: Antitrust's Latest Problem Child,* 41 FORDHAM L. REV. 569, 580-81 (1973) for discussion of the Northern Securities case.

272 *See, e.g.,* cases involving the motion picture industry, United States v. Paramount Pictures, Inc., 334 U.S. 131, 152 (1948) (acquisitions which were the fruits of or utilized in the restraint of trade should be divested); Schine Chain Theatres v. United States, 334 U.S. 110, 128 (1948) (divestiture remedy necessary to deprive the defendant of the monopoly in film exhibition obtained in wrongful conduct); United States v. Crescent Amusement Co., 323 U.S. 173, 188-90 (1944) (divestiture of corporations affiliated with the corporate defendant justified in order to avoid

Another line of cases suggests that where the combination or conspiracy is itself the gravamen of the violation, the defendant may be dissolved to the extent necessary to dissipate the illegal combination,[273] although on occasion divestitive relief has gone beyond this principle.[274]

Thus, the proximate cause limitation has appeared in numerous contexts. Yet underlying most, if not all, of these cases is the premise that because of its supposed severity, divestiture should be granted only where the reprehensibility of the defendant's conduct demands it. Divestiture is viewed as a remedy designed to end the specific conspiracy, combination, acquisition, or use of properties which formed the predicate of the violation, and a means of depriving the defendants of the benefits of their wrongdoing.[275] Thus the focus or, more appropriately, misfocus is similar to that resulting from the substantial violation approach.[276] The two methods are reminiscent of a phrase often used disparagingly to describe the entirety of Anglo-American jurisprudence — "marching backward into the future."

---

future combination among conspirators); United States v. Paramount Pictures, Inc., 85 F. Supp. 881 (S.D.N.Y. 1949) (divorcement of defendant's business as exhibitors from its business as producers and distributors a necessary remedy where vertical integration was accomplished by exclusion of competition and price fixing).

273 *See, e.g.*, United States v. American Tobacco Co., 221 U.S. 106, 185 (1911) (horizontal conspiracy ordered dissolved although American's ties with its subsidiaries left intact by the district court); Standard Oil Co. v. United States, 221 U.S. 1, 77-78 (1911) (holding company dissolved although regionally/vertically-integrated participants left intact); United States v. Corn Products Refining Co., 234 F. 964, 1015-18 (S.D.N.Y. 1916), *cert. denied*, 249 U.S. 621 (1919) (illegal combination of plants into Corn Products Refining Co. could only be dealt with effectively by divestiture); United States v. Eastman Kodak Co., 226 F. 62, 81 (W.D.N.Y. 1915), *decree entered*, 230 F. 522 (W.D.N.Y. 1916), *cert. denied*, 255 U.S. 578 (1921) (defendant forced to divest itself of property obtained wrongfully from competitors); United States v. International Harvester Co., 214 F. 987, 1001 (D. Minn. 1914) (consolidation of five competitors in harvesting equipment market with aggregate market share of 80-85% ordered dissolved although not along the lines of the original companies).

274 United States v. The Pullman Co., 53 F. Supp. 908 (E.D. Pa. 1944), 50 F. Supp. 123 (E.D. Pa. 1943) (although Pullman acquired its market power in market for the manufacture of "sleepers" via the acquisition of every other sleeping car manufacturer in the United States between 1867 and 1900, the offense found to exist was the extension via long-term contracts of this market power to the market for the servicing of these cars: therefore, the decree ordered the separation of the Pullman Co. (servicing) from Pullman Standard (manufacturing) via the simplest method, *i.e.*, divestiture).

275 *See, e.g.*, United States v. Crescent Amusement Co., 323 U.S. 173, 189 (1944).

276 See text accompanying notes 51-59 *supra*.

The retributive and specific deterrence underpinnings of such reasoning are manifest, yet the punitive goals thereby implicated are rejected by the prevailing rhetoric and wisdom governing the application of equitable remedies. The broad consensus of opinion dictates that the Sherman Act § 2 ought not be treated as a criminal statute and that those criminal sanctions available to the Government should not be utilized[277]—it is difficult to apply punitive measures where moral indignation is lacking.[278] In short, the argument for increased utilization of divestiture rejects its implementation as a penalty against wrongdoers,[279] and views the employment of this remedy exclusively as a means of restoring competition[280] even though any decree requiring divestiture doubtless will be viewed by those controlling the target firm as punitive in nature.[281]

The remedial orientation of divestiture was apparent from the first occasions on which it was used. In both the *Standard Oil*[282] and *American Tobacco*[283] cases of 1911, the Supreme Court's rationale for granting divestiture was founded on its belief that only a remedy which dissolved the accumulated market power of the defendants could guarantee future competition in the relevant markets. While in both of these cases the divested property could be said to be proximately related to the Sherman Act violation,[284] the language chosen by the Court indicated that divestiture was being used prospectively as a restorative device, rather than retrospectively as a punitive one.[285]

---

277 *See, e.g.*, P. AREEDA, *supra* note 52, at 52-53. *See also* Turner, *supra* note 56, at 1221-22.

278 A. NEALE, *supra* note 68, at 402.

279 ATTORNEY GENERAL'S NAT'L COMM. TO STUDY THE ANTITRUST LAWS 355 (1955).

280 *See, e.g.*, Timken Roller Bearing Co. v. United States, 341 U.S. 593, 603 (1951); United States v. Paramount Pictures, Inc., 334 U.S. 131 (1948); United States v. Griffith, 334 U.S. 100, 109 (1948); Schine Chain Theatre v. United States, 334 U.S. 110 (1948); International Salt Co. v. United States, 332 U.S. 392, 401 (1947).

281 A. NEALE, *supra* note 68, at 401.

282 Standard Oil Co. v. United States, 221 U.S. 1, 78 (1911) (Sherman Act violation based on the formation of a national stock holding company which was ordered dissolved).

283 American Tobacco Co. v. United States, 221 U.S. 106, 187 (1911) (Sherman Act violation based on control exercised by American over accessory and subsidiary companies through stock ownership which was ordered divested).

284 See note 73 *supra*.

285 See text accompanying notes 82-83 *supra*. *See also* Comment, *Private Divestiture: Antitrust's Latest Problem Child*, 41 FORDHAM L. REV. 569, 580-82 (1973).

This emphasis, although frequently ignored in later Sherman Act § 2 cases,[286] continues to specify the appropriate focus in the remedy formulation proceeding on a prospectively oriented search for a solution to the problem of overwhelming market power and its performance manifestations.[287] The object of the remedy is to eliminate the undesirable consequences of market power and to assure the public freedom from any recurrence[288] — in short, to determine how to open a market to competition.[289] In order to pursue this goal effectively, the courts must ponder alternative scenarios with regard to the industry in question.[290] The district court must determine whether competitors can be expected to flourish and whether there are foreseeable changes in market conditions that would obviate the need for judicial intervention.[291] In partic-

286 See text accompanying notes 53-57, *supra*.

287 *See* A. NEALE, *supra* note 68, at 410.

288 *Cf.* Ford Motor Co. v. United States, 405 U.S. 562, 573 (1972) (divestiture ordered in Clayton Act § 7 case).

289 *See* International Salt Co., Inc. v. United States, 332 U.S. 392, 401 (1947); United States v. Griffith, 334 U.S. 100, 109 (1948).

290 *See, e.g.*, United States v. Aluminum Co. of America (Alcoa), 91 F. Supp. 333, 416-17 (S.D.N.Y. 1950) (the district court here painted a wide range of scenarios before rejecting divestiture on "practicality" and "necessity" grounds, see text accompanying notes 300-49 *infra*). *See also* A. NEALE, *supra* note 68, at 412 (discussion of *Alcoa* case).

291 *E.g.*, United States v. Aluminum Co. of America, 91 F. Supp. 333, 346-47 (S.D.N.Y. 1950).

It is noteworthy that the trial court may take judicial notice of market structure information where the appropriate remedy is concerned, but not with regard to liability. *See* United States v. Aluminum Co. of America, 148 F.2d 416, 445-46 (2d Cir. 1945), *remedy reviewed*, 91 F. Supp. 333 (S.D.N.Y. 1950), *remedy reviewed*, 153 F. Supp. 132 (S.D.N.Y. 1957) (petition to extend district court's jurisdiction over *Alcoa* decree declined since Alcoa's market share had declined and competition was found to be thriving).

In the *Alcoa* case, the district court denied dissolution relief on the basis of a wide-ranging prospective analysis of the market structure of the aluminum industry. The court argued that dissolution would tend to weaken the American aluminum industry in world markets. Further, Alcoa only had two plants which were competitive with Alcoa's chief competitors, Reynolds and Kaiser. The problem of finding sufficient, qualified managerial talent to staff newly formed competitors, and the difficulty of preserving vigorous research departments, were among the factors which led the court to decide against dissolution. The court did force Alcoa to divest its Canadian subsidiary, however. United States v. Aluminum Co. of America, 91 F. Supp. 333, 416-18 (S.D.N.Y. 1950). But while general domestic dissolution was not granted, the court did not limit its analysis of the remedy question to the specific conduct Alcoa had engaged in or the specific property involved in the violation. The court considered the overall market situation prospectively. *Id. See also*, United States v. Pullman Co., 50 F. Supp. 123, 127 (E.D. Pa. 1943), *aff'd per curiam*, 330 ular, the court must look to see if in fact the market power related to the conduct in question can be controlled or dissolved.[292]

To limit the extent of divestiture to the specific property involved in the violation is to reduce divestiture to an "unmerger" decree of use only in Clayton Act § 7 cases.[293] In Sherman Act § 2 cases where market power has not been acquired through a recent marger, the district court would be unable to provide adequate structural relief.[294] Given the finding that conduct-directed remedies are generally ineffective in correcting market performance,[295] such a limitation of divestiture would totally emasculate § 2.

On occasion, courts have recognized the limitations inherent in the proximate cause to the formulation of monopolization remedies. Thus, while paying lip service to the proximate cause doctrine by alluding to the need for establishing a causal link between the target property and the violation, several courts have recognized that a prospective assessment of the probable effects of the various remedies on market power is desirable. In *Crescent Amusement*, for example, the divestiture order was justified in part by the proximate cause reasoning that the defendant ought not be allowed to reap the benefits of its illegal combinations and acquisitions.[296] Additionally, however, the court explicitly recognized that, without complete divestiture and severing of ties, the propensity for restrictive conduct on the part of the defendants would be likely to continue.[297] Yet where the courts have recognized the necessity for an examination of present market condi-

U.S. 806 (1947); United States v. Paramount Pictures, 334 U.S. 131 (1948) (divestiture ordered here means that structural elements are given new emphasis). *See generally*, C. KAYSEN & D. TURNER, *supra* note 23, at 21.

292 *See, e.g.*, Standard Oil Co. v. United States, 221 U.S. 1, 77-78 (1911).

293 15 U.S.C. § 18 (1970). The term "unmerger" as used in this article, refers to the severance of the acquired firm from the acquiring firm and thus restoration of the status quo prior to the merger.

294 For example, in United States v. National Lead Co., 332 U.S. 319 (1947), the overwhelming combined market power of National Lead and DuPont made possible and profitable the restrictive patent licensing and cartel agreements. Yet the court, for a variety of reasons, including the proximate cause doctrine, rejected divestiture of both of the defendant's dual-plant operations. *Id.*, at 351-55.

295 See text accompanying notes 170-89 *supra*.

296 United States v. Crescent Amusement Co., 323 U.S. 173, 189 (1944). *See also*, Schine Chain Theatres v. United States, 334 U.S. 110, 128 (1948).

297 323 U.S. at 189-90. *See also*, United States v. Imperial Chemical Industries, Ltd., 105 F. Supp. 215, 236-38 (S.D.N.Y. 1952).

tions with an eye toward remedial modification, all too often the ultimate divestiture decree has been limited to the specific property involved in the violation.[298] To escape this misdirected focus on past events, the judiciary must turn to guidelines other than the proximate cause formula which fails to prevent future market failure, a failure which it shares with the substantial violations approach.

C. *Necessity*

Because of the several alternatives to the divestiture remedy from which the district court may choose in attempting to achieve the purposes of the Sherman Act,[299] the courts generally have held that divestiture ought to be granted only where it is *necessary* — where no other type of relief would be adequate to remedy the situation.[300] However, the rhetoric employed by courts rarely de-

---

298 *See, e.g.*, Schine Chain Theatres v. United States, 334 U.S. 110, 128 (1948) (the Court makes some suggestion that divestiture may go beyond the proximate cause principle but ultimately does not so hold, *id.* at 129-30).

299 See text accompanying notes 65-76 *supra*. Effects similar to those achieved by divestiture can be pursued through many different technical forms. For example, where divestiture of capital stock is sought, the court can: (1) enjoin specific acts of the company; (2) enjoin the voting of the stock for certain purposes; (3) sterilize the stock completely; (4) place the stock in the hands of a trustee; or (5) order partial divestiture. *See generally* Wise, *Three D's of Antitrust Enforcement* in HOFFMAN'S ANTITRUST LAW AND TECHNIQUES 407 (M. Hoffman & A. Winard ed. 1963). The alternatives to divestiture include: (1) fines imposed by the Government, *e.g.*, 15 U.S.C. §§ 1-2 (1970); (2) treble damages, 15 U.S.C. §§ 1-5 (1970); and (3) imprisonment, *e.g.*, 15 U.S.C. §§ 1-2 (1970). Other seldom-invoked penalties are contained in Sherman Act § 6, 15 U.S.C. § 31 (1973).

300 *See, e.g.*, Van Cise, *supra* note 51, at 151-53, (". . . divestiture should only be ordered to the extent that it is absolutely necessary." *Id.* at 151); Turner, *supra* note 56, at 1224. A majority of the members of the Attorney General's Nat'l Comm. to Study the Antitrust Laws 355 (1955) treated divestiture as an exceptional remedy where judicial timidity was fully justified. *But see* the vigorous dissent from this point of view. *Id.* at 356-57. Based on his survey of the 24 cases where divestiture was ordered in the Attorney General's 1955 Report, *supra* note 79, at 354 n.13, Wise found two requirements governing the application of divestiture:

> First, divestiture relief must be necessary and the circumstances such that no other relief will be adequate. Second, divestiture relief must be practicable and the circumstances such that divestiture is feasible. The application of these two principles is dependent upon the circumstances in each individual case.

Wise, *Three D's of Antitrust Enforcement*, *supra* note 299, at 409. Note that Wise's generalization of the necessity standard does not say for what divestiture must be necessary before it can be applied. *See also* Timken Roller Bearing Co. v. United States, 341 U.S. 593, 603-04 (1951); United States v. General Electric Co., 115 F.

fines their perception of the proper function of remedies under Sherman Act § 2. Consequently, it is often difficult to determine for what purpose divestiture must be necessary before it can be granted.

While typically there are no clear-cut distinctions in the case law, the purposes which the courts have suggested that remedies must serve under the Sherman Act § 2 can be classified into three groups. First, there is the goal of preventing the repetition of the specific conduct which constituted the predicate of the violation.[301] Under this rubric, the monopolization remedy is viewed as a means whereby the industry or firm under scrutiny can be returned to the status quo as it existed before the illegal conduct occurred; and, the effect is to leave the monopoly power of the defendant and the ability to exercise that power intact except in the rare instance where the court perceives that the specific conduct in question is preventable only through divestiture.[302] Often, the basis for this restrictive interpretation of the purpose of the remedy in a monopolization case is the supposed harshness of the divestiture decree,[303] the possible existence of economies,[304] or the perceived presence of substantial technical problems in implementing divestiture.[305]

A second and more ambiguous purpose which sometimes influences remedy selection in monopolization cases is reflected in the

---

Supp. 835, 871, 875 (D.C.N.J. 1953); United States v. Union Shoe Machinery Co., 110 F. Supp. 295, 347-48 (D. Mass. 1953), *aff'd per curiam*, 347 U.S. 521 (1954).

301 *See, e.g.*, Timken Roller Bearing Co. v. United States, 341 U.S. 593, 603-04 (1951); United States v. Imperial Chemical Industries, 105 F. Supp. 215, 236-38 (S.D.N.Y. 1952) (only those subsidiaries that were involved in the illegal geographic market allocation scheme were ordered divested and only because there were no other means for remedying the illegal action).

302 *See generally* United States v. Imperial Chemical Co., 105 F. Supp. 215, 236-38 (S.D.N.Y. 1952).

303 *See, e.g.*, United States v. United Steel Co., 251 U.S. 417, 453 (1920); United States v. Hartford-Empire Co., 323 U.S. 386 (1945); United States v. United Shoe Machinery Co., 110 F. Supp. 295 (D. Mass. 1953). See generally text accompanying notes 190-249 and notes 88-100 *supra* for a discussion of the basis of the relative harshness of structural and conduct remedies, and of the timidity of the federal judiciary toward divestiture, respectively.

304 *See generally* United States v. Aluminum Company of America, 148 F.2d 416, 446 (2d Cir. 1945).

305 *See generally* Timken Roller Bearing Co. v. United States, 341 U.S. 593, 603-04 (1951).

notion that the remedy ought to provide a state of "workable"[306] or " substantial"[307] competition; or, that it should "pry open" the market in question, promoting an increase in competition.[308] Such rhetoric often has been present in cases where divestiture is made contingent upon the effectiveness of a conduct injunction.[309] In these cases, divestiture is ordered only after an instituted conduct remedy subsequently has been determined insufficient in establishing the desired market condition.[310]

An example of this is the *United Shoe Machinery* case.[311] In the original district court decision, Judge Wyzanski denied the Government's request for dismemberment of the defendant's only shoe machinery manufacturing plant since it was deemed "imprac-

---

306 *See* United States v. United Shoe Machinery Corp., 266 F. Supp. 328, 330 (D. Mass. 1967) (ten year review of Judge Wyzanski's 1953 decree, 110 F. Supp. 295).

307 *See* United States v. Eastman Kodak Co., 1954 Trade Cas. ¶ 67,920 at 70,009 (W.D.N.Y. 1954); United States v. Pitney-Bowes, Inc., 1959 Trade Cas. ¶ 69,235, at 74,864 (S.D. Cal. 1959).

308 International Salt Co., Inc. v. United States, 332 U.S. 392, 401 (1947). *See also* United States v. Griffith, 334 U.S. 100, 109 (1948); United States v. General Electric Co., 115 F. Supp. 835, 871 (D.N.J. 1953).

309 *See, e.g.,* cases cited in notes 306-07, *supra.*

310 *Id.*

311 United States v. United Shoe Machinery Corp., 110 F. Supp. 295 (D. Mass. 1953), *aff'd per curiam,* 347 U.S. 521 (1954) (initial judgment and decree), *decree reviewed,* 266 F. Supp. 328 (D. Mass. 1967) (decree held to have been effective), *rev'd in part,* 391 U.S. 244, 251-52 (1968).

> We find nothing in the 1953 decree, as amended, or in the District Court's opinion relating thereto which presents an obstacle or embarrassment to the application of this principle in the present case. If the decree has not, after 10 years, achieved its "principal objects," namely, "to extirpate practices that have caused or may hereafter cause monopolization, and to restore workable competition in the market"—the time has come to prescribe other, and if necessary more definitive, means to achieve the result. A decade is enough. Even if we should assume that paragraph 18, as the District Court now states, had only the limited purpose of calling for a 10-year report as to whether the decree was "gradually eroding United's 1953 power to monopolize the market," 266 F. Supp., at 330, its specific provisions did not exhaust the District Court's power. Relief in a Sherman Act case "should put an end to the combination and deprive the defendants of any of the benefits of the illegal conduct, and break up or render impotent the monopoly power found to be in violation of the Act." *United States* v. *Grinnell Corp.,* 384 U.S. 563, 577 (1966). The District Court should proceed to determine whether the relief in this case has met the standards which this Court has prescribed. If it has not, the District Court should modify the decree so as to achieve the required result with all appropriate expedition (citations omitted).

Id. at 251-52 (Fortas, J.).

ticable."[312] However, United Shoe was ordered to divest its nail, tack, and eyelet subsidiaries and branches within nine months, since there were no similar practicality problems and since the monopoly power United Shoe enjoyed in these markets depended almost entirely upon the defendant's monopoly in the shoe machinery industry.[313] To counter United Shoe's monopoly power in the shoe machinery business, Judge Wyzanski ordered the defendant to refrain from particular marketing practices: specifically, the defendant could no longer offer its machinery on a lease-only basis, nor could it utilize certain tying arrangements.[314] Only if these conduct remedies failed to promote workable competition in the shoe machinery market would divestiture be ordered.[315]

The decline in United Shoe's market share over the next ten years (from 85 percent in 1953 to 62 percent in 1963[316]) persuaded Judge Wyzanski that workable competition had been restored.[317] The Supreme Court did not wholly agree.[318] The Court held that the Government may have been entitled to more substantial relief than the conduct injunctions because the purpose of the decree — achieving workable competition — had not been achieved necessarily by a reduction in United Shoe's market share to 62 percent.[319]

The *United Shoe* experience exemplifies the problems of vagueness encountered under a "workable competition" standard. The term has no structural or performance-oriented content.[320] As a result, the courts, the defendant, and the defendant's competitors do not know what will be required in order to establish acceptable market conditions.

A criterion for granting divestiture which is based upon strictly quantifiable market structure standards, such as the one employed in the *IBM* case[321] or those proposed under the various legislative

312 110 F. Supp. 295, 348 (D. Mass. 1953).
313 *Id.* at 351.
314 *Id.* at 351-54.
315 266 F. Supp. 328, 330 (D. Mass. 1967).
316 *Id.* at 331-32.
317 *Id.*
318 391 U.S. 244, 251-52 (1968). See quotation in note 130 *supra*.
319 *Id.*
320 *See* Baldwin, *supra* note 28, at 141-42.
321 United States v. International Business Machines Corp., 1956 Trade Cas.

deconcentration proposals,[322] suffers from a different shortcoming. The defendant faced with such a contingent order of divestiture may find ways of avoiding the specific standard which are less than socially optimal. This manner of reaction is termed "anticipatory degradation."[323] Anticipatory degradation is an umbrella term which includes a variety of actions a current or prospective defendant could take in order to avoid penalties that would otherwise be impaired by a per se market share standard. A defendant might, for example, raise prices to the point where new entry would be encouraged and thus, in effect, trade market share for profits.[324] Or a defendant could discontinue a complete line of his product or sell its least modern plants.[325] Further, a defendant corporation might elect to avoid price wars, delay the introduction of new products, or avoid discounting.[326] All of these responses have fairly high social costs.[327]

A third and final interpretation of the necessity requirement — one suggested by the *United Shoe*[328] court — is that the remedy decreed ought to decrease the "monopoly power" either in its capacity as an element of the offense or in its capacity as an explanation for the conduct predicate. (Monopoly power can be defined

---

¶ 68,245 (S.D.N.Y. 1956). This consent decree contained a contingent structural relief provision which provided that if IBM's share of the card tabulating market did not fall below 50 percent in seven years as a result of certain conduct injunctions, divestiture was to be ordered: in fact, this is what occurred. United States v. International Business Machines Corp., 1963 Trade Cas. ¶ 8824.46 (S.D.N.Y. 1963) (IBM required to divest itself of 16 presses for tabulating cards).

322 *See, e.g.*, Hart Bill § 101(b)(3).

323 *See generally* R. Pitofsky, Lecture to Advanced Antitrust class at Harvard Law School, December 18, 1975. *See also* W. SHEPHERD, THE TREATMENT OF MARKET POWER 127-29 Figure 4.4 (1975) for an analytical treatment of the phenomena of business firms giving up some market share in exchange for higher profits and, perhaps, antitrust freedom.

324 R. Pitofsky, *supra* note 323.

325 *Id.*

326 *Id.*

327 The problem of anticipatory degradation is most acute where a defendant faces per se liability standards with divestiture as a certain remedy, or where a defendant faces a divestiture remedy if a prior long-term conduct injunction "fails" and such failure is defined quantitatively. However, the problem of anticipatory degradation is less serious under the traditional Sherman Act § 2 application, as neither the initial test for liability nor the second-line choice between remedy types employs any clear-cut structural standards.

328 See note 130 *supra*.

as ". . . the power to control prices or exclude competition").[329] Baldwin is expansive in his praise of such a remedial standard:

> The definition focuses on possession of power rather than on the set of structural conditions which provide the power. Hence it avoids both the insuperable difficulty of specifying the great number of combinations of structural factors which might lead to an unacceptable degree of economic power and the inappropriateness of singling out one measurable aspect of structure as a meaningful standard. On the other hand, since the definition is addressed to existence of monopoly power rather than its use, a decree provision framed in such terms would be sensitive to structural as well as behavioral considerations.
>
> . . . The present proposal is not for incorporation of a proscription into the body of substantive antitrust law, but rather for a remedial standard to be applied to a firm which already has been found to be in violation of Section 2 of the Sherman Act (or in some instances, perhaps, of Section 1) in a civil case and under existing criteria of guilt.[330]

Baldwin is proposing this as the contingency standard to be employed in the determination of the effect of and the necessity for structural relief in the event of its failure.[331] However, it is arguable that such considerations occasionally have influenced the courts in the initial granting of divestiture in § 2 cases.[332] The

---

329 United States v. E.I. du Pont de Nemours & Co., 351 U.S. 377, 391 (1956).

330 Baldwin, *supra* note 28, at 145-46.

331 *Id.* at 146.

332 *See generally* United States v. E.I. du Pont de Nemours & Co., 353 U.S. 586 (1957), 366 U.S. 316 (1961); United States v. Paramount Pictures, Inc., 334 U.S. 131 (1948); United States v. Crescent Amusement Co., 323 U.S. 173, 189-90 (1944); United States v. Pullman Co., 50 F. Supp. 123, 127 (E.D. Pa. 1943), *aff'd per curiam* 330 U.S. 806 (1947); United States v. International Harvester Co., 214 F. 987, 1001 (D. Minn. 1926), *aff'd* 274 U.S. 693 (1927). *But see* United States v. Hartford-Empire Co., 323 U.S. 386 (1945) (Court was unwilling to order divestiture where market power was due to the aggregation of patents because in the Court's view the technical strength of the industry depended on the defendant). In the *Standard Oil* case the Court described the purpose of Section 2 relief as follows:

> 1st. To forbid the doing in the future of acts like those which we have found to have been done in the past which would be violative of the statute.
>
> 2d. The exertion of such measure of relief as will effectually dissolve the combination found to exist in violation of the statute, and *thus neutralize the extension and continually operating force which the possession of the power unlawfully obtained has brought and will continue to bring about.*

Standard Oil Co. v. United States, 221 U.S. 1, 78 (1911) (emphasis added). *See also*

Corn Products Refining case, in particular, draws the distinction between this formulation of the purpose of the remedy in cases and the two mentioned earlier.[333] The Court argued that if only the abusive *exercise* of monopoly power was the problem, then the question arises as to what is necessary to prevent repetition of such conduct.[334] Sometimes a conduct-prohibiting injunction would suffice.[335] However, if monopoly power itself was condemned, divestiture would be the logical remedy.[336] Ultimately, the Court in *Corn Products* ordered the defendant to divest itself of several of its plants, holding that an injunction was inadequate to meet the requirements of the statute.[337]

Thus, under the third interpretation of the "necessity" guideline, divestiture is the remedy generally relied upon. This is in stark contrast to the initial formulation of this rubric — necessary only to prevent the specific conduct leading to current liability — where divestiture is looked to only infrequently. The second interpretation with its resultant contingent divestiture remedy is directly related to current belief concerning the various structure-conduct links discussed heretofore.[338] The courts, in employing the conduct remedy directly and the divestiture remedy only contingently, have revealed that the present perceptions of the judiciary are not necessarily synonomous with widely accepted economic learning.[339]

Yet the "necessity" guideline can provide a framework under which the divestiture presumption is warranted. Of course, the view of formulae as necessary to cancel the specific conduct which led to the § 2 violation is inappropriate under the reasoning used

---

United States v. American Tobacco Co., 221 U.S. 106, 187 (1911). *See generally* W. SHEPHERD, THE TREATMENT OF MONOPOLY POWER 64 (1975) (one purpose of the antitrust laws is to abate existing market power); Sherman & Tollinson, *Public Policy Toward Oligopoly: Dissolution and Scale Economics*, 4 ANTITRUST L. & ECON. REV. 77, 81-82 (Summer 1971); Mason, *Preface to* C. KAYSEN & D. TURNER, ANTITRUST POLICY at xvi (1959).

333 United States v. Corn Products Refining Co., 234 F. 964 (S.D.N.Y. 1916), *appeal dismissed,* 249 U.S. 621 (1919), *decree modified* Oct. 18, 1921.

334 *Id.* at 1015.

335 *Id.*

336 *Id.*

337 *Id.* at 1015-17.

338 See text accompanying notes 120-51 *supra.*

339 *Id.*

to criticize the substantial violation and proximate cause guidelines — it dwells upon the past rather than upon control of future events. But under either the second or third set of necessity goals, increased use of divestiture is indicated. For the second grouping, the initial, rather than contingent, use of divestiture is dictated, as it has become apparent that conduct remedies are inappropriate in the vast majority of cases. Finally, for direct dissipation of market power, divestiture continues to be the best solution.

In sum, the "necessity" requirement may be a misnomer, as the current economic learning suggests that generally direct structural relief is the *necessary* remedy, but that on occasion other forms of relief also might be *necessary*. The choice must be made not in terms of *necessity*, but in terms of relative effectiveness — that is, excess of expected benefits over real economic costs. Thus, a more accurate characterization of this criterion would emphasize the desirability of divestiture as presumptively the most effective relief in light of the current economic learning.[340]

D. *Practicality*

The fourth and final criterion which the federal courts have utilized to evaluate the divestiture remedy in monopolization cases is perhaps the best conceived. The requirement, simply stated, is that divestiture ought to be granted only where it is "practicable"[341] or clearly in the "public interest."[342] However, in implementing this guideline the courts have granted divestiture only

---

340 See text accompanying notes 170-89 *supra.*

341 United States v. United Shoe Machinery Corp., 110 F. Supp. 295 (D. Mass. 1953) (see text accompanying note 130 *supra* for discussion of this case); United States v. Nat'l Lead Co., 332 U.S. 319, 351-53 (1947) (see text accompanying notes 64-69 *supra* for discussion of this case). United States v. United States Steel Corp., 251 U.S. 417, 453 (1920) (divestiture denied since no feasible way of divesting combination while protecting its foreign trade was suggested to the court); United States v. Terminal R.R. Ass'n of St. Louis, 224 U.S. 383 (1912); *decree modified,* 236 U.S. 194 (1915); United States v. General Electric Co., 115 F. Supp. 835, 864 (D.N.J. 1953) (court refused to order divestiture of General Electric's incandescent-lamp business because it was neither feasible nor necessary. Court also refused to grant divestiture of the defendant's partial interests in certain foreign companies, which were parties to certain agreements found to be illegal: conduct injunctive relief deemed sufficient); United States v. American Can Co., 234 F. 1019 (D. Md. 1916), *appeal dismissed,* 256 U.S. 706 (1921).

342 *See* Van Cise, *supra* note 251, at 152-53.

where the expected benefits to be gained from divestiture to the general public and to shareholders, creditors, and other parties were directly associated with the enterprise found in violation of the Sherman Act § 2 proscriptions.[343]

There are two basic problems with this formulation of the practicality requirement. First, the judiciary typically has underestimated the benefits to be derived from divestiture remedies.[344] Second, the courts generally have overstated both the costs to the public at large[345] and the costs to the individuals associated with

---

343 *See* cases cited in note 341 *supra*. *See also* Justice Frankfurter's vigorous dissent in the *DuPont* case, United States v. E.I. du Pont de Nemours & Co., 366 U.S. 316, 335 (1961).

Professor William Shepherd describes the current remedy selection procedure as follows:

> An effective remedy was likely to be ordered if (1) there was a technical basis for dividing, such as decentralization and an origin in recent mergers, (2) the product was relatively simple and standardized, so that innovation or national military involvement were not major questions, (3) only a moderate weakening in the monopolists' total position was involved — so that potential competitors were active and expectations were not sharply reversed by divestiture (but not so severe a weakening as to make action superfluous) — and (4) *the stock was closely held, so that the expected or actual impact was not widespread.*

W. SHEPHERD, THE TREATMENT OF MARKET POWER 190 (1975) (emphasis added). Another commentator has noted that:

> Ideally, account should not be taken of these considerations until divestiture has been determined to be the necessary remedy. In actual practice, however, the interests of the corporate defendant, its stockholders and employees are often weighed with the public's stake in a freely competitive economy to determine initially the feasibility of a divestiture order. To counter the government's demands for a decree of divestiture, counsel for the defendant corporation will usually demonstrate to the court the hardships which a divestiture order would work, and will suggest one or more alternative remedies which it is felt will accomplish the same end, but the hardships to private interests. If the argument for the alternative remedy fails and divestiture is ordered, the "hardships" argument is again employed, this time to attempt to mold a divestiture plan most favorable to the private interests involved.

Comment, *Aspects of Divestiture as an Antitrust Remedy*, 32 FORDHAM L. REV. 135, 142 (1963).

344 See text accompanying notes 170-89 *supra* for a discussion of the strong causal connection running from market structure elements and market performance elements.

345 The real economic costs to the public at large due to divestiture include the possible loss of economies of scale as well as the possible negative impact on incentives of the firm in question. The general nonpervasiveness of economies of scale have been discussed elsewhere in this article, see text accompanying notes 190-210 *supra*. *See also* Note, *The Industrial Reorganization Act: An Antitrust Proposal to Restructure the American Economy*, 73 COLUM. L. REV. 635, 654 (1973) (discussion

the enterprise directly, often relying upon mere speculation rather than realistic expectation.[346]

A third, and perhaps the most fundamental, criticism of this remedy selection applied by the federal judiciary is that it balances the wrong things. On the cost side of the ledger *only* real economic costs to the public should be included. Only as between equally effective remedies should the costs to third parties be considered in choosing a remedy.[347] These contentions are developed in prior

---

of the Hart Bill). Some have even argued that the existence of scale economies are the only effective argument against divestiture of a defendant found in violation of the Sherman Act's monopolization proscription. Sherman & Tollinson, *Public Policy toward Oligopoly: Dissolution and Scale Economies,* 4 ANTITRUST L. & ECON. REV. 77, 89-90 (Summer 1971). *See also,* W. SHEPHERD, THE TREATMENT OF MARKET POWER 184 (1975) (The Sherman Act § 2 has no provisos about scale economies, good behavior, harm to innocent investors, or intent. However, application is unpredictable and infrequent.)

Also to be noted is the possible adverse effect of divestiture on business acumen. Williamson argues that we ought not let the mere presence of business acumen deter a finding of liability under the Sherman Act § 2. *See* Williamson, *supra* note 150, at 1516-18. Rather, in a remedy proceeding, if it is apparent that the defendant's position (i.e., dominant firm position) was obtained through business acumen, the court ought to ask whether managerial talent is divisible. If not, then market dominance is probably the price we must pay for excellence. *Id.* If managerial talent is divisible, there is no difficulty in this regard in breaking up the firm. *Id.* Further, "business acumen" really may be a code word for relative efficiency. The market may have suffered from what Williamson terms "default failure" by the competitors of the defendant, thus creating merely the impression of a high degree of business acumen. *Id.* A possible example of default failure is the diesel locomotive manufacturing industry, dominated by General Motors. *Id.* at 1525. *See also Hearings Pursuant to S. Res. 61 Before the Subcomm. on Antitrust and Monopoly of the Senate Comm. on the Judiciary,* 84th Cong., 1st Sess., pt. 8, at 3946-98 (1955) (testimony of C.R. Osborne). In short Williamson argues that if dissolution will provide efficacious relief to single-firm dominance, a monopolization suit is warranted. Divestiture should be granted unless: (1) the dominant firm possesses *absolute* managerial superiority that will be dissipated if divestiture is attempted; or (2) the disincentives to incipient dominant firms would be substantial if divestiture were carried. *Id.* at 1526.

A. Neale has correctly noted that where there is only simple financial control and integration over separate operating units divestiture is an easy remedy. He cites as examples the Standard Oil, American Tobacco, DuPont (explosives), and railroad cases as examples of this type case. A. NEALE, *supra* note 68, at 412-13 (1970). Further, Neale argues that the practicality standard as interpreted by the courts is faulty because it may lead to discrimination against those industries which have certain types of connections easily undone. *Id.* at 414. *See* Fraidin, *Dissolution and Reconstruction: A Structural Remedy, and Alternatives,* 33 GEO. WASH. L. REV. 899, 906-11 (1965) for a discussion of the "unmerger" cases as he calls them.

346 *See* cases cited at note 103 *supra.*

347 *See, e.g.,* Fraidin, *supra* note 345, at 903-05 who notes that:

1. Expected diseconomies should not necessarily bar structural relief. It

sections.[348] Proper application of the practicality guideline requires the balancing test previously described.[349] Hence, the presumption toward divestiture is appropriate under this guideline, as well as under the necessity approach, as a structural remedy will be generally prescribed by the balancing test.

Conclusion

The primary deficiency of current antitrust policy is not its restrictive and cumbersome liability standard, but the lack of effective remedies for such violations.[350] This article has attempted to provide the basis for revised guidelines for the granting of Sherman Act § 2 monopolization remedies.

First, divestiture ought to be the presumed remedy for monopolization offenses. This guideline is based on the findings that

---

> is legitimate, indeed necessary, that a court appraise the diseconomies involved in the dissolution of a monopolizing corporation. But such inquiry should not stop on discovering that some diseconomies will result from decreasing the concentration in the industry. That is question begging. The difficult task facing the court is to balance the predicted positive and negative results of a proposed decree. Absent a contrary public policy, such as military necessity or conservation, the only "indissoluble" corporation is the declining-cost company where insuperable economies of scale exist. Theories of dissolution that emphasize the physical barriers to dissolution, such as a one-plant monopoly, may err in overrating short-run diseconomies (footnotes omitted).

Fraidin approaches a formulation of the proper balancing test from the perspective of one arguing against a test which would hold that the mere likelihood of real economic costs to the public ought to be determinative against a divestiture remedy. Rather, he argues for a test which would balance these costs against the expected real economic benefits — but not the costs to third parties.

Another commentator has noted that "Ideally, account should not be taken of these considerations [i.e., harm to third parties] until divestiture has been determined to be the necessary remedy." Comment, *Aspects of Divestiture as an Antitrust Remedy*, 32 FORDHAM L. REV. 135, 142 (1963). See note 343 *supra*, for quotation from same source.

Cf. United States v. E.I. du Pont de Nemours & Co., 366 U.S. 316, 327-29 (1960). The Court posited two requirements: (1) the district court must choose a remedy that effectively redresses the violation; and (2) the adverse tax and market consequences of total divestiture cannot save a partial divestiture if partial divestiture is not as effective. *Id.* Note that this was a Clayton Act § 7 case brought in 1948 challenging the acquisition of stock in General Motors by the defendant in 1917-1919.

348 See text accompanying notes 211-39 *supra*.

349 See text accompanying notes 170-89 *supra*.

350 See text accompanying notes 77-87 *supra*.

divestiture tends to be more effective[351] and less harsh than conduct-related relief,[352] and, additionally, supported by legal precedent.[353] The industrial organization literature demonstrates the substantial causal links between the elements of market structure and performance.[354] Conversely, the links between structure and conduct and between conduct and performance are less clear-cut.[355] Consequently, remedies directed exclusively at conduct will improve performance only ambiguously and uncertainly given the likelihood that market power will be effectuated through forms of market conduct not covered by the conduct injunction.

This presumption in favor of the divestiture remedy for monopolization violations ought to be rebuttable only where the defendant can show that substantial real economic costs will outweigh the benefits accruing from divestiture due to increased competition.[356] The possible harm to those parties with a direct financial interest in the scrutinized business(es) ought not be included in this calculation both under a reasonable interpretation of the law and on policy grounds.

Further, structural remedies in general, and divestiture in particular, may actually be less severe than conduct regulatory and prohibitory injunctions in terms of both their real economic costs[357] and the impact on those with a direct financial interest in the defendant corporation.[358] Divestitures also may involve less expenditure of judicial resources than would a conduct regulatory injunction in force over several years.[359] In short, divestiture may be less harsh than conduct injunctions which are even less effective.

Finally, there is precedent under a sensible interpretation of existing guidelines for this proposed Sherman Act § 2 remedy selection process. First, the requirements that divestiture ought to be ordered only where the violation of the liability standards are

---

351 See text accompanying notes 170-89 *supra*.
352 See text accompanying notes 190-249 *supra*.
353 See text accompanying notes 51-59 *supra*.
354 See text accompanying notes 170-89 *supra*.
355 See text accompanying note 170 *supra*.
356 See text accompanying notes 211-39 *supra*.
357 See text accompanying notes 190-210 *supra*.
358 See text accompanying notes 211-39 *supra*.
359 See text accompanying notes 241-49 *supra*.

substantial[360] or only with regard to the property directly related to the conduct element of the violation[361] ought to be disregarded. Not only do these requirements ignore the need for prospectively oriented remedy proceedings, but they tend to obfuscate the trial court's view of the current industrial organization learning.

The guidelines that divestiture be both necessary — or rather, relatively effective — and practical, if sensibly interpreted, ought to be retained. To a great extent, the necessity requirement's precedential value for more vigorous use of divestiture depends upon the court's interpretation of the purpose a monopolization remedy should serve. If the first, narrower interpretation is adopted, divestiture will almost always be deemed "unneccessary" given the simpleness or simple-mindedness required to frame a conduct injunction around the conduct found to be the predicate for liability.[362] However, even here a strong argument for divestiture could be made on the basis that it might in fact be more effective in preventing the conduct in question than a conduct remedy.[363] The second standard moves closer to the mark of a usable structural standard for deciding on an appropriate decree but it suffers from the problem of vagueness.[364] On the other hand, a precise, per se market share test may lead to anticipatory degradation which, while not always undesirable, may have serious negative social consequences.[365] Lastly, the "necessity standard" purpose most supportable on the basis of the Sherman Act legislative history as well as the theoretical and empirical findings of industrial organization economists, is the goal of dissipating monopoly power.[366] The remedy in Sherman Act § 2 cases ought to be designed to vitiate the monopoly power that formed an element of the monopolization violation. While in some cases divestiture may not be necessary for this purpose, it would seem to be the logical remedy in most cases brought under § 2.

---

360 See text accompanying notes 251-59 *supra.*
361 See text accompanying notes 260-98 *supra.*
362 See text accompanying notes 299-301 *supra.*
363 See text accompanying note 302 *supra.*
364 See text accompanying notes 306-26 *supra.*
365 See text accompanying notes 56, 323-27 *supra.*
366 See text accompanying notes 328-40 *supra.*

Finally, the practicality requirement, if demarcated with care, provides support for presumptive use of divestiture. For in considering alternative remedies the court ought to weigh the expected benefits of a given remedy against the remedy's expected real economic costs.[367] The trial court should look to prior experiences with the divestiture remedy to spark its imagination for creative routes of asset dispositions in order to enhance the benefits of divestiture.[368] Conversely, the court ought not confuse the real economic costs of a given remedy with the possible costs to those with a direct financial interest in the defendant corporation. Only as between equally effective remedies should these third-party costs enter the cost-benefit calculus. Finally, the court must not disregard the evidence which suggests that divestiture may be less harsh in terms of both real economic costs to consumers and costs to third parties than conduct remedies.[369]

---

367 See text accompanying notes 341-49 *supra.*

368 See text accompanying notes 211-39 *supra.*

369 See text accompanying notes 190-249 *supra.*

STATUTE

TO RIGHT MASS WRONGS: A FEDERAL CONSUMER CLASS ACTION ACT

JAMES ANDREW HINDS, JR.*

Modern society seems increasingly to expose men to . . . group injuries for which individually they are in a poor position to seek legal redress, either because they do not know enough or because such redress is disproportionately expensive. If each is left to assert his rights alone if and when he can, there will at best be a random and fragmentary enforcement, if there is any at all. This result is not only unfortunate in the particular case, but it will operate seriously to impair the deterrent effect of the sanctions which underlie much contemporary law. The problem of fashioning an effective and inclusive group remedy is thus a major one.[1]

The concerns of early commentators about the lack of effective legal remedies for small claimants remain as valid today for consumers as they were when written over thirty years ago. Indeed, recent decisions have all but ended the usefulness of small claimant class actions in the federal courts. Mr. Hinds, after an examination of the current state of consumer remedies and proposals, sets forth a comprehensive federal class action statute. Under the Hinds proposal, some of the more notable innovations in both substantive and procedural areas serve as a basis to broaden the concept of consumer fraud, to lower jurisdictional and evidentiary barriers, and to liberalize the requirement of notice to the greatest extent consistent with due process. The statute emphasizes the active role of the trial court, the need for adequate incentives for private enforcement, and the realities of class action litigation.

---

*Member of the Class of 1976 at Harvard Law School. The author wishes to thank Robert J. Hoelscher, of the Class of 1977 at Harvard Law School, for his substantial assistance in the preparation of this work.

1 Kalven & Rosenfield, *The Contemporary Function of the Class Suit*, 8 U. CHI. L. REV. 684, 686 (1941).

Introduction

In the past fifteen years consumer protection has become a matter of increasing public concern.[2] Numerous studies have concluded that unscrupulous business tactics contribute to social unrest in poverty areas[3] and cause severe financial distress to consumers.[4] These tactics concern the business community as well, because the unethical merchant harms the competitive position of the honest businessman. The difficulties involved in attacking such practices, however, are many. The number and variety of techniques employed to defraud consumers make it difficult to provide an analytically comprehensive definition of "consumer fraud." Legislators have traditionally resorted to nonexclusive listings of "types" of abuses, focusing primarily on misrepresentations of quality, price, warranty, and use.[5] Moreover, although

---

2 *See generally* D. CAPLOVITZ, THE POOR PAY MORE (2d ed. 1967); Eovaldi & Gestrind, *Justice for Consumers: The Mechanisms of Redress,* 66 NW. U.L. REV. 281 (1971); Hester, *Deceptive Sales Practices and Form Contracts — Does the Consumer Have a Private Remedy?,* 1968 DUKE L.J. 831; Johnson, *Consumer Rights and the Regulatory Crisis,* 20 CATH. U. L. REV. 424 (1971); Saxbe, *The Role of the Government in Consumer Protection: The Consumer Frauds and Crimes Section of the Office of the Ohio Attorney General,* 29 OHIO ST. L.J. 897 (1968); Tydings, *The Private Bar — Untapped Reservoir of Consumer Power,* 45 NOTRE DAME LAW. 478 (1970); Wade & Kamenshine, *Restitution for Defrauded Consumers: Making the Remedy Effective Through Suit by Governmental Agency,* 37 GEO. WASH. L. REV. 1031 (1969); White, *Representing the Low-Income Consumer in Repossession, Resales and Deficiency Judgment Cases,* 64 NW. U.L. REV. 808 (1970); Comment, *Class Actions for Consumer Protection,* 7 HARV. CIV. RIGHTS-CIV. LIB. L. REV. 601 (1972); Note, *Consumer Legislation and the Poor,* 76 YALE L.J. 745 (1967).

3 H. EDELHERTZ, THE NATURE, IMPACT AND PROSECUTION OF WHITE COLLAR CRIME 11 (1970); NATIONAL ADVISORY COMMISSION ON CIVIL DISORDERS, REPORT 139 (1968).

4 The Chief of the Antitrust Division in the Department of Justice has estimated that price-fixing and other anticompetitive actions by businesses raise price levels by as much as $80 billion each year. 121 CONG. REC. 7618 (daily ed. May 7, 1975). It has been estimated that several billion dollars more are taken from the public each year through consumer fraud — more than all the auto thefts, burglaries, robberies, embezzlements, larcenies, and forgeries combined. W. MAGNUSSEN & J. CARPER, THE DARK SIDE OF THE MARKETPLACE 8 (1968); PRESIDENT'S COMMISSION ON LAW ENFORCEMENT AND ADMINISTRATION OF JUSTICE, THE CHALLENGE OF CRIME IN A FREE SOCIETY 33 (1967).

5 *See, e.g.,* California Consumers' Legal Remedies Act, CAL. CIV. CODE § 1770 et seq. (West 1973); MASS. GEN. LAWS ANN. ch. 93A, § 9 (1972). Most of the state statutes employing the "types" approach are similar in substance to NATIONAL CONSUMER LAW CENTER, BOSTON COLLEGE LAW SCHOOL, NATIONAL CONSUMER ACT § 3.201 (First Final Draft 1970).

many federal laws regulate consumer transactions,[6] the responsibility for protecting the consumer has traditionally rested with the states and their common law doctrines.

Presently pending in Congress are companion House[7] and Senate[8] bills to establish a federal agency to represent the interests of consumers. Although each bill has been passed by the house of its origin, no conference committee has yet been formed to resolve the differences.[9] The primary purpose of this work is to focus on the major deficiency in the measures as they are presently written — the lack of a private right of action through the class action device — and discuss how best to provide such a remedy by amendment to the present bills.

I. The Inadequacy of Traditional Consumer Remedies

A consumer remedy can be judged by its effectiveness in accomplishing two principal objectives. First, it must compensate those

6 *See, e.g.*, the antitrust laws, 15 U.S.C. §§ 1-31 (1970); Truth-in-Lending Act, 15 U.S.C. § 1601 et seq. (1970); Federal Trade Commission Act § 5(a)(1), 15 U.S.C. § 45(a)(1)((1970).

7 H.R. 7575, 94th Cong., 1st Sess. (1975).

8 S. 200, 94th Cong., 1st Sess. (1975).

9 34 Cong. Quarterly 575 (1975).

The earliest attempt to generate federal consumer protection legislation was in 1961, when the late Senator C. Estes Kefauver introduced a bill to establish a Department on consumers. S. 1688, 84th Cong., 1st Sess. (1961). The more recent push for a consumer agency began in 1965 when Representative Benjamin S. Rosenthal introduced a bill that would have created a Cabinet-level Department of Consumer Affairs. H.R. 7179, 89th Cong., 1st Sess. (1965). But no concrete action, other than extensive hearings, resulted. *See, e.g.*, H.R. Rep. No. 141, 85th Cong., 1st Sess. (1963); H.R. Rep. No. 921, 88th Cong., 2d Sess. (1964); H.R. Rep. No. 1851, 90th Cong., 2d Sess. (1968). In 1969 and 1970, the Senate Subcommittee on Executive Reorganization considered a number of proposals for consumer representation, including separate bills introduced by Senators Joseph Montoya, Jacob Javits, and Charles Percy. S. 3165, S. 3240, S. 3097, 91st Cong., 1st Sess. (1969).

In late 1970, the Senate Government Operations Committe reported a bill that represented an amalgam of the many measures introduced in prior sessions. It established a council of consumer advisors to assist the President; an independent consumer protection agency with advocacy and product-testing functions; and a funding program for state and local consumer activities. S. 4459, 91st Cong., 2d Sess. (1970). A similar bill in the House died in the Rules Committee. H.R. 18214, 91st Cong., 2d Sess. (1970). In the 92d and 93rd Congresses, comprehensive consumer agency packages were killed by Senate filibusters. H.R. 10835, 92d Cong., 1st Sess. (1971); S. 3970, 92d Cong., 2d Sess. (1972); S. 707, 93d Cong., 1st Sess. (1973).

who have been injured by restoring to them the sum they have lost. Second, it must deter potential wrongdoers by removing the profit from the proscribed activity.[10]

Forums such as Better Business Bureaus, although serving a valuable educational function, clearly fail to provide a full remedy because they lack the power to compel a business to restore money wrongfully taken.[11] Small claims courts are scarcely more effective. Some jurisdictions are still without tribunals of this type, and despite efforts in many states to simplify small claims actions, many such courts confuse the unwary consumer with complicated procedures.[12] Moreover, small claims courts can become overrun by victims of large-scale frauds.[13] Although dishonest merchants are in theory subject to criminal sanctions, in practice they have little to fear. Prosecutors are constantly pressured to devote their attention to more violent types of crime.[14] Federal agencies empowered to deal with various aspects of consumer protection are too widely dispersed even to communicate among themselves, let alone create effective redress mechanisms.[15] There is the further problem of

---

10 *Cf.* C. McCormick, Handbook on the Law of Damages § 137, at 560 (1935).

11 Wade & Kamenshine, *Restitution for Defrauded Consumers: Making the Remedy Effective Through Suit by Government Agency*, 37 Geo. Wash. L. Rev. 1031, 1049-50 (1969).

12 *Hearings on S. 2246, S. 3092, S. 3201, Before the Subcomm. on Consumers of the Senate Comm. on Commerce*, 91st Cong., 1st Sess., ser. 48, pt. 1, at 175-76 (1969) (statement of Bess Myerson Grant).

13 *Id.* at 176, 184.

14 *Cf.* Note, *Translating Sympathy for Deceived Consumers into Effective Programs for Protection*, 114 U. Pa. L. Rev. 395, 426-27 (1966).

15 For example, within the Department of Justice alone there are a number of units dealing with consumer affairs:

> The Antitrust Division is responsible for maintaining competitive conditions in the free enterprise system. The Consumer Affairs Section of Antitrust, in particular, is responsible for the institution of civil and criminal proceedings in cases referred to the Department by other agencies, such as the Federal Trade Commission and the Food and Drug Administration, which have primary responsibility for consumer protection activities. The Criminal Division's role is to oversee the enforcement of the majority of the federal criminal statutes, many of which directly affect the consumer. The United States Attorneys' offices are responsible for the prosecution of consumer-related crimes and regulatory violations. A number of the larger United States Attorneys' offices have divisions specifically designated as consumer protection units. The Drug Enforcement Administration . . . services consumers of legitimately manufactured drugs through its regulatory and compliance program.

40 Fed. Reg. 55173 (1975). Recently a plan of reorganization was proposed by the

limited resources. It seems safe to say that no federal administrative body has ever had enough staff and money to execute fully the functions for which it was created. Finally, experience suggests that even in the best organized and most heavily funded agencies, the passage of time brings deterioration and less enthusiastic enforcement of the initial policies which gave them birth. John Kenneth Galbraith notes succinctly:[16]

> [R]egulatory bodies, like the people who comprise them, have a marked life cycle. In youth they are vigorous, aggressive, evangelistic, and even intolerant. Later they mellow, and in old age — after a matter of ten or fifteen years — they become, with some exceptions, either an arm of the industry they are regulating or senile.

It has long been recognized by commentators, scholars, and other authorities that the multitude of federal agencies whose policies and decisions affect consumers have not protected consumer interests as Congress expected when the agencies were created.[17] The report accompanying the proposed Senate bill identifies the primary reason for this failing as "the present lack of effective representation of consumer interests before Federal agencies."[18]

Department of Justice to organize these varied activities to "ensure that the needs and views of consumers are duly considered in the Department's decision-making process." *Id.*

16 J. GALBRAITH, THE GREAT CRASH, 1929, at 171 (1955); *see generally* W. GELLHORN, INDIVIDUAL FREEDOM AND GOVERNMENTAL RESTRAINTS (1956).

17 *See, e.g.*, S. REP. No. 883, 93d Cong., 2d Sess. 1-2, 6-7 (1974); H.R. REP. No. 962, 93d Cong., 2d Sess. 6 (1974); Moss v. CAB, 430 F.2d 891, 893 (D.C. Cir. 1970); Office of Communications of the United Church of Christ v. FCC, 359 F.2d 994, 1003-04 (D.C. Cir. 1966). *See generally* Stewart, *The Reformation of American Administrative Law*, 88 HARV. L. REV. 1669, 1713-16 (1975); OFFICE OF CONSUMER AFFAIRS, U.S. DEPT. OF HEALTH, EDUCATION, AND WELFARE, FEASIBILITY STUDY FOR HANDLING CONSUMER COMPLAINTS (1975); Lazarus & Onek, *The Regulators and the People*, 57 VA. L. REV. 1069 (1971); Bonfield, *Representation for the Poor in Federal Rulemaking*, 67 MICH. L. REV. 511 (1969); Cramton, *The Where, Why and How of Broadened Public Participation in the Administrative Process*, 60 GEO. L.J. 525 (1972); Jaffe, *Ecological Goals and the Ways and Means of Achieving Them*, 75 W. VA. L. REV. 1 (1972).

It should be noted that the pending legislation's definitions of "consumer" and "interests of consumers" encompass a broad range of interests including problems associated with products, services, transactions, and other areas and information relating to them. H.R. REP. No. 425, 94th Cong., 1st Sess. 9 (1975).

18 S. REP. No. 66, 94th Cong., 1st Sess. 9 (1974). *See also* H.R. REP. No. 425, 94th Cong., 1st Sess. 2 (1975).

A. *The Imbalance in Voice*

As the Senate report states, consumers are usually poorly organized, under-funded, and ill-equipped to present an effective case before federal agencies.[19] Frequently, consumers do not learn until too late about proposed decisions which affect their interests. It is true that in recent years there has been increasing awareness on the part of consumers of how governmental decision-making affects their everyday lives. Organizations like Common Cause and the various Nader groups now exist, and effective proponents have begun to come forth on behalf of consumers. But the resources of the business community enable it to influence policy far more than can inadequately financed consumer groups. Effective representation by business interests in matters affecting consumers is virtually assured because the same decision that may make only a two or three cent difference to a consumer may determine whether a business or an entire industry is able to increase or maintain its profits. In such cases business interests are ready to undergo great expense to assure that their views are heard. Yet consumers, even if they concentrate their resources and succeed in making their views known in one case, must ignore or inadequately address numerous contemporaneous matters affecting them. And since many decisions must rest on detailed and highly technical data, it is difficult for agencies to reach a conclusion in the best interests of the entire public when they hear only one or two views of a multi-faceted problem.

B. *Representation of Consumer Interests at the Federal Level*

To correct this imbalance, the legislation currently before Congress establishes a consumer protection agency (CPA) to represent the interests of consumers in the proceedings and activities of other federal agencies and to provide the decisionmakers with alternative points of view. Neither the House nor the Senate bill creates an

---

19 Representation at the state level is no better. *See generally* Leflar & Rogol, *Consumer Participation in the Regulation of Public Utilities: A Model Act*, 13 HARV. J. LEGIS. 235, 241-52 (1976); Reed, *Legislating For the Consumer: An Insider's Analysis of the Consumers' Legal Remedies Act*, 2 PACIFIC L.J. 1, 3-4 (1971).

all-powerful regulatory body or strips authority from any existing agency. Instead, the functions of the CPA are those of an advocate, complaint collector, and information clearing house.[20]

The Administrator is allowed to intervene before federal agency proceedings whenever he determines that the proceedings may substantially affect an interest of consumers. Limitations on this right of intervention are similar to those applicable to all other "parties" to such proceedings.[21] In agency hearings[22] the Administrator acts as a party representing the interests of consumers, while in "notice and comment" rulemaking[23] the Administrator is allowed to participate in any manner appropriate to the nature of the proceeding. In addition, both bills authorize the Administrator to appear in federal courts to secure judicial review, intervene as of right, and otherwise participate in civil proceedings involving the review or enforcement of any federal agency action which may substantially affect the interests of consumers.[24]

---

20 Section 5 in the House bill, note 7 *supra,* sets out the functions of the agency as follows:

> (1) to represent the interests of consumers before Federal agencies;
> (2) to encourage and support research, studies, and testing;
> (3) to submit annual recommendations to the Congress and President on measures to improve the Federal Government protection and promotion of consumer interests;
> (4) to publish and distribute materials to inform consumers of matters of interest to them;
> (5) to conduct conferences and investigations, including economic surveys, concerning the needs, interests, and problems of consumers;
> (6) to cooperate with other levels of government and with private enterprise in the protection of consumer interests; and
> (7) to keep committees of Congress informed and respond to queries by Members.

Section 5 of the Senate bill, note 8 *supra,* provides additional functions for the agency, including publishing a consumer register of actions of Congress, federal agencies, and federal courts of interest to consumers, making publicly available consumer complaints, and encouraging the development of informal dispute settlement procedures involving consumers.

21 S. 200 § 6(a) (1975); H.R. 7575 § 6(a) (1975).

22 Administrative Procedure Act §§ 4(b), 5-8, 5 U.S.C. §§ 553(c), 554-57 (1964).

23 Administrative Procedure Act § 4, 5 U.S.C. § 553 (1964).

24 S. 200 § 6(c)(1); H.R. 7575 § 6(c). The power granted the Administrator by the Senate bill exceeds that normally held by parties to agency actions, for under this version the CPA would have power to initiate judicial review of decisions in situations where it has not otherwise participated at the agency level and where such decisions substantially affect an interest of consumers. The rationale for this expanded right is that the efficiency of administration

> would be undercut if the Administrator were to be given standing only

The Administrator is also expected to serve as the focal point for consumer complaints. This function will promote a greater awareness by the Administrator of the actual problems of the consumer and will provide a method for monitoring the responsiveness of other federal agencies to consumer problems. Since the CPA will also forward the complaints received to the businesses involved, it should also be in a position to promote the informal resolution of problems. But the Administrator will have no authority to impose any solution on any person or governmental agency.[25]

The grant of powers to the CPA serves also to define the agency's limitations. The very nature of the administrative process dictates that the agency's effectiveness will be largely limited to prospective policy choices. Even in administrative and judicial proceedings in which past practices are adjudicated, there will generally be no place for the agency in the monetary redress of consumer grievances. Moreover, there are the prospects of inadequate resources and eventual agency senility.

All these considerations point to the utility — indeed, the necessity — of a private right of action as an additional weapon for consumer protection. But the individual lawsuit prosecuted by a private attorney is essentially useless as a remedy. Because most consumer frauds involve only small amounts, even a successful plaintiff could expect that his total expenditures in bringing his claim to court would exceed any possible recovery. Few people, no matter how high their income, can be expected to bring a lawsuit under such circumstances. Moreover, consumers, particularly from low-income groups, are unlikely to seek an attorney on their own, both because they lack knowledge of their rights and because they are hesitant about becoming involved in the legal system.[26]

> where he had participated at the agency level . . . [and he would naturally] be encouraged to intervene or participate in every action in order to preserve his right to participate at the judicial level.

S. REP. No. 66, 94th Cong., 1st Sess. 24 (1975). The Senate, however, limits the Administrator's right to use this initiation of review right. *See id.* at 25.

25 S. REP. No. 66, 94th Cong., 1st Sess. 11, 30 (1975).

26 *See* Rice, *Remedies, Enforcement Procedures and the Duality of Consumer Transaction Problems,* 48 B.U.L. REV. 559, 567-68 (1968). Legal aid organizations offer potential solutions to part of this private litigation problem, but all too often the defrauded consumer is too poor to afford the cost of individual litigation while having an income too high to qualify for legal aid. Furthermore, legal aid

As a result, injured consumers are effectively deprived of recompense while defrauding merchants continue their deceptive practices and clear a wrongful profit.

Substantive law problems also exist. Except in a few progressive states, state common law and statutes fall far short of granting an adequate remedy to consumers. Actions at common law for deceit or fraud generally require the plaintiff to show several separate elements — misrepresentation of material facts, reliance, and causal connection — each of which is often difficult to prove.[27] The Uniform Deceptive Trade Practices Act was designed especially to combat consumer fraud,[28] but only a few states have adopted this progressive legislation. Even where legislation exists, many courts have been loath to allow consumer class actions.[29]

---

groups often suffer from low budgets and inadequate staffing and cannot handle the crush of numerous claimants victimized by large-scale frauds.

27 W. PROSSER, HANDBOOK OF THE LAW OF TORTS, § 105, at 685-86 (4th ed. 1971). *But cf.* CAL. CIV. CODE § 1780 (West 1973).

28 *See* Dole, *Consumer Class Actions Under the Uniform Deceptive Trade Act,* 1968 DUKE L.J. 1101.

29 The opposing lines of state class action practice are illustrated by developments in New York and California. New York courts have read their class action rule quite narrowly, while California has been in the forefront of liberal interpretation.

Decisions in New York have flowed from a doctrine that no class action can be maintained unless all members have a "common interest" that would require compulsory joinder. Society Milion Athena, Inc. v. National Bank, 281 N.Y. 282, 292, 22 N.E.2d 374, 377 (1939). The strictness of this approach has meant that although injunctive relief may be had by a class, money damages, which are thought to be peculiarly individual, are usually not available. In Kovarsky v. Brooklyn Union Gas Co., 279 N.Y. 304, 18 N.E.2d 287 (1938), the plaintiff sought to maintain a class action for small, but illegal, charges for resumption of service after temporary discontinuance. Although suit was permitted for injunctive and declaratory relief, an accounting for the overcharges was dismissed on grounds that appeared close to those of the "common interest" doctrine. 279 N.Y. at 314, 18 N.E.2d at 290. *See* Starrs, *The Consumer Class Action,* 49 B.U.L. REV. 407, 435-36 (1969).

In a recent decision, the New York Court of Appeals followed this line of reasoning and dismissed a consumer class action for fixed penalties under the New York Retail Installment Sales Act. Hall v. Coburn Corp., 26 N.Y.2d 396, 259 N.E.2d 740, 311 N.Y.S.2d 281 (1970). According to the court, "[s]eparate wrongs to separate persons, though committed by similar means and even pursuant to a single plan, do not alone create a common or general interest in those who are wronged." 26 N.Y.2d at 400, 259 N.E.2d at 721, 311 N.Y.S.2d at 283, *quoting* Society Milion Athena, Inc. v. National Bank, *supra* (Lehman, J.). This rationale, which essentially removes the possibility of recovering money damages for members of a consumer class action, has been the subject of severe criticism. *See* Dole, *Consumer Class Actions Under Recent Consumer Credit Legislation,* 44 N.Y.U.L. REV. 80, 105 (1969); Comment, *Class Actions for Consumer Protection,* 7 HARV. CIV. RIGHTS-CIV. LIB. L. REV. 601, 625 (1972).

In contrast, a consumer class action can serve both goals of an adequate remedy. First, it can provide compensation for each individual at very little individual expense because the lawsuit is conducted by the attorney of the party representative and the costs of litigation are distributed among the class members as a whole. Second, and more important, the class remedy can serve as both a specific and a general deterrent by demonstrating to unscrupulous merchants the degree of success possible to class plaintiffs, the high damage awards available, the heavy litigation expenses required to defend, and the extensive unfavorable publicity involved. In a large and complex economy, this is an essential addition to the limited deterrent effect of an agency's powers.

One should also examine the class action in light of the policy considerations underlying the need for private enforcement of consumer remedies. If the maintenance of individual actions exposes litigants to disproportionate expenditures of money, time, and effort, then a denial of class relief is, in effect, tantamount to a denial of justice. Moreover, consumer class actions produce a beneficial psychological effect by channeling the efforts of numerous consumers against a common foe within the existing judicial structure. Without class actions, the maxim *lex semper dabit remedium*[30] becomes a hollow mockery and the legal system is perceived as an abettor of injustice.

California, however, has maintained flexible standards for consumer class actions under a statutory procedure similar to that of New York. CAL. CIV. PRO. CODE § 382 (West 1973). Consumer class actions are now maintainable under the California Consumers' Legal Remedies Act. CAL. CIV. CODE § 1760 et seq. (West 1973). As far back as 1948, the California Supreme Court rejected the New York compulsory joinder rule. *Compare* Weaver v. Pasadena Tournament of Roses Ass'n, 32 Cal. 2d 833, 198 P.2d 514 (1948), *with* Society Milion Athena, Inc. v. National Bank, *supra*. In Chance v. Superior Court, 58 Cal. 2d 275, 373 P.2d 849, 23 Cal. Rptr. 761 (1962), it repudiated the "common fund" requirement that had inhibited previous attempts to maintain class actions, holding that only an ascertainable class and a well-defined community of interest in questions of law and fact were needed. Daar v. Yellow Cab Co., 67 Cal. 2d 695, 433 P.2d 732, 63 Cal. Rptr. 724 (1967) required that class members be specifically identified only when absent members came forward to collect damages and indicated that "community of interest" means there need only be substantial issues which are better resolved in common. This liberal approach led to the holding in Vasquez v. Superior Court, 4 Cal. 3d 800, 484 P.2d 964, 94 Cal. Rptr 796 (1971), that a class action for fraud could be maintained by consumers where there was a common scheme to defraud. See text accompanying notes 93-96 *infra*.

30 The law will always give a remedy. H. BROOM, LEGAL MAXIMS 118 (10th ed. 1939).

One argument advanced against consumer class action legislation is that individual damages are so small that no one is really hurt. But society cannot allow a wrongdoer to profit by his misdeeds without encouraging others to engage in similar conduct to increase their profits — and penalize those that remain honest. If, on the other hand, the possibility of profit diminishes because a practical legal means exists to recover the improper gains, business methods will change. A combination of this economic effect with general ethical principles dictates that ill-gotten profits should not be retained — even if the monetary wrong per individual is minute, and even if the victim cannot be located.

Possibly the greatest benefit of the consumer class action will inure to society as a whole. Once businesses realize that they can be sued for seemingly insignificant misrepresentations and overcharges, they will be more likely to deal fairly with consumers. Judge Weinstein found a similar therapeutic value in securities fraud class actions:[31]

> Those who criticize the class actions on the grounds that it stirs up plaintiffs and serves only to provide fees for attorneys overlook the fact that we are not dealing with the traditional lawsuit which concerns primarily those litigants before the court. The public's concern with openness and honesty in public securities markets gives it an interest no less significant than that of particular plaintiffs and defendants.

II. The Federal Forum: An Inadequate Remedy

Federal court suits could help fulfill the goals of compensation and deterrence, but numerous jurisdictional and procedural barriers impede small claimants. For an action to be heard in a federal court, the statutory prerequisites to jurisdiction must first be satisfied.[32] Jurisdiction over the subject matter of the suit is generally based on either a federal question[33] or diversity of citi-

---

31 Dolgow v. Anderson, 43 F.R.D. 472, 487 (E.D.N.Y. 1968).

32 C. Wright, Handbook of the Law of Federal Courts § 7, at 15 (2d ed. 1970) [hereinafter cited as Wright].

33 28 U.S.C. § 1331 (1970).

zenship.[34] Any class action must then satisfy the requirements of Rule 23 of the Federal Rules of Civil Procedure.

At least in theory, a consumer class action might be brought under any of the three categories in Rule 23(b). It is conceivable that the prosecution of individual actions against a business could create the risk that "varying adjudications . . . would establish incompatible standards of conduct for the party opposing the class" as required by Rule 23(b)(1). But the small individual amounts typically involved make the risk of multiple actions remote.

If the primary relief requested is an injunction to halt a particular practice, or if declaratory relief is sought, Rule 23(b)(2) would be applicable.[35] A determination under this subsection could provide a number of flexible and powerful equitable remedies to the plaintiff class, including the recall of defective goods or an across-the-board rescission of contracts deemed oppressive. Yet however great the deterrent effect of an injunction, the consumer is not likely to receive full compensation for past wrongs done to him.

If the plaintiffs are pursuing a primarily monetary remedy, a Rule 23(b)(3) class action may exist.[36] The first prerequisite of a (b)(3) class action is that questions common to the class as a whole must predominate over individual issues; this requirement could be met if the plaintiff proves that defendant engaged in a course of conduct common to the class.[37] Thus, for example, in a class

---

34 28 U.S.C. § 1332 (1970). Construction of this requirement raises numerous problems. *See generally* WRIGHT, *supra* note 32, §§ 24-31.

35 Although a suit seeking predominantly money damages would not qualify under this part of the rule, several courts have allowed 23(b)(2) actions involving damages or back pay when the primary relief sought was an injunction. *See* Nix v. Grand Lodge of Int'l Ass'n of Machinists, 479 F.2d 382, 385 (5th Cir. 1973); Sabala v. Western Gillette, Inc., 371 F. Supp. 385, 391 (S.D. Tex. 1974). *See also* Advisory Committee Note, 39 F.R.D. 98, 102 (1966); Rice, *Remedies, Enforcement Procedures and the Duality of Consumer Transaction Problems,* 48 B.U.L. REV. 559, 567-78 (1968).

36 *See generally* 7A C. WRIGHT & A. MILLER, FEDERAL PRACTICE & PROCEDURE: CIVIL §§ 1772-1774 (1972) [hereinafter cited as WRIGHT & MILLER].

37 There is little authority regarding the concept of predomination under 23(b)(3). It is clear that the mere existence of common questions is not sufficient. Instead the court must evaluate the relationship between common and individual issues to determine predominance. *See, e.g.,* Harrigan v. United States, 63 F.R.D.

action against an allegedly deceptive advertiser, whether the defendant's advertising is deceptive and whether the plaintiff class is likely to be damaged by the defendant's conduct are common issues and could well be found to predominate over individual issues such as reliance.[38] The case becomes more troublesome if the fraud is alleged to have been perpetrated by oral misrepresentations. Here individual issues may well predominate unless overwhelming evidence establishes a common solicitation pattern.[39]

The second prerequisite of Rule 23(b)(3) is that relief under that subdivision be superior to that provided through other means. Given the small individual amounts in question and the lack of incentive to bring suit, the class device may not only be a superior means; it may be the sole avenue of relief.[40]

Despite the usefulness of the class action, a combination of jurisdictional requirements and procedural complexities have all but denied access to the federal courts for small claimant classes. The Supreme Court in *Snyder v. Harris*[41] precluded aggregation of damages by class action claimants in order to meet the jurisdictional amount requirement for diversity jurisdiction,[42] holding that the amount in controversy could not be aggregated unless the cause of action alleged in behalf of the class was "joint and common"[43] as to the whole class. Mr. Justice Black, speaking for the majority, concluded that the 1966 revision of Rule 23 could not have changed the traditional requirements of jurisdictional amount.[44] And while the majority denied that its decision marked a return to the application of the narrow and confusing pre-1966

---

402 (E.D. Pa. 1974); Minnesota v. United States Steel Corp., 44 F.R.D. 559 (D. Minn. 1968).

38 *See* Dole, *The Emergence of Deceptive Advertising as a Group Tort: A Possible Consequence of the 1966 Federal Rule Amendment with Respect to Class Actions*, 62 Nw. U.L. Rev. 661 (1968).

39 An example might be a phone solicitation program under which the seller's agents read a prescribed text to each consumer. *See* Vasquez v. Superior Court, 4 Cal. 3d 800, 811-14; 484 P.2d 964, 971-73, 94 Cal. Rptr. 796, 803-05 (1971).

40 *But cf.* Katz v. Carte Blanche Corp., 496 F.2d 747 (3d Cir.), *cert. denied*, 419 U.S. 885 (1974) (test case, followed by individual suits, held superior in small claimant action), *noted in* 88 Harv. L. Rev. 825 (1975).

41 394 U.S. 332 (1969).

42 28 U.S.C. § 1332 (1970).

43 394 U.S. at 341.

44 *Id.* at 338.

categories of class suits,[45] the implication was undeniable: the Supreme Court was not going to accept the new Rule 23 as a panacea for small claimants.

The restrictive view of the Court in *Snyder* has been widely criticized.[46] Requiring all claimants to meet the ambiguous "joint and common" test in effect denies them their most likely avenue into the federal courts — diversity jurisdiction. The strict reading of the statutes in *Snyder* to limit federal jurisdiction seemed to reflect a fear that the federal courts would be engulfed by controversies of a purely local nature.[47] Nor was the case an aberration. The Court's adherence to *Snyder* led to the rejection of an attempt by a class representative to meet the jurisdictional amount for the whole class through his own claim, when none of the other class members had a $10,000 claim.[48]

Some loopholes in the aggregation prohibition exist.[49] In cases where a class of horse owners sued race track owners for failure to pay money under a purse contract,[50] and an employee class sued for failure to pay shares under a profit sharing pension plan,[51] there was a finding of a trust fund of sorts in which the class members had sufficiently unified interest to allow aggregation, although the claims were not technically "joint." *Snyder* was also avoided by having a state bring an action on behalf of its citizens in *The Drug Cases*,[52] an antitrust class action brought by several

45 *Id.* at 341; *see* FED. R. CIV. P. 23 (1939), setting forth what came to be known as "true," "hybrid" and "spurious" class actions. *See generally* Z. CHAFEE, SOME PROBLEMS OF EQUITY 245-65 (1950).

46 *See, e.g.*, WRIGHT, *supra* note 32, § 36, at 122-23. Indeed, Professor Wright has suggested that Congress amend the diversity statute to allow aggregation. *Id.*, § 73, at 317.

47 *See* 394 U.S. at 339-40. Yet at least one proposed statute would expressly incorporate state law into a federal consumer class action. S. 1378, 92d Cong., 1st Sess., § 4 (1971). *See generally* Newberg, *Federal Consumer Class Action Legislation: Making the System Work,* 9 HARV. J. LEGIS. 217, 234-36 (1971).

48 Zahn v. International Paper Co., 414 U.S. 291 (1973), *aff'g* 469 F.2d 1037 (2d Cir. 1972).

49 *See generally* Bangs, *Revised Rule 23: Aggregation of Claims for Achievement of Jurisdictional Amount,* 10 B.C. IND. & COM. L. REV. 601 (1969).

50 Berman v. Narragansett Racing Ass'n, 414 F.2d 311 (1st Cir. 1969).

51 Dierks v. Thompson, 414 F.2d 453 (1st Cir. 1969).

52 West Virginia v. Charles Pfizer & Co., 440 F.2d 1079 (2d Cir.), *cert. denied,* 414 U.S. 871 (1971) (approval of settlement of multiple private antitrust actions against drug manufacturers). *See generally* Wolfram, *The Antibiotics Class Actions,* 1976 A.B. FOUND. RES. J. 251.

states against manufacturers of antibiotic drugs. Although the court's opinion did not deal with the aggregation problem (because there is no amount in controversy requirement in actions brought under the antitrust laws),[53] its broad discussion of the ability of the state to maintain the suit might have provided a method for avoiding the jurisdictional amount problem in other contexts.[54] But later cases have demanded that the state possess an interest independent of its citizens as a prerequisite to a suit.[55]

Aggregation problems have not been alone in closing the federal courts to small claimant classes. Equally significant is a fundamental aspect of class action procedure: notice to absent members. In *Eisen v. Carlisle & Jacquelin*,[56] one of several million investors who had dealt on an odd-lot basis during a four-year period brought an antitrust action in which his own stake was only seventy dollars. When combined with the claims of other class members, however, the total damages sought amounted to several million dollars.

In the eight published opinions in *Eisen* since it was filed in 1966,[57] the courts have essentially rewritten the law concerning

53 Clayton Act § 4, 15 U.S.C. § 15 (1970).

54 *See generally* Note, *The Role of the States in Treble Damage Recovery Under the Federal Antitrust Laws: Rule 23 Class Actions and Parens Patriae*, 22 DRAKE L. REV. 155 (1972); Comment, *Suits by a State as Parens Patriae*, 48 N.C.L. REV. 963 (1970).

55 Hawaii v. Standard Oil Co., 405 U.S. 251 (1972), was an antitrust action for injuries to the state's "economy and prosperity" in which the Court held that the State could sue in its "proprietary capacity," but that to allow damages for injury to the State's "general economy" would open the door to duplicate recoveries. *See also* California v. Frito-Lay Inc., 474 F.2d 774 (9th Cir. 1973), *cert. denied*, 412 U.S. 908 (1973). In *The Drug Cases*, however, this "proprietary capacity" requirement would have been met, since the states had paid for the drug purchases of welfare recipients.

56 417 U.S. 156 (1974).

57 41 F.R.D. 147 (S.D.N.Y. 1966) (class certification denied lack of predomination of common questions); 370 F.2d 119 (2d Cir. 1966) (*Eisen I*) (denial of certification held reviewable final order; "death knell" doctrine); 391 F.2d 555 (2d Cir. 1968) (*Eisen II*) (denial of certification rev'd); 50 F.R.D. 471 (S.D.N.Y. 1970) (discussion of manageability and notice); 52 F.R.D. 253 (S.D.N.Y. 1972) (discussion of notice, costs of notice, and "fluid recovery"; preliminary hearing on the merits ordered); 54 F.R.D. 565 (S.D.N.Y. 1972) (order after preliminary hearing on the merits assessing 90% of notice costs on defendants); 479 F.2d 1005 (2d Cir. 1973) (*Eisen III*) (action dismissed as unmanageable; limited notice, hearing on distribution of notice costs, allocation of costs, and "fluid recovery" rejected; rehearing en banc denied); 417 U.S. 156 (1974) (*Eisen IV*) (court agrees that action is unmanageable, and that limited notice and hearing on allocation of cost thereof not authorized, but does not reach "fluid recovery" issue and remands for further determinations).

Rule 23(b)(3) class actions. Judge Tyler of the District Court attempted to institute a modified notice program under which the costs of notice were divided between the plaintiff and defendant[58] to satisfy the notice requirements of Rule 23(c)(2). But the court of appeals, and subsequently the Supreme Court,[59] rejected his limited notice plan as not enough to give the "best notice practicable."[60] As applied to *Eisen*, the "best notice practicable" meant that of the six million odd-lot investors joined in the suit, two million — those determined by the district court to be reasonably identifiable from computerized records kept by the defendants — had to receive individual notice by mail. Notice by publication for the remaining four million "unidentifiables" apparently was permissible. In so holding, the Court appeared to follow *Mullane v. Central Hanover Bank & Trust Co.*,[61] decided twenty-four years earlier. But it remains open to question whether the *Mullane* notion of individual notice wherever possible is mandatory as a matter of due process when the cost of such notice means that small claimant class actions cannot exist.[62]

A reading of the Advisory Committee's note to the revision of Rule 23 does not conclusively argue for either side in *Eisen*. Supporting the more restrictive view is the Committee's characterization of Rule 23(b)(3) as a category that "encompasses those cases in which a class action would achieve economies of time, effort and expense, and promote uniformity of decision as to persons similarly situated . . ."[63] There are commentators who suggest that the rule was designed primarily as an advanced joinder device to include in a single action all persons of a definable group in a situation where, absent the class action, at least a portion of

---

58 50 F.R.D. 471, 473 (S.D.N.Y. 1970).

59 479 F.2d at 1015, *aff'd* 417 U.S. 156, 178-79.

60 FED. R. CIV. P. 23(c)(2).

61 339 U.S. 306 (1950). The Court there held that notice by publication could not satisfy due process where the names and addresses of interested parties were known. Only where the names and addresses of interested parties were not known — and not reasonably ascertainable — was notice by publication approved. *Id.* at 317.

62 The Court, although it discussed *Mullane*, appeared to base its insistence on individual notice on Rule 23(c)(2) alone. *See* 417 U.S. at 177. See text accompanying notes 109-18 *infra*. *See generally* Note, *Managing the Large Class Action: Eisen v. Carlisle & Jacquelin*, 87 HARV. L. REV. 426, 434-41 (1973).

63 Advisory Committee Note, 39 F.R.D. 69, 102-03 (1966).

the group could and would sue as individuals.[64] Moreover, the Rule's requirement of predomination of common questions supports the Committee's conclusion that the device would not be useful in cases where individual questions of damages — or, arguably, reliance — still existed.[65] On the other hand, the Committee also appears to have expressly sanctioned the use of class actions for small claimants:[66]

> . . . the class may have a high degree of cohesion and prosecution of the action through representatives would be quite unobjectionable, *or the amounts at stake for individuals may be so small that separate suits would be impracticable.*

Yet despite the ambiguity in the historical background of Rule 23 and the desirability of allowing class actions to rectify consumer wrongs, resort to a federal forum is now largely impossible. No existing federal legislation grants consumers the right to sue for damages or equitable relief based on deceptive advertising, fraud, or various other related injuries; such actions must be predicated on state law, diversity of citizenship, and an amount in controversy in excess of $10,000 for each individual. A statutory remedy, then, is essential.

III. A Federal Consumer Class Action Proposal

A. *Jurisdiction*

Several considerations suggest that jurisdiction over most consumer class actions be vested exclusively in the federal courts. It may well be that there is no special competence in the area of consumer claims that dictates exclusive federal jurisdiction, but federal courts do possess far more experience in the management of complex class actions than the courts of most states.[67] In a case

---

64 *See* Landers, *Of Legalized Blackmail and Legalized Theft: Consumer Class Actions and the Substance-Procedure Dilemma,* 47 S. Calif. L. Rev. 842, 846 (1974). *See also* Wright, *Recent Changes in the Federal Rules of Procedure,* 42 F.R.D. 437, 567 (1966); Kaplan, *Continuing Work of the Civil Committee: 1966 Amendments of the Federal Rules of Civil Procedure,* 81 Harv. L. Rev. 356, 390 (1967).

65 *See* Advisory Committee Note, 39 F.R.D. 98, 103 (1966).

66 *Id.* at 104.

67 See text accompanying note 29 *supra.*

where a defendant is accused of a wrong that affects a geographically dispersed group, the waste of resources and the unfairness to the defendant that would arise in the prosecution of numerous identical claims in the courts of different states is apparent.[68]

The power of Congress to vest exclusive jurisdiction of certain actions in the federal courts, so long as the subject matter is within the legislative power — and hence the judicial power — of the national government is clear enough.[69] Similarly, Congress can provide that claims may be aggregated to meet a jurisdictional amount requirement.[70] Yet Congress may also decide that actions having an aggregate amount in controversy below a certain figure do not merit the attention of the federal judiciary and thus may provide that such actions shall not be heard in the federal courts. But it might well be stretching the point too far to provide for a minimum individual claim. If the aggregate harm to society is great enough, it should not matter that individual harms, as in *Snyder* or *Eisen*, are small.[71]

---

68 Indeed, the res judicata questions that might arise when one of these cases is concluded are disturbing. *See generally In re* Air Crash Disaster Near Dayton, Ohio, 350 F. Supp. 757 (S.D. Ohio 1972), *rev'd sub nom.* Humphreys v. Tann, 487 F.2d 666 (6th Cir. 1973); Note, *Collateral Estoppel of Nonparties,* 87 HARV. L. REV. 1485, 1492 (1974).

69 *See* HART & WECHSLER'S THE FEDERAL COURTS AND THE FEDERAL SYSTEM 418-19 (2d ed. P. Bator, P. Mishkin, D. Shapiro, & H. Wechsler 1973) [hereinafter cited as HART & WECHSLER]; *see also* Note, *Exclusive Jurisdiction of the Federal Courts in Private Civil Actions,* 70 HARV. L. REV. 509 (1957).

70 This follows from both the absence of any jurisdictional amount requirement, and the grant of congressional power to make "Regulations" concerning jurisdiction, found in Section 2 of Article III of the Constitution. *See also* Snyder v. Harris, 394 U.S. 332, 341-42 (1969) (dictum).

71 Section 6(b) of the model statute, *infra,* sets a jurisdictional amount requirement of $50,000, but sets no minimum figure for damages to the representative or any other class member. To be sure, this may exclude many meritorious claims where the damages, perhaps because of the local nature of the harm, simply are insufficient. Consumers thus are left to such state remedies as may exist. In the wake of *Snyder* and *Eisen,* some states have been re-examining their attitude toward small claimant class actions. *See, e.g.,* McMonagle v. Allstate Ins. Co., 227 Pa. Super. 205, 324 A.2d 414 (1974) (4 separate opinions on class actions procedure under state rule comparable to FED. R. CIV. P. 23 (1939)), *rev'd on other grounds,* 331 A.2d 467 (Pa. 1975) (court notes, in dictum, that it will not consider itself bound by federal decisions). But, with some exceptions, *cf.* note 29 *supra,* states have been far less accommodating toward class actions than the federal courts. Whether Congress could constitutionally require the states to adopt liberalized substantive and procedural rules for consumer class actions is an exceedingly complex question beyond the scope of this work. Senator Bayh's consumer class

Several class actions, directed against the same conduct, may be expected under such a jurisdictional grant, and it is quite possible that the asserted classes may be in conflict with the interests of each other or that two representatives in different courts may purport to represent the same class. It therefore becomes essential that there be some means to resolve class conflicts and coordinate all pretrial proceedings. Fortunately, this power already exists in the Judicial Panel on Multidistrict Litigation, established in 1968[72] to handle complex class action litigation.

As Professor Handler has noted, the Panel's power to order consolidated pretrial proceedings under one judge has been widely used[73] to define the scope of all classes involved[74] and to coordinate discovery. The savings of resources for both the parties and the judicial system under this unitary pretrial adjudication are apparent from *The Drug Cases,* in which 66 separate civil actions pending in 10 districts against 5 defendant drug companies were consolidated[75] and eventually settled.[76]

Opponents of consumer class actions argue that relaxation of standards will result in a flood of litigation and increased burdens on the courts.[77] This contention fails to note that the class action often encourages judicial economy in cases where individuals would otherwise bring suit. Admittedly, the typical consumer class

---

action bill of 1971 provided for just such an imposition. S. 1378, 92d Cong., 1st Sess., §§ 3-7 (1971). *See generally* Dice v. Akron, C.&Y.R.R., 342 U.S. 359 (1952); Note, *Is Federalism Dead? A Constitutional Analysis of the Federal No-Fault Automobile Insurance Bill: S. 354,* 12 HARV. J. LEGIS. 668 (1975). *See also* HART & WECHSLER, *supra* note 69, at 567-73.

72 Act of April 29, 1968, Pub. L. No. 90-296, § 1, 82 Stat. 109 (codified at 28 U.S.C. § 1407 (Supp. 1967)).

73 Handler, *The Shift from Substantive to Procedural Innovations in Antitrust Suits — The Twenty-Third Annual Antitrust Review,* 71 COLUM. L. REV. 1, 15-16 n.78 (1971), sets forth a long list of cases coordinated under the complex litigation statute between 1968 and 1970.

74 *Id.* at 16.

75 *See, e.g., In re* Antibiotic Drugs, 295 F. Supp. 1402 (Jud. Panel Multidist. Lit. 1968).

76 West Virginia v. Chas. Pfizer & Co., 314 F. Supp. 710 (S.D.N.Y. 1970), *aff'd* 440 F.2d 1079 (2d Cir.), *cert. denied,* 414 U.S. 871 (1971). See § 6(e) of model statute *infra.*

77 *See generally* McGough & Lerach, *Termination of Class Actions: The Judicial Role,* 33 U. PITT. L. REV. 445 (1972). *See also* Gardner, *Consumer Report: Proposals to Help Buyers Fight Back When Wronged Get Increased Consideration,* 5 NAT. J. REP. 981 (1973).

action will involve individual amounts too small to justify separate suits. But it is not to be assumed that these actions are without merit, and to foreclose these suits is therefore to deny the redress of substantive wrongs. If class actions increase the burden on the courts, the short answer is to press for expansion of facilities, not rejection of remedies.

B. *Protections and Prohibitions*

The scope of those to be protected by the statute should be as broad as possible if the remedy is to reach all in need.[78] Practical considerations, however, suggest two modifications to a definition reaching individuals only.

A business corporation, given its greater resources, is more likely to bring suit if it has been defrauded by a supplier than most individual consumers. Yet this does not compel the conclusion that a class action statute should not grant an equal remedy to business entities. It may well be that small corporations and partnerships lack the resources necessary to bring actions after suffering injury. When this happens, the entity's competitive position may be damaged, because it must absorb the loss or pass the injury through the market by raising its price. In either case, both the retail system and the individual consumer suffer from the lack of a remedy. And even if each member of a group of similarly injured businesses is able and willing to sue alone, courts should be able to hear a consolidated action whenever possible to prevent waste of the judicial system's resources. Indeed, this was one of the basic ideas that encouraged the development of the class action.[79] The danger of the multiple liability of a defrauding supplier to both a business and its customers, although significant, can be defused by consolidated assessment of business and consumer damages and possible application of the "passing-on" defense.[80]

---

78 See § 3 of model statute *infra*.

79 *See* Advisory Committee Note, 39 F.R.D. 98, 102-03 (1966) (FED. R. CIV. P. 23(b)(3)).

80 Hanover Shoe Co. v. United Shoe Machinery Co., 392 U.S. 481 (1968), severely limited the use of the "passing-on" defense in a private antitrust action, and *In re* Western Liquid A Asphalt Cases, 487 F.2d 191 (9th Cir. 1973), *cert. denied*, 415

The second modification is somewhat more difficult to justify. It seems clear that a consumer who refrains from a purchase because of misrepresentations suffers damage just as an actual purchaser does. Yet the damages are far more difficult to measure. It may be that an individual paid a higher price for a similar product; but not all consumer products are fungible, and the precise amount of damage would be difficult to calculate. An even greater assessment problem exists if a consumer does without the product entirely. Even if one later buys the product originally refused at a higher price, in which case the damages admit of easier proof, it becomes no easier to calculate the loss suffered during the time one uses either another product or none at all.

In any case, damages are not the sole issue for individual adjudication. A similar problem was noted by the Supreme Court in *Blue Chip Stamps v. Manor Drug Stores*,[81] in which the Court refused to permit securities fraud actions under Rule 10b-5[82] in cases where no purchase or sale of stock occurred.[83] Although the Court relied largely on construction of the applicable language, the concurring opinion noted the problems of subjective proof raised by a contrary holding.[84] One can much more confidently infer that a reasonable person would have reconsidered his purchase had he been aware of the true state of a material fact[85] and thus grant damages than one can be certain that the proximate cause of a failure to purchase was a particular misstatement. Other, more personal considerations may have been equally material — such as a sudden shortage of needed cash. Given all these individual matters at issue, a class action for damages would appear unworkable and unfair to defendants.[86] In any case, nothing

---

U.S. 919 (1974), refused to hold actions by remote purchasers precluded even where "passing on" of illegal overcharges could be demonstrated. The law in this area, however, is not settled. *See generally* 2 M. HANDLER, TWENTY-FIVE YEARS OF ANTITRUST 780-86 (1973).

81 421 U.S. 723 (1975).

82 17 C.F.R. § 240.10b-5 (1975) (promulgated in 1942 under the Securities Exchange Act of 1934 § 10(b), 15 U.S.C. § 78j(b) (1970)).

83 *See also* Birnbaum v. Newport Steel Corp., 193 F.2d 461 (2d Cir.), *cert. denied*, 343 U.S. 956 (1952).

84 421 U.S. at 757-58 & n.2 (Powell, J., concurring).

85 *Cf.* Mills v. Electric Auto-Lite Co., 396 U.S. 375, 384-85 (1970).

86 *Cf.* Blue Chip Stamps v. Manor Drug Stores, 421 U.S. 723, 758 (1975) (Powell, J., concurring) ("The subjective issues would be even more speculative in . . . class actions . . .").

would prevent action by consumer groups or competitors to urge the Federal Trade Commission to take action against misleading representations.[87]

In defining the kinds of activities to be proscribed by a consumer protection statute, two common approaches exist: the "types" listing and the blanket prohibition. In the "types" approach, a series of proscriptions are outlined which, when read together, purport to constitute the total universe of acts and practices covered by the statute.[88] The major difficulty with this approach is that no legislator has such Delphic insight as to make the original classifications inclusive enough to cover all possible future events. The "types" can be avoided by artful counsel, and cases which might have appeared clear in the mind of the legislator can be excluded by narrow judicial interpretation of the statute. Given these difficulties, an administrative structure for interpretation seems essential. Not only should there be power to make clarifying regulations to insure the application of the statutory purpose to new exigencies; there should also be power to exempt certain practices which, arguably, are found not to be within the legislature's intention.[89] No doubt the parameters of the "types" will also be developed by litigation. But reasonably explicit standards furnish the predictability needed for both fairness to the regulated industry and the workable functioning of the statute.[90]

In a blanket prohibition of the kind employed in Section 5(a) of the Federal Trade Commission Act, all "unfair or deceptive acts and practices" are made illegal.[91] The vagueness of this sort of statute aggravates the problems of predictability seen in the "types" approach, for the merchant can hardly know which acts

87 *See* Federal Trade Commission Act § 5(b), 15 U.S.C. § 45(b) (1973).

88 *See* California Consumers' Legal Remedies Act tit. 1.5, ch. 3, CAL. CIV. CODE § 1770 (West 1973), *as amended*, (West Supp. 1976); NATIONAL CONSUMER LAW CENTER, BOSTON COLLEGE LAW SCHOOL, NATIONAL CONSUMER ACT § 3.201 (First Final Draft 1970).

89 An act so drafted would not be unique. *See* Securities Exchange Act of 1934 § 16(b), 15 U.S.C. § 77(c) (1970).

90 Section 4(b) of the model statute, *infra*, an optional section, thus permits the Administrator of the Consumer Protection Agency to make such rules and regulations as are needed to make more clear and workable the statutory proscriptions, and also to exempt such acts and practices as he may find to be "not comprehended within the purpose of this Act."

91 15 U.S.C. § 45(a) (1970). *See also* MASS. GEN. LAWS ANN. ch. 93A, § 2 (1975).

or practices he presently conducts may be prohibited by the statute and cannot adjust his activities accordingly. Indeed, the dangers of haphazard development of guidelines through litigation are even more manifest here than under a "types" approach. Unless standards are established by an administrative apparatus before the legislation goes into effect, a private enforcement model would seem unworkable. Yet the administrative agency, knowing that its decisions could serve as a green light for plaintiffs' attorneys, could well become more hesitant in its adjudication of practices or promulgation of regulations, and the "slowdown effect" would probably be detrimental to consumer protection in general. Viewed together, these difficulties counsel rejection of the blanket prohibition.

C. *Fraud as a Common Question*

Without careful drafting, the multiplicity of elements of the cause of action may make a consumer class remedy unworkable. Most consumer schemes involve some kind of fraud; and if the plaintiff must prove individual reliance upon alleged material misrepresentations, there must be a trial with examination of the plaintiff concerning the written or oral statements made by the defendant and his agents.[92] If this is necessary, it is difficult to see how the needed predomination of common questions can exist.

Recent state and federal decisions, however, may have altered the substantive law of misrepresentation and fraud. In *Vasquez v. Superior Court*,[93] a class suit by purchasers of frozen food and freezers alleged knowingly false representation, reliance, and damages — all necessary elements in an action for fraud in California.[94] Rejecting defendant's arguments that the oral sales pitch made to each individual and the lack of proof as to absent members made a class action impossible, the California Supreme Court held that the standard form contract employed raised "at least a rebuttable implication" that there might also have been uniformity in the oral representations, their falsity, and their ma-

---

92 *See* W. PROSSER, HANDBOOK OF THE LAW OF TORTS § 103 (4th ed. 1971).

93 4 Cal. 3d 800, 484 P.2d 964, 94 Cal. Rptr. 796 (1971).

94 CAL. CIV. PRO. CODE § 382 (West 1973).

teriality.[95] From proof of these elements, said the court, an inference or a presumption of reliance could be drawn.[96]

In *Dolgow v. Anderson,*[97] a securities fraud action, the defendant opposed certification on the grounds that each member of the class must separately prove reliance, compliance with the statute of limitations, and damages since the alleged misrepresentations "were contained in a variety of statements, each of which were [sic] not necessarily brought to the attention of all investors. . . ."[98] The court, however, held that the plaintiffs need only allege a "continuous and common plan" to manipulate the price of stock in the defendants' statements to investors to establish the necessary commonality.[99]

Critics of these changes insist that reliance by each individual class member upon the misrepresentation or omission remains a necessary element,[100] and the procedural context in which both *Vasquez* and *Dolgow* arose lends some support to this argument.[101] But one noted commentator has suggested that it is not altogether clear that reliance is an independent and individual element of the cause of action,[102] and there is considerable case law to support

---

95 4 Cal. 3d at 813, 484 P.2d at 972, 94 Cal. Rptr. at 304 (1971).

96 Some authorities hold that reliance may be presumed — with the implication that the risk of non-persuasion is shifted to the defendant. RESTATEMENT OF CONTRACTS § 479 (1934); 12 S. WILLISTON, CONTRACTS, § 1515, at 480 (3d ed. W. Jaeger 1970). *See generally* Note, *The Reliance Requirement in Private Actions Under SEC Rule 10b-5,* 88 HARV. L. REV. 584 (1973). *But cf.* McNabb v. Thomas, 190 F.2d 606, 611 (D.C. Cir. 1951). The *Vasquez* court, however, refused to decide whether this presumption or a mere inference, which kept the risk of non-persuasion with the plaintiff, would arise. 4 Cal. 3d at 814, 484 P.2d at 973, 94 Cal. Rptr. at 805 (1971). *See generally* Note, *supra,* at 597-602 (1975).

97 43 F.R.D. 472 (E.D.N.Y. 1968).

98 *Id.* at 488.

99 *Id.* at 489.

100 *See, e.g.,* Handler, *The Shift from Substantive to Procedural Innovations in Antitrust Suits — The Twenty-Third Annual Antitrust Review,* 71 COLUM. L. REV. 1, 6-7 (1971); Simon, *Class Actions — Useful Tool or Engine of Destruction,* 55 F.R.D. 375 (1972); Note, *Impact of Class Actions on Rule 10b-5,* 38 U. CHI. L. REV. 337, 345-56 (1971).

101 The decision in *Vasquez* arose on a petition for a writ of mandate to compel the trial court to vacate its order sustaining demurrers, while *Dolgow* came on a motion for certification of a class action. *See* 4 Cal. 3d at 806, 484 P.2d at 967, 94 Cal. Rptr. at 799; 43 F.R.D. at 478.

102 Homburger, *Private Suits in the Public Interest in the United States of America,* 23 BUFFALO L. REV. 343, 370 (1974).

him.[103] Indeed, "[w]hat the critics have branded as a blatant disregard of substantive law engendered by the class action rule may simply be a new development in the substantive law, based on policy considerations quite unrelated to the class action rule."[104] At any rate, the power of Congress to draw on these decisions and fashion a new substantive basis for the action is incontestable.[105]

D. *Notice and Due Process*

The central procedural problems in class actions concern what is sometimes loosely called "manageability,"[106] and run throughout the life of the suit. The procedure must not only ensure that the litigation fully and finally disposes of the outstanding questions in a fair and expeditious manner; it must also deter the frivolous suit. Even the most notable defenders of the class action concede its inherent capacity for abuse,[107] and this possibility seems particularly strong in the consumer context, given the likelihood of widely dispersed class members and huge potential damages. But it does not follow that the present barriers to class actions are needed; instead, a procedure that revamps the current certification and trial practices is possible.

The needs of the parties and the court at the outset of the liti-

---

103 *See, e.g.*, Mills v. Electric Auto-Lite Co., 396 U.S. 375, 384-85 (1970); Affiliated Ute Citizens v. United States, 406 U.S. 128, 153-54 (1972); Royal Air Prop., Inc. v. Smith, 312 F.2d 210 (9th Cir. 1962); Ellis v. Carter, 291 F.2d 270 (9th Cir. 1961); Kardon v. National Gypsum Co., 73 F. Supp. 798 (E.D. Pa. 1947). *See generally* Note, *The Reliance Requirement in Private Actions Under SEC Rule 10b-5*, 88 HARV. L. REV. 584 (1975).

104 Homburger, *supra* note 102, at 371.

105 *Cf.* Securities Exchange Act of 1934 § 16(b), 15 U.S.C. § 77(c) (1970), which conclusively presumes that all short-swing profits from the purchase or sale of securities by officers, directors, and beneficial owners were made on the basis of inside information.

Section 7(d) of the model statute, *infra*, goes beyond *Vasquez* and *Dolgow* and shifts the risk of non-persuasion on all the elements of fraud to the defendant upon a prima facie showing by the plaintiff. Given the statutory basis, any argument that the *Dolgow* decision used a procedural device to enlarge a substantive right, and thus violated the Rules Enabling Act, 28 U.S.C. § 2072 (1970), becomes moot.

106 *See* FED. R. CIV. P. 23(b)(3)(d).

107 *See generally* Weinstein, *Some Reflections on the "Abusiveness" of Class Actions*, 58 F.R.D. 299 (1973).

gation are clear enough. Fairness to both plaintiff and defendant, as well as avoidance of waste of the resources of the parties and the judicial system, demand that the common cognizability of the claims be determined as soon as possible. In addition, all involved should learn whether the interests of the class are as homogeneous and as properly represented as is asserted by the class representative. It is to these ends that any notice requirements should be structured, avoiding the overly narrow reading of Rule 23 and *Mullane* given in *Eisen IV*.[108]

The *Eisen* Court seems to have forgotten that *Mullane* itself required, at bottom, a flexible approach to notice[109] based on a balancing between the individual interests involved and the interest of the judicial system in the final resolution of disputes.[110] It would seem that if the *Mullane* Court had construed the requirement that the "best notice practicable"[111] be given to absent members as strictly as did the appellate courts in *Eisen*,[112] it would have required that individual service of process be made on each ascertainable class member. That it did not do so,[113] and that it emphasized that "notice reasonably certain to reach most of those interested in objecting is likely to safeguard the interests of all,"[114] is evidence enough that due process need not be a straightjacket.

Another element deserves consideration. An absence of notice in *Mullane* would have meant that the right of the beneficiaries to object to the management of their trust funds would have been foreclosed by an impending consolidation into common funds. Thus a claim that might well be pressed in any case would be

---

108 See text accompanying notes 58-62 *supra*.

109 *See* 339 U.S. at 317-18:

> . . . [W]e have no doubt that . . . impracticable and extended searches are not required in the name of due process. The expense of keeping informed from day to day of substitutions . . . would impose a severe burden . . . . [W]ithin the limits of practicability notice must be such as is reasonably calculated to reach interested parties.

110 *Id.* at 314.

111 To be sure, Rule 23(c)(2) was not promulgated until 1966, but it purported to follow *Mullane*. *See* Advisory Committee Note, 39 F.R.D. 98, 106-07 (1966).

112 *See* 479 F.2d at 1009, 1017 & n.21 (2d Cir. 1973), 417 U.S. at 176-77 (1974).

113 Instead, the Court held that mailed notice was sufficient. 339 U.S. at 319.

114 *Id.*

prevented by an absence of notice. In a consumer class action, however, the problem of foreclosure is pre-existing; without a class device free from excessive procedural difficulties and costs, actions usually cannot be brought at all. It is difficult to see why the notice requirements of cases in which personal or property rights will be lost through the action of an adverse party must be read to apply to cases in which a representative attempts to gain recovery for absent members who apparently have the same interests as their champion.[115]

It follows that the central inquiry should be directed to the nature of notice needed to insure that an individual's interests are properly advocated.[116] This is not so much a question of whether there are due process concerns of notice, for indisputably there are;[117] instead, the question is "what process is due"[118] in order to ensure adequate representation.

An assessment of adequacy of representation requires, in turn, an examination of the nature of the class interests, and it is critical to distinguish among congruent, diverging, and conflicting interests. It is said that the named plaintiff's interests must be fully coterminous with those of the class he represents in order to insure that the advocacy of the class interest is as zealous as possible.[119] To be sure, there is a danger that when interests are diametrically opposed to each other, token advocacy will lead to

---

115 Given the widely dispersed character of the class, reliance on Schroeder v. City of New York, 371 U.S. 208 (1962), for the proposition that individual notice is necessary as a matter of due process whenever practicable seems clearly inapposite. Even though *Schroeder* involved a single individual's real property rights where the person's identity was "very easily ascertainable," 371 U.S. at 212-13, it seemed to be read by both the Second Circuit and the Supreme Court in *Eisen IV* as requiring that the same sort of individual notice is essential in a 23(b)(3) class action. *See* 479 F.2d at 1017 n.21; 417 U.S. at 175.

116 The Supreme Court curtly rejected the argument that adequate representation is "the touchstone of a class action" by noting that "quite apart from what due process may require, the command of Rule 23 [(c)(2)] is clearly to the contrary." 417 U.S. at 177. But this does not foreclose a notice scheme that both overcomes the obstacles of Rule 23 and is harmonious with the demands of due process.

117 *Cf.* Hansberry v. Lee, 311 U.S. 32 (1940).

118 Morrissey v. Brewer, 408 U.S. 471, 481 (1972).

119 *See* WRIGHT, *supra* note 32, § 72, at 308; 7A WRIGHT & MILLER, *supra* note 36, § 1768, at 638.

an unfavorable judgment binding on the entire class.[120] Moreover, although collateral attack is available to question the prior judgment, economy and fairness demand that this remedy be used most sparingly as a means of uncovering conflicting interests. But this argument loses most of its force when applied to cases where interests merely diverge, *i.e.*, are primarily concerned with only slightly different wrongs or parties. Here the concern for wholehearted advocacy often degenerates into formalism, as, for example, where it was held that a class representative could not prosecute a claim which involved precisely the same wrong and legal theory against a second defendant who had wronged the class but not the representative himself.[121] This attitude mindlessly equates slightly differing desires with antagonism to the needs of others. It does little but require the needless substitution of another representative or the wasteful dismissal of a meritorious action which has already consumed resources of the parties and the courts. Most importantly, it ignores the realities of the class suit.

Particularly in a consumer class action statute, the central concepts are that it is a group that is represented, and that the cause of action and the relief sought are more social than private. Hence it is the "private attorney general,"[122] seeking both monetary reward[123] and the enforcement of public policies,[124] who controls the fundamental decisions and direction of the suit — not the named plaintiff.

The question thus becomes not whether the representative's

---

120 Should the interests of the class members attach to largely different claims or theories, there is the danger that one representative will press part of the suit with such little vigor that some members are denied due process of law. Hansberry v. Lee, 311 U.S. 32 (1940); Battles v. Braniff Airways, Inc., 146 F.2d 336 (5th Cir. 1944), *cert. denied*, 325 U.S. 871 (1945); McCarthy v. Director of Selective Service Sys., 322 F. Supp. 1032 (E.D. Wis. 1970), *aff'd on other grounds* 460 F.2d 1089 (7th Cir. 1972).

121 La Mar v. H & B Novelty & Loan Co., 489 F.2d 461 (9th Cir. 1973).

122 Associated Indus. v. Ickes, 134 F.2d 694, 704 (2d Cir. 1943) (Frank, J.).

123 A quite similar monetary structure exists in the contingent fee personal injury suit. Here the injured party lacks sufficient funds to pay for the litigation expenses himself. Hence the attorney finances the action, gaining reimbursement for his efforts only if he gains a judgment or adequate settlement. The party's interests, to be sure, are more personal than related to those of society.

124 *Cf.* Newman v. Piggie Park Enterprises, 390 U.S. 400, 402 (1968).

injuries or theories are precisely the same as those of all members of the class, but rather whether the representative's attorney sufficiently presses the desires of the group he seeks to represent.[125] Subject to due process limitations, this should be — and, indeed, has often been — the interpretation given to the present requirement that a representative "fairly and adequately" protect the interests of the class,[126] even in complex cases.[127] As Professor Wright has noted: "[T]he quality of the representation is more important than the numbers, and . . . even a single representative of the class may be enough."[128] When the possibility of significantly diverging or conflicting interests exists, however, the court has ample power to order the creation of subclasses or to limit a representative's role in the litigation.[129]

The nature of the interests involved will, to be sure, affect the court's decision on whether common questions predominate over individual issues in the suit. But there is more to be considered. First is the question of stereotypical conduct toward the class which, if found, should serve to provide a common issue on the various elements of a fraud claim.[130] A second question is whether the trial can be structured to resolve damages as a common issue and to provide for decision on individual damages in a later proceeding.[131]

A somewhat extensive use of the court's resources may be necessary to determine these issues. But long-term judicial economy,

---

125 This does not require a wholesale destruction of the theory of standing to sue. *Cf.* U.S. CONST. art. III, § 2 ("cases and controversies"). Indeed, the courts are willing to permit named plaintiffs to remain in a class action in many cases even though their right to assert a claim has passed because of the passage of time or other inexorable or fortuitous circumstances. *See* Sosna v. Iowa, 419 U.S. 393, 400-03 (1975); Roe v. Wade, 410 U.S. 113, 125 (1973). *See generally* WRIGHT, *supra* note 32, § 12, at 34-35.

126 FED. R. CIV. P. 23(a)(4).

127 *See, e.g.*, Sosna v. Iowa, 419 U.S. 393, 403 & n.13 (1975) (constitutional attack on durational residency requirement); Dolgow v. Anderson, 43 F.R.D. 472, 495-97 (E.D.N.Y. 1968) (securities fraud); Siegel v. Chicken Delight, Inc., 271 F. Supp. 722 (N.D. Cal. 1967) *aff'd*, 448 F.2d 43 (9th Cir.), *cert. denied* 405 U.S. 955 (1972) (antitrust tie-in claim).

128 WRIGHT, *supra* note 32, § 72, at 309 (citations omitted).

129 FED. R. CIV. P. 23(c)(4), 23(d)(1)-4. The provisions of these rules are incorporated in §§ 8(b)(1), 8(c)(2)-(3) of the model statute *infra*.

130 See text accompanying notes 93-104 *supra*.

131 See text accompanying notes 172-84 *infra*.

fairness to the parties, and deterrence against strike suits demand that scrutiny of the adequacy of representation, delineation of class interests, and determination of whether common, manageable issues predominate should be made as soon as practicable after the action is brought, as under prior practice.[132] This does not mean, however, that the court's findings need be as much a shot in the dark as is now the case. A flexible procedure can be devised that will avoid both the Scylla of an overly lenient judicial attitude[133] and the Charybdis of *Eisen.*

The approach used by Judge Tyler in *Eisen*[134] commends itself to these needs. If it is not clear to the court that common questions do indeed predominate and that the class's interests are adequately represented in the litigation, there should be power to compel the representative to show that what he has pleaded is probably the case. To this end, the court should be able to order such notice as it believes necessary to encourage other class members with differing interests to come forward and present their objections[135] or motions for intervention.[136] If at all possible, individual notice by mail should be sent to those people who have the largest monetary stakes in the litigation since such people are prima facie the most likely to intervene.[137] A further necessity should be publication in such newspapers or periodicals as are most likely to reach the attention of other members.[138] Finally, notice might well be sent to a further sample of alleged class members.[139] The available evidence indicates that few members, when

---

132 *See* FED. R. CIV. P. 23(c)(1).

133 Indeed, it has been argued that the lack of a method for making preliminary findings has made Rule 23(b)(3) far too receptive to marginal or frivolous claims. Katarincic & McClain, *Federal Class Actions Under Rule 23: How to Improve the Merits of Your Action Without Improving the Merits of Your Claim,* 33 U. PITT. L. REV. 429, 443 (1972); *cf., e.g.,* Esplin v. Hirschi, 402 F.2d 94, 99 (10th Cir. 1968) (". . . [I]f there is to be an error made, let it be in favor and not against the maintenance of the class action, for it is always subject to modification. . . ."). This attitude has sometimes caused class determinations to be unreasonably delayed. *See, e.g.,* Berland v. Mack, 48 F.R.D. 121 (S.D.N.Y. 1969) (two years).

134 52 F.R.D. 253 (S.D.N.Y. 1971).

135 *Cf.* Mullane v. Central Hanover Bank & Trust Co., 339 U.S. 306, 314 (1950).

136 FED. R. CIV. P. 24.

137 *See* Eisen v. Carlisle & Jacquelin, 52 F.R.D. 253, 267 (S.D.N.Y. 1971).

138 *Id.* at 268.

139 *Id.* at 267 & n.11.

notified, seek to exclude themselves from the action[140] — if, indeed, the members understand what the notice is about.[141] But any significant response may well inform the court about the nature of the class or suggest that the representation is not as adequate as alleged.[142] Should this be the case, the court has ample power to divide the class, dismiss a part of the alleged class from the litigation, or limit the participation of any parties seeking to intervene.[143]

The notice should have, however, a more practical effect in addition to satisfying a court's due process concerns. The cost of the notice, since it will ordinarily be borne by the representative, should serve as a more than adequate deterrent to the bringing of strike suits. It does not follow, of course, that the merits of the claim must be in direct proportion to the size of the class represented. But no rational plaintiff's attorney can be expected to bring an action unless he is reasonably certain that the costs of litigation can later be recovered.

Notice alone may well not be enough to persuade the court that a proper class action is before it. Hence the additional procedural tool of a preliminary hearing may be necessary. An examination may be warranted to determine several issues: whether the class is prima facie adequately represented; the extent to which common questions exist and predominate;[144] and whether the plaintiff should be required to bear the cost of notice.[145] If it appears diffi-

---

140 *See* Pomerantz, *New Developments in Class Actions — Has Their Death Knell Been Sounded?*, 25 Bus. Law. 1259, 1266 (1970).

141 Miller, *Problems in Giving Notice in Class Actions*, 58 F.R.D. 313, 321-22 (1973), sets forth some confused replies from persons notified of the settlement in *The Drug Cases.*

142 Eisen v. Carlisle & Jacquelin, 52 F.R.D. 253, 268 (S.D.N.Y. 1971); Feder v. Harrington, 52 F.R.D. 178, 182 (S.D.N.Y. 1970).

143 Fed. R. Civ. P. 23(c)(4)(B), (d)(1), (d)(3). The costs might be paid from the fund created by settlement or trial. Or the court might order that they be paid by the losing party. See text accompanying notes 257-62 *infra.*

144 Given the substantive provisions of § 7 of the model statute, *infra*, this burden should not be excessively high; see text accompanying notes 93-104 *supra.* The court should be flexible in determining whether any individual issues can be resolved in a later proceeding. See text accompanying notes 172-84 *infra.*

145 One court, in making this decision, considered the following factors: the merit of the plaintiff's claim; the defendant's desire for res judicata as to the class; the ability of the named plaintiffs to pay for the notice; the percentage of the class's claim held by the representatives and the financial responsibility of the

cult to resolve the questions of adequate representation and common questions without notice and intervention, a hearing on the merits, somewhat in the nature of a motion for a preliminary injunction,[146] could be used to make an assessment of notice costs.[147] Some discovery, the cost of which would be borne by each party, will probably be needed;[148] but this has the salutary effect of both enlightening the court and deterring the baseless claim.[149] The prospect of hearing followed by notice and then further hearing will undoubtedly strike some as more than a court can bear. Nevertheless, the court's flexibility, the co-operation of counsel, and the general principle that the meritorious action should not be denied its day in court ought to outweigh such momentary considerations. Indeed, the potential avoidance of later wasteful litigation from use of these screening devices makes the alternative appear penny wise and pound foolish. None of these devices, in any case, need preclude the later use of summary judgment[150] as an additional method to terminate the frivolous suit.

E. *Settlement*

The active role of the court in the management of class actions extends to scrutiny of proposed settlements. Prior practice provides a useful background, for many guidelines[151] have developed

---

named plaintiffs. Feder v. Harrington, 52 F.R.D. 178, 182 (S.D.N.Y. 1970); *see also* Berland v. Mack, 48 F.R.D. 121, 132 (S.D.N.Y. 1969); Eisen v. Carlisle & Jacquelin, 54 F.R.D. 565 (1972).

146 *Cf.* Eisen v. Carlisle & Jacquelin, 52 F.R.D. 253, 271 (S.D.N.Y. 1971).

147 The concern voiced by Judge Medina in *Eisen III*, 479 F.2d 1005, 1115 (2d Cir. 1973), that such a determination would prove prejudicial is amply rebutted in Note, *Managing the Large Class Action: Eisen v. Carlisle & Jacquelin*, 87 HARV. L. REV. 426, 443-46 (1973).

148 Judge Tyler ordered such discovery in *Eisen*. *See* 52 F.R.D. at 272.

149 Should the parties fear that discovery might be used for mere snooping, the court has ample power to issue whatever protective orders are necessary. FED. R. CIV. P. 26(c).

150 FED. R. CIV. P. 56. Indeed, the court might employ the little-used practice of apportioning the costs of discovery according to the preliminary merits as a further deterrent to the doubtful claim. *See* River Plate Corp. v. Forestal Land, Timber & Ry. Co., 185 F. Supp. 832, 838 (S.D.N.Y. 1960).

151 *See, e.g.*, Teachers Ins. & Annuity Ass'n v. Beame, 67 F.R.D. 30, 34 (S.D.N.Y. 1975); American Finance Sys. Inc. v. Harlow, 65 F.R.D. 572 (D. Md. 1974); Held v. Missouri Pacific R.R., 64 F.R.D. 346, 347-48 (S.D. Tex. 1974); Wainwright v.

under Rule 23(e)'s requirement that no class action "be dismissed or compromised without the approval of the court."[152]

Generally, the court's duty is to act as the guardian of the rights of absent parties[153] and to ensure that a fair and reasonable compromise[154] has been reached without fraud or collusion. Courts have laid down general principles of review[155] that recognize the general policies of judicial economy and private resolution that favor settlements.[156] But the burden of showing that the settlement is fair rightly rests on the proponents of the plan.[157] As a further aid to determine the propriety of the proposal, the court needs the power to order that notice be given to absent class members.[158]

Kraftco Corp., 58 F.R.D. 9, 11 (N.D. Ga. 1973). *See generally* MANUAL FOR COMPLEX LITIGATION § 1.46 (1973); 3B J. MOORE, FEDERAL PRACTICE, ¶ 23.80 (2d ed. 1975).

152 *Cf.* FED. R. CIV. P. 41(a)(1) (voluntary dismissal of other actions without court approval permitted). The clear purpose of this provision is "to protect the nonparty members of the class from unjust or unfair settlements affecting their rights when the representatives become faint-hearted before the action is adjudicated or are able to secure satisfaction of their individual claims by a compromise." 7A WRIGHT & MILLER, *supra* note 36, § 1797, at 226.

153 *See, e.g.,* Norman v. McKee, 431 F.2d 769, 774 (9th Cir. 1970), *cert. denied,* 401 U.S. 912 (1971) (proposed settlement disapproved because settlement would benefit only future investors and not present members of the affected class).

154 *See, e.g.,* Percodani v. Riker-Maxson Corp., 50 F.R.D. 473, 477 (S.D.N.Y. 1970) (class action seeking $12,000,000; proposed settlement of $1,800,000, absent proof that the defendant was unable to pay more, disapproved as not fair and reasonable where court determined that plaintiff's claim had considerable merit).

155 As one court put it:

> (1) A compromise settlement involves mutual sacrifice in order to prevent unprofitable litigation.
> (2) Vindictive motives or pressures are not proper in settlement negotiations.
> (3) The recommendation of acceptance by experienced and competent counsel is a fact entitled to great weight.
> (4) The role of the court is somewhat constrained in that its business judgment is not to be substituted for that of the parties who worked out the settlement accord unless the settlement, taken as a whole, appears so unfair on its face as to preclude judicial approval.

Percodani v. Riker-Maxson Corp., 50 F.R.D. 473, 477-78 (S.D.N.Y. 1970).

156 *See* Feder v. Harrington, 58 F.R.D. 171, 174 (S.D.N.Y. 1970).

157 The proponents of the settlement have the burden of proving:

> (1) that it is not collusive but was arrived at after arm's length negotiation; (2) that the proponents are counsel experienced in similar cases; (3) that there had been sufficient discovery to enable counsel to act intelligently; and (4) that the number of objectants or their relative interest is small.

Id. at 174-75 (citations omitted).

158 Under present practice, the manner of notice of settlement is for the dis-

Whether a proposed settlement is fair depends on more than the likelihood of success at trial. A court must also consider the probable amount of recovery, discounted by the costs of litigation; the plan for distribution of the settlement; and the extent to which any class members, after notification, object to the proposal.[159] In some cases, however, a court ought to require an evidentiary hearing,[160] particularly if serious objections are raised to the settlement's amount[161] or validity.[162] The necessity for and terms of the hearing should be a matter for the trial court's discretion and reversible only upon a clear showing of abuse.[163]

The response by absent class members to notice of settlement should indicate to the court whether such a hearing is required. Thus the nature and extent of notice[164] should be designed to inform and reach those parties most likely to object.[165] As a deterrent against collusive settlements, the notice should fully disclose the consequences of the proposal, the amount and division of the settlement fund, and the counsel fees and other expenses to be charged against the class.[166] A history of the litigation and the circumstances of settlement should be included to make the details of the plan understandable.

---

cretion of the court. FED. R. CIV. P. 23(e). *See generally* 7A WRIGHT & MILLER, *supra* note 36, § 1797, at 232.

159 *See* Cherner v. Transitron Electronic Corp., 221 F. Supp. 48, 53 (D. Mass. 1963) (Wyzanski, J.). *See also* Grunin v. International House of Pancakes, 513 F.2d 114 (8th Cir. 1975); Norman v. McKee, 431 F.2d 769 (9th Cir. 1970), *cert. denied,* 401 U.S. 912 (1971); Detroit v. Grinnell Corp., 356 F. Supp. 1380 (S.D.N.Y. 1972) (antitrust claim), *aff'd,* 495 F.2d 448 (2d Cir. 1974). *See generally* 7A WRIGHT & MILLER, *supra* note 36, § 1797, at 230-31.

160 See § 8(a) of model statute, *infra.*

161 *Cf.* Detroit v. Grinnell Corp., 356 F. Supp. 1380, 1384-86 (S.D.N.Y. 1972) (hearing refused where objector's expert would show only that provable damages would be 3%-7% of total billings and settlement would provide for recovery of 3.2%-3.7% of total billings).

162 *Cf.* Calhoun v. Cook, 487 F.2d 680 (5th Cir. 1973), *rev'g* 362 F. Supp. 1249 (N.D. Ga. 1973) (entry of order enforcing an alleged settlement without a hearing improper where fact concerning the existence of agreement in dispute).

163 *See The Drug Cases,* 440 F.2d 1079, 1085 (2d Cir. 1971), *cert. denied,* 404 U.S. 871 (1971). *See also* Girsh v. Jepson, 521 F.2d 153 (3d Cir. 1975); Cranston v. Harden, 504 F.2d 566 (2d Cir. 1974); Kohn v. American Metal Climax, Inc., 489 F.2d 262 (3d Cir. 1973).

164 Sections 8(c)(4)(i)-(ii) and 8(d) of the model statute, *infra,* follows present practice under Rule 23(e). See note 158 *supra.*

165 *Cf.* Mullane v. Central Hanover Bank & Trust Co., 339 U.S. 306, 319 (1950).

166 *See* MANUAL FOR COMPLEX LITIGATION § 1.46 (1973).

F. *Common Damages and Fluid Recovery*

Even if notice problems are resolved, there remains the question of how damage issues are to be adjudicated within the class action structure. The Second Circuit in *Eisen II* noted its reluctance "to permit actions to proceed where they are not likely to benefit anyone but the lawyers who bring them."[167] There is, indeed, some weight to the view that vindication of rights, enforcement of public policies, and deterrence of wrongdoers are all frustrated when it becomes difficult to channel the monetary recovery to the people affected. In credit card sales, the records of transactions are usually in the possession of the defendant, and notifying the class members and obtaining accurate records of damages is relatively simple. When sales are in cash, however, no records may be available, and calculation of damages is more difficult.

Some courts have thought it advantageous to establish the extent of the claims and claimants early in the trial and have ordered that class members be notified to file a statement of their claims within a reasonable period or lose the right to assert them.[168] This procedure provides the court and the litigants with a fairly good idea of the identity of the potential claimants and the amount of their asserted damages, but would be counterproductive in small claimant class actions where the absent class members have little incentive to respond even if found.[169] In a statute that is grounded on a policy of maximum compensation to all class members, this procedure is unacceptable on its face.

Some courts seem to have treated the problem of ascertainment of damages as irrelevant to class action certification.[170] Others, however, have concluded that damage problems make otherwise cognizable class actions unmanageable per se.[171] The possibilities

167 391 F.2d 555, 567 (2d Cir. 1968).

168 *See, e.g.,* Biechele v. Norfolk & W.Ry., 309 F. Supp. 354 (N.D. Ohio 1969); Philadelphia Elec. Co. v. Anaconda Am. Brass Co., 43 F.R.D. 452 (E.D. Pa. 1968); Harris v. Jones, 41 F.R.D. 70 (D. Utah 1966).

169 See text accompanying notes 140-41 *supra*.

170 *See, e.g.,* Green v. Wolf Corp., 406 F.2d 291 (2d Cir. 1968), *cert. denied,* 395 U.S. 977 (1969); Foster v. Detroit, 405 F.2d 138 (6th Cir. 1968); Alameda Oil Co. v. Ideal Basic Indus., Inc., 326 F. Supp. 98 (D. Colo. 1971); Illinois v. Harper & Row Publishers, Inc., 301 F. Supp. 484 (N.D. Ill. 1969).

171 *See, e.g., Eisen III,* 479 F.2d 1005, 1018 (2d Cir. 1973).

that a lawsuit will consume tremendous litigant and judicial resources before it is suddenly dismissed, or that class claims with merit will never reach the court because of damage ascertainment problems, compel the consideration of some innovative alternatives. In cases where there has been a fixed overcharge, such as was asserted in *Eisen*,[172] damage to the class of purchasers is relatively easy to calculate from the records of sales kept by defendants. Indeed, data on quantity and price can suffice even if records have not been kept for each member of the class.[173] Alternatively, econometric models and volume sales breakdowns can be used. If these methods prove technical or confusing to the trier of fact, there is always the possibility that a special master can be appointed to aid in sorting out expert testimony or in making calculations.[174]

From time to time it has been asserted that even in these cases a calculation of each individual's damage must be made separately because to do otherwise would deprive a party of either due process[175] or the right to jury trial.[176] This contention, viewed against long-standing precedent, seems plainly unfounded. It is well settled that when a wrong has been proved, the difficulty of ascertainment of damages is not to be "confused with right of recovery."[177] The cases fairly shout that almost any assessment which is not wholly speculative will be accepted, especially when the conduct of the defendant has made it difficult for the plaintiff to show how much damage he has suffered.[178] In employment discrimination cases, the difficulties of individual determinations of back pay have not prevented courts from requiring that an overall award of back pay be computed — and reversing lower courts

---

172 Eisen v. Carlisle & Jacquelin, 52 F.R.D. 253, 261-62 (S.D.N.Y. 1971).

173 *Cf.* Philadelphia v. American Oil Co., 53 F.R.D. 45, 71 (D.N.J. 1971).

174 *See* FED. R. CIV. P. 53.

175 *Eisen III*, 479 F.2d 1005, 1018 (2d Cir. 1973) (dictum).

176 Handler, *The Shift from Substantive to Procedural Innovations in Antitrust Suits — The Twenty-Third Annual Antitrust Review*, 71 COLUM. L. REV. 1, 7 (1971).

177 Story Parchment Co. v. Paterson Parchment Paper Co., 282 U.S. 555, 565 (1931).

178 Bigelow v. RKO Radio Pictures, Inc., 327 U.S. 251, 264 (1946); Union Carbide & Carbon Corp. v. Nisley, 300 F.2d 561, 590 (10th Cir.), *petition for cert. dismissed*, 371 U.S. 801 (1962); F.W. Woolworth Co. v. NLRB, 121 F.2d 658 (2d Cir. 1941).

that fail to do so.[179] Particularly when strong public policies are involved, a defendant should properly bear the risk of whatever uncertainty his wrongful conduct has produced.[180] And this principle should apply with equal force to cases where the individual's amount of possible recovery is problematical but his membership in the class of affected persons is not.[181] One should keep in mind that the defendant also gains an important benefit — a final, binding adjudication of the rights of the class against him.[182]

After adjudication of the total damages, notice should go out informing class members of the opportunity to present their claims[183] to whatever administrative or quasi-judicial body the court establishes to draw up claim forms and process individual applications.[184] Uncollected damages will probably remain. But rather than permit the defendant to retain what has been determined to be the fruits of wrongdoing, there should be a number of alternatives available to the court. One possibility suggested by a number of commentators is a "fluid recovery" distribution to a "next-best" class composed to the greatest extent possible of absent class members.[185] This method of finding a next-best re-

---

179 *See* Pettway v. American Cast Iron Pipe Co., 494 F.2d 211, 251-63 (5th Cir. 1974).

180 Bigelow v. RKO Radio Pictures, Inc., 327 U.S. 251, 265 (1946).

181 *But cf.* Malina, *Fluid Class Recovery as a Consumer Remedy in Antitrust Cases,* 42 N.Y.U.L. REV. 477, 487-91 (1962).

182 *Cf. In re* Antibiotic Antitrust Actions, 333 F. Supp. 278, 281, 287 (S.D.N.Y. 1971), *mandamus denied,* 440 F.2d 119 (2d Cir. 1971).

183 Shapiro, *Processing the Consumer's Claim,* 41 ANTITRUST L.J. 257, 268-69 (1971), includes several suggestions on use of the media to insure maximum notice. See also § 8(c)(4)(ii) of model statute *infra.*

184 In *The Drug Cases,* some 42,000 individual claims were processed by an attorneys' committee which examined all relatively large claims and investigated a sample of others. A wide variety of methods were used to inform class members of their opportunity to file claims, and the total cost of processing was relatively small. *See* Eisen v. Carlisle & Jacquelin, 52 F.R.D. 253, 259 (S.D.N.Y. 1971) (description of claim procedure). *See also* Shapiro, *Processing the Consumer's Claim,* 41 ANTITRUST L.J. 257 (1972). In Cherner v. Transitron Electronic Corp., 201 F. Supp. 934 (D. Mass. 1962), a class action under the Securities Act of 1933, the use of a special master to process shareholders' claims resulted in significantly higher administration expenses relative to those in *The Drug Cases. See* Eisen v. Carlisle & Jacquelin, *supra,* at 259. *See generally* Dam, *Class Actions: Efficiency, Compensation, Deterrence, and Conflict of Interest,* 4 J. LEGAL STUDIES 47 (1975); Jacoby & Cherkasky, *The Effects of Eisen IV and Proposed Amendments of Federal Rule 23,* 12 SAN DIEGO L. REV. 1 (1974); Note, *Damage Distribution in Class Actions: The Cy Pres Remedy,* 39 U. CHI. L. REV. 448 (1972) [hereinafter cited as *Cy Pres*].

185 See *Cy Pres, supra* note 184, at 452-65; Jacoby & Cherkasky, *supra* note 184, at 20-24.

cipient for the funds derived by the class suit is analogous to the doctrine of cy pres in the law of charitable trusts.[186]

The court's function here is to apply its discretion to develop a remedial scheme that effectuates as closely as possible the intent of the legislature in providing legal remedies for consumers. When it is impossible to reach the class of absent consumers actually harmed by the defendant's activities, the court should apply the funds to a group whose members most nearly approximate the affected class. If it becomes impossible to ascertain such a group, the court can attempt to determine which alternate recipients the Congress would prefer; in some cases this may mean devoting the funds to a broader public service in order to benefit society as a whole.[187]

The number and kind of approaches available for disposition of the uncollected damages is limited only by the imagination of the court. One commentator has suggested three possible approaches under a cy pres-like theory: distribution to those class members who come forward to collect their damages; distribution through the state in its capacity as *parens patriae* or by escheat; and distribution through the market.[188]

Under the first approach, those members who come forward divide the fund among themselves. These persons may be the "best class" in the sense that they have been interested enough to come forward to press their claims, but it is apparent that this approach results in a windfall for them. Given the realities of the small claimant action, where many may be expected not to bother with their potential — even if provable — claims, exclusive reliance on this approach seems clearly deficient. In a *parens patriae* or escheat remedy, the government would take all uncollected

---

186 The cy pres doctrine is a rule of construction used by courts to effectuate testamentary charitable gifts that would otherwise fail because an inherent flaw rendered the gift impossible. The court attempts to determine the testator's intent and seeks an alternative that will best serve the testator's original purpose. *See* E. FISCH, THE CY PRES DOCTRINE IN THE UNITED STATES 128, 216-18 (1950). *See also Cy Pres, supra* note 184, at 452-63.

187 In *The Drug Cases,* 440 F.2d 1079 (2d Cir. 1970), *aff'g* 314 F. Supp. 710 (S.D.N.Y. 1970), the Attorney General of Illinois, on behalf of Illinois consumers, recovered $4.5 million in a private antitrust action and subsequently announced that a substantial portion of the damages would be used to finance future public health projects in the state. [1971-73 Transfer Binder] CCH CONSUMERISM 162 (1971).

188 *See generally Cy Pres, supra* note 184, at 452-63.

funds and distribute them to a much wider class of persons using the same service, or living in the same general geographic area. The money, however, should not go into the general fund because such an unrestricted recovery, even though a deterrent, would appear to be more a fine or penalty than compensatory damages.

Judge Tyler's opinion in *Eisen*[189] is perhaps the best example of a plan of fluid recovery through distribution in the market.[190] Under his procedure, there would be an initial period after gross damages were determined during which each class member could come forward and prove his damages. Any uncollected residue would then be used to reduce the odd-lot differential in all transactions involving the defendants. The court indicated that this method would satisfy both its concern that class members actually share in any eventual recovery and its desire to avoid the retention of illegal profits by the defendants.[191]

Employment of a fluid recovery device presents several difficulties. In a market distribution of damages through price reduction, a court would be required to ensure that the reduction equals the uncollected or uncollectable damages as closely as practicable. Moreover, the court would have to prevent any quality reduction by the defendant that would decrease both the value of the good or service and the compensation received by the purchasing class. Finally, a lowering of price may adversely affect competition.[192]

189 52 F.R.D. 253 (S.D.N.Y. 1971).

190 *See also* Daar v. Yellow Cab Co., 67 Cal. 2d 695, 433 P.2d 732, 63 Cal. Rptr. 724 (1967), in which illegal overcharges were returned to the class of cab riders by a general reduction of fares designed to deplete the settlement fund over a number of years. *Daar* is an uncertain authority for fluid recovery outside the settlement context, however, since the court stated that the individuals injured by the overcharges would ultimately have to prove their individual claims. 67 Cal. 2d at 706, 713, 433 P.2d at 740, 745, 63 Cal. Rptr. at 732, 737.

191 Eisen v. Carlisle & Jacquelin, 52 F.R.D. 253, 264-65 (S.D.N.Y. 1971). *See also Cy Pres, supra* note 184, at 458-63. This plan, however, was rejected on appeal as beyond a court's authority under Rule 23. *Eisen III*, 479 F.2d 1005, *reh'g en banc denied*, 479 F.2d 1020 (2d Cir. 1973), *aff'd on other grounds*, 417 U.S. 156 (1974). Critics have assailed this holding as a mechanical and unsympathetic reading of Rule 23 that completely ignores the court's discretion in fashioning relief. *See* Note, *Managing the Large Class Action: Eisen v. Carlisle & Jacquelin*, 87 HARV. L. REV. 426, 451 (1973). *But see* Note, *Eisen v. Carlisle & Jacquelin III: Applying the Axe to the Green Bay Tree*, 35 U. PITT. L. REV. 450, 463 (1973).

192 *See* Note, *Managing the Large Class Action: Eisen v. Carlisle & Jacquelin*, 87 HARV. L. REV. 426, 447 n.119 (1973).

Depending on purchaser responsiveness to the price reduction, the adjusted rate could be attractive enough to divert demand from competing products or suppliers. This would in effect punish the competitors of the defendant by pulling business away from merchants who have acted fairly all along, thus lessening the deterrent effect of the class remedy. Furthermore, if the court attempts to set the price reduction at a point that minimizes competitive effects, the problem of class members dropping out of the relevant geographic or product market with the passage of time is greatly exacerbated.[193]

But to leave the defendant with his ill-gotten gains is even more unpalatable. Therefore a court might require studies of price elasticity for the product in question in order to tailor more precisely its remedy. Or a court might adopt the compromise suggested by one commentator and require a substantial enough price reduction to distribute the damage fund rapidly, with the stipulation that the existence of the lower price not be advertised to the public in order to lessen the effect on competition.[194] The assumption that the damaged class members would be those most likely to continue to consume the product, however, seems open to question.

The possibility that a heavy exaction from a fluid recovery scheme might financially destroy a defendant should be a matter of some concern, and it appears that some courts have decided cases on just this ground.[195] But there is no good reason to make this fear a conclusive rationale for the denial of relief. The antitrust laws are evidence enough that national policy does not re-

---

193 Concern for the low correlation between consumers harmed by the defendant's illegal conduct and the class of persons that benefits from the fluid recovery scheme has led to a conclusion that fluid recovery should not be used unless there is a nearly perfect fit. Jacoby & Cherkasky, *supra* note 184, at 20-24. But the same argument would appear to apply, albeit indirectly, to recoveries in shareholders' derivative suits, in which investors who had owned stock while the wrongful conduct was committed do not receive any benefit from the corporation's recovery, and shareholders who invest after the suit receive a windfall of sorts. The "perfect fit" contention, then, seems to prove too much. *See* 73 COLUM. L. REV. 1641, 1648-49 (1973).

194 *Cy Pres*, *supra* note 184, at 463.

195 *Cf.* Ratner v. Chemical Bank N.Y. Trust Co., 54 F.R.D. 412, 416 (S.D.N.Y. 1972) (class action under Truth in Lending Act; statutory $100 bounty per class member would yield damages of $13 million; class action dismissed).

gard a business's desires as the sole value in the economic system of this country.[196] A similar view applies equally here. If a business is able to survive only by adopting unscrupulous tactics not used by its honest competitors, both those competitors and society suffer. A notion that argues that any conduct necessary to remain in the marketplace is justified invites both economic harm and a general deterioration of a society's moral bases. It follows that the merchant who stays in business only because of his dishonesty does not deserve a place in the economy.

Despite the possible difficulties, the court should feel free to experiment in awarding relief in consumer class actions. Indeed, it need not feel bound by cy pres-like fluid recovery schemes,[197] for it has or can command the tools[198] necessary to shape the remedy to meet the exigencies of each case. The primary consideration must be that difficulties in administration should not be allowed to destroy the usefulness of the class action procedure envisioned in the statute.

G. *Appeals*

The substantive and procedural complexities of the class actions heighten the importance of appellate jurisdiction in the federal

---

196 It is doubtful that anyone has ever expressed — or could ever express — this policy with more force than did Judge Learned Hand in his celebrated *Alcoa* opinion:

> [T]here are some contracts restricting competition which are unlawful, no matter how beneficent they may be; no industrial exigency will justify them; they are absolutely forbidden. . . . We have been speaking only of the economic reasons which forbid monopoly; but, as we have already implied, there are others, based upon the belief that great industrial consolidations are inherently undesirable, regardless of their economic results. . . . [A]mong the purposes of Congress in 1890 was a desire to put an end to great aggregations of capital because of the helplessness of the individual before them. . . . Throughout the history of these statutes it has been constantly assumed that one of their 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.

United States v. Aluminum Co. of America, 148 F.2d 416, 427-29 (2d Cir. 1945).

197 *See generally* Weinstein, *Some Reflections on the "Abusiveness" of Class Actions,* 58 F.R.D. 299, 303-06 (1973).

198 The scope of judicial power already exercised under Rule 23(d) is incorporated in § 8(c) of the model statute, *infra. See* MANUAL FOR COMPLEX AND MULTIDISTRICT LITIGATION, §§ 1.6, 2.1, 5.5. *See generally* 7A WRIGHT & MILLER, *supra* note 36, § 1791, at 192-94.

court system. Generally, appeals from district courts are not permitted from any decision other than a final judgment,[199] but the scope of this rule is best understood through an examination of its exceptions.[200]

One method of appeal before final judgment on the merits is the use of the extraordinary writs.[201] The Supreme Court has flatly stated that they are not to be used "as substitutes for appeals. As extraordinary remedies, they are reserved for really extraordinary cases."[202] But the concept of mandamus and prohibition to decide critical jurisdictional questions, though narrow,[203] has been broad enough to reach a clear "usurpation of power"[204] by a lower court. In *Schlagenhauf v. Holder*,[205] the writ was held properly used in a case of first impression under the Federal Rules of Civil Procedure where an abuse of power by the lower court appeared obvious.[206] Similarly, the writs might be used as a "shakedown cruise" for the procedural rules of a consumer class action statute and as a later remedy for grievous abuses of the judicial power.

Section 1292 of the Judicial Code[207] contains a collection of congressionally created exceptions to the final judgment rule. Some of the interlocutory appeals permitted under this section have a relatively long history of judicial treatment and are fairly

---

199 *See* 28 U.S.C. § 1291 (1970). *See generally* WRIGHT, *supra* note 32, § 101, at 452-53.

200 *See generally* Note, *Appealability in the Federal Courts*, 75 HARV. L. REV. 351 (1961).

201 28 U.S.C. § 1651(a) (1970) provides: "The Supreme Court and all courts established by Act of Congress may issue all writs necessary or appropriate in aid of their respective jurisdictions and agreeable to the usages and principles of law." *See generally* Note, *Supervisory and Advisory Mandamus Under the All Writs Act*, 86 HARV. L. REV. 595 (1973).

202 *Ex parte* Fahey, 332 U.S. 258, 260 (1947).

203 *See, e.g.*, Will v. United States, 389 U.S. 90 (1967) (mandamus will not lie for Government to test a pretrial discovery order in a criminal case). *See generally* HART & WECHSLER, *supra* note 69, at 170-72.

204 DeBeers Consol. Mines, Ltd. v. United States, 325 U.S. 212, 217 (1945). Cases where the right to jury trial has been improperly denied are illustrative. *See, e.g.*, Beacon Theatres, Inc. v. Westover, 359 U.S. 500, 511 (1959).

205 379 U.S. 104 (1964) (mandamus to obtain revision of trial court's medical examination order under Rule 35).

206 *Id.* at 106-12. *See also* LaBuy v. Howes Leather Co., 352 U.S. 249 (1957), in which, as Chief Justice Warren later said, "a district judge displayed a persistent disregard of the Rules of Civil Procedure promulgated by this Court. . . ." Will v. United States, 389 U.S. 90, 96 (1967).

207 28 U.S.C. § 1292 (1970).

precise in scope.[208] But in other cases, appealability of an interlocutory order depends primarily on judicial discretion. In these certified interlocutory appeals, created by Congress in 1958, a district court must state in its order that there is "a controlling question of law as to which there is a substantial ground for difference of opinion and that an immediate appeal from the order may materially advance the ultimate termination of the litigation," and the court of appeals has discretion to refuse an appeal from the order.[209]

These requirements have at times been evaded. In one personal injury case, the district court refused to certify an interlocutory appeal and the court of appeals nevertheless took the case and reached the merits.[210] Moreover, the Supreme Court agreed that the order was appealable, saying that the court of appeals "properly implemented the same policy Congress sought to promote in [providing for certified interlocutory appeals] by treating this obviously marginal case as final and appealable. . . ."[211]

Certified interlocutory appeals are anything but a model of doctrinal clarity,[212] and much of the difficulty in ascertaining whether an appeal will be permitted is due to the Supreme Court's apparent policy of emphasis on the practical considerations involved.[213] This policy of practicality, which seems to have gained increasing

208 28 U.S.C. § 1292(a) (1970). This subsection permits interlocutory appeals from many orders in injunction, reecivership, admiralty, and patent cases. *See generally* WRIGHT, *supra* note 32, § 102, at 458-61.

209 Interlocutory Appeals Act of 1958, Pub. L. No. 85-919; 72 Stat. 1770, *amending* 28 U.S.C. § 1292 (codified at 28 U.S.C. § 1292(b) (1970)).

210 Gillespie v. United States Steel Corp., 321 F.2d 518 (6th Cir. 1963). The court seemed to base its decision on the ground that there was no prejudice to the defendant since its position on the merits was correct. *Id.* at 520-32 *passim*. Professor Wright calls this theory "intriguing," WRIGHT, *supra* note 32, § 101, at 457, but it would seem more accurate to call it dangerous. Arguably, a party that was confident of its position on the merits would not contest the appeal of an interlocutory order, and an appellate court might feel itself bound to accept the appeal by a preliminary consideration of the merits. Such a doctrine seems to concern itself more with the desires of the parties than with the policies behind the final judgment rule, and the Supreme Court quite properly did not base its decision on appealability on the lower court's grounds. Gillespie v. United States Steel Corp., 379 U.S. 148, 152 (1964).

211 *Id.* at 154.

212 *See generally* Note, *Interlocutory Appeals in the Federal Courts Under 28 U.S.C. § 1292(b)*, 88 HARV. L. REV. 607 (1975).

213 *See* WRIGHT, *supra* note 32, § 101, at 456-58.

acceptance in many federal courts,[214] is best seen in the "collateral order" exception to the final judgment rule.

Collateral orders were characterized by the Supreme Court in *Cohen v. Beneficial Industrial Loan Corporation*[215] as orders within "that small class which finally determine claims of right separable from, and collateral to rights asserted in the action, too important to be denied review and too independent of the cause itself to require that appellate consideration be deferred until the whole case is adjudicated."[216] The limits of the doctrine, even after almost thirty years, are not clear,[217] but *Eisen IV* is an indication of its reach. Previous cases had treated an order requiring that the plaintiff post security for costs as appealable,[218] and this rationale was held to apply to Judge Tyler's order that the defendants pay for almost all the cost of notice.[219] This view of collateral orders does not require that a party be put "out of court" as to all or part of his claim. Instead, it seems to state, as did one court of appeals,[220] that

> . . . an order, otherwise unappealable, determining substantial rights of the parties which will be irreparably lost if review is delayed until final judgment may be appealed immediately. . . .

It should not follow, however, that a statute that provides for

---

214 *Id.* at 456 & n.40, 458 & n.52.

215 337 U.S. 541 (1949).

216 *Id.* at 546.

217 *See* WRIGHT, *supra* note 32, § 101, at 455-58.

218 Cohen v. Beneficial Indus. Loan Corp., 337 U.S. 541 (1949) (shareholder's derivative action under New Jersey law, by which security was required). *See also* McClure v. Borne Chemical Co., 292 F.2d 824 (3d Cir.), *cert. denied,* 368 U.S. 939 (1961).

219 According to the Court:

> Like the order in *Cohen,* the District Court's judgment on the allocation of notice costs was "a final disposition of a claimed right, which is not an ingredient of the causes of action and does not require consideration with it," and it was similarly appealable as a "final decision" . . . In our view the Court of Appeals therefore had jurisdiction to review fully the District Court's resolution of the class action notice problems in this case, for that court's allocation of 90% of the notice costs to respondents was but one aspect of its effort to construe the requirements of Rule 23(c)(2) in a way that would permit petitioner's suit to proceed as a class action.

417 U.S. 156, 172 (1974) (citations omitted). The Court did not, however, reach Judge Tyler's decision certifying the case as a class action.

220 United States v. Wood, 295 F.2d 772, 778 (5th Cir. 1961), *cert. denied,* 369 U.S. 850 (1962).

flexible notice orders and trial court allocation of the costs of notice should also provide that these decisions are appealable. The holding in *Eisen IV* dealt with a case fraught with important implications for small claimant class actions, and it is therefore understandable that both the court of appeals and the Supreme Court should have reached the problems presented on a collateral order theory.[221] Yet a statute that called for the minimum amount of notice consistent with due process[222] would require judicial formulation of standards capable of reasonably easy application by lower courts.[223] To freeze the application of these standards into a notion of "rights" and provide for automatic appeals would change the present use of the collateral order doctrine as a means of reaching untested procedural issues affecting the essential nature of the conduct of certain types of suits, and could well burden the appellate courts with insubstantial matters.

The more desirable practice would permit the first few consumer class actions to become test cases in which notice and cost determinations are appealed as certified interlocutory orders.[224] After standards are set by the appellate courts, attacks on flagrantly erroneous trial court orders could be made through the extraordinary writs.[225] Alternatively, appeals could be taken and accepted as collateral orders if the effect of the order in question would be to end a party's conduct of a viable lawsuit.[226] In either case, review would not become the general practice.

---

221 Indeed, the majority of the court of appeals seemed to base its refusal of rehearing en banc on the ground that the Supreme Court was certain to grant certiorari. *See Eisen III*, 479 F.2d at 1020 (Kaufman, J., concurring in denial of rehearing en banc).

222 *Compare* § 8(b)(2) of the model statute, *infra*, which calls for notice calculated to reach "a substantial percentage" of absent class members and sets forth numerous factors in the shaping of proper notice, *with* FED. R. CIV. P. 23(c)(2), which demands "the best notice practicable under the circumstances, including individual notice to all members who can be ascertained through reasonable effort." See also text accompanying notes 108-45 *supra*.

223 This sort of balancing and setting of standards already occurs in administrative law cases where the Supreme Court has set forth standards on "what process is due" in an administrative hearing. *See, e.g.*, Goss v. Lopez, 419 U.S. 565, 577-84 (1975) (public school suspensions).

224 28 U.S.C. § 1292(b) (1970).

225 28 U.S.C. § 1651. See text accompanying notes 201-06 *supra*.

226 As the Supreme Court has said, the concept of finality is to be given a "practical rather than a technical construction." Cohen v. Beneficial Indus. Loan

Even if the judge-made collateral order exception is a matter of peculiar judicial competence, one corollary of that exception appears so clearly grounded in reason and practicality that it should be made a part of any consumer class action statute. The "death knell" doctrine, which holds appealable the denial of class certification, is another product of the *Eisen* litigation. When the case first came to the district court, Judge Tyler dismissed the class action but permitted the representative to pursue his individual claim.[227] It was abundantly clear that Morton Eisen would not continue to press his claim for seventy dollars, and he asked that the order be reviewed as a collateral order. The court of appeals agreed to hear him:[228]

> Dismissal of the class action . . . will irreparably harm Eisen and all others similarly situated. . . . Where the effect of a district court's order, if not reviewed, is the death knell of the action, review should be allowed.

The court's conclusion was clearly right on collateral order grounds alone; all the requirements of *Cohen* were met.[229] Moreover, the case was, for all practical purposes, dismissed in toto. It is literally true, that any order which puts a party "out of court," whether on procedural grounds or on the merits, is a "final decision"[230] and should be appealable as of right. The same sort of practical construction seen in collateral order cases[231] should be applied to the "out of court" requirement for application of the "death knell" doctrine. The Second Circuit's suggested test in *Korn v. Franchard Corp.*[232] — that "a plaintiff simply cannot continue his lawsuit alone"[233] — seems as workable a test as any, although some problems remain.

There is not necessarily any bright-line way to determine whether the plaintiff is "out of court." As the *Korn* court recog-

---

Corp., 337 U.S. 541, 546 (1949). *See generally* HART & WECHSLER, *supra* note 69, at 1553-55.

227 41 F.R.D. 147 (S.D.N.Y. 1966).

228 *Eisen I*, 370 F.2d 119, 172 (Kaufman, J.).

229 See text accompanying notes 214-16 *supra*.

230 28 U.S.C. § 1291 (1970). *But cf.* FED. R. CIV. P. 54(b).

231 See text accompanying notes 213-21 *supra*.

232 443 F.2d 1301 (2d Cir. 1971).

233 *Id.* at 1306.

nized, the complexity of the case is a most helpful indication;[234] the possible recovery to the plaintiff alone is another.[235]

A second problem is whether an order granting class treatment should also be appealable as of right. Since certification of a class increases the pressure for settlement enormously,[236] it is arguable that the *in terrorem* effect of certification needs to be lessened by a right to appeal. But pressure to settle out of court is not of the same nature as being put out of court by judicial decision, even if the effect is substantially similar. If the burden on the defendant is great and the basis of the lower court's decision sufficiently dubious, the certified interlocutory appeal and the collateral order doctrine itself provide adequate remedies.[237]

Finally, there is the appeal from an order dismissing a part of the alleged class and restricting the named plaintiff to representation of a subclass.[238] Clearly, the diminution in the possible fee to the plaintiff's attorney does not present a cognizable ground for appeal as of right. Therefore "death knell" treatment should be reserved for cases where the subclass order is equivalent to total dismissal; any other sufficiently prejudicial orders can be put under certification or collateral order rubrics.[239]

To incorporate the "death knell" doctrine into statutory provision will by no means solve all the appeal problems of class actions. Indeed, the precise scope of the doctrine will probably be a matter of considerable controversy,[240] as is now the case with other appellate jurisdiction problems.[241] But difficulties with the

234 *Id.* at 1307.

235 *Id.* at 1306-07. *See also* Green v. Wolf Corp., 406 F.2d 291, 295 n.6 (2d Cir. 1968), *cert. denied,* 395 U.S. 977 (1969) (securities fraud; representative's claim of less than $1,000; "death knell" held applicable).

236 *See, e.g.,* Herbst v. International Tel. & Tel. Corp., 495 F.2d 1309 (2d Cir. 1974) (complex securities claim brought by representative who held 100 or 200 of 22 million shares certified as class action).

237 *See, e.g., id.* at 1312-13 (certification held appealable).

238 FED. R. CIV. P. 23(c)(4)(B).

239 Section 8(e) of the model statute, *infra,* thus provides for an appeal as of right only from orders denying any class or subclass certification that put a party "out of court." Other statutory and judge-made appeals are preserved and left to judicial expansion or contraction.

240 The Second Circuit's troubles with the doctrine are well summarized in Parkinson v. April Industries, Inc., 520 F.2d 650 (1975).

241 *Compare, e.g.,* Wright, *The Doubtful Omniscience of Appellate Courts,* 41 MINN. L. REV. 751 (1957), *with* Carrington, *The Power of District Judges and the Responsibility of Courts of Appeals,* 3 GA. L. REV. 507 (1969).

margins of the doctrine should not outweigh its essential soundness and require, as some courts have,[242] that the procedural killing of a class action remain hidden from review.

H. *Professional Responsibility and Counsel Fees*

Class actions raise serious and largely unexplored problems of legal ethics. Many of these problems arise because traditional rules governing professional conduct are vague and often not suited to this type of litigation.

A common complaint voiced by opponents of the class device is that such actions open up the opportunity for widespread solicitation of clients.[243] *Halverson v. Convenient Food Mart, Inc.*,[244] is an excellent example of the problems inherent in class litigation. In *Halverson*, an association of franchises hired a lawyer to assert certain rebate rights. The attorney drafted a letter for the association's president in order to solicit more franchises to join both the association and the proposed lawsuit. Reversing the trial court's dismissal of the case, the court of appeals held that the only impropriety involved was the attorney's failure to cite the possible disadvantages of the suit in his letter. Because he had previously been retained by the association, it was proper to inform association members of their opportunity to join a mutually rewarding lawsuit. Moreover, the attorney had the right to solicit non-association members for the suit because a "[l]awyer whose client will benefit from joinder of others similarly situated may seek out claimants if his motive is not to secure fees for himself."[245] Given the facts of the case, this holding seems correct. But it also seems to contradict the Code of Professional Respon-

---

242 *See, e.g.*, Hackett v. General Host Corp., 455 F.2d 618 (3d Cir.), *cert. denied*, 407 U.S. 925 (1972) ("death knell" doctrine rejected). *See also* Parkinson v. April Industries, Inc., 520 F.2d 650, 658-60 (Friendly, J., concurring) (abolition of "death knell" advocated); Samuel v. University of Pittsburgh, 506 F.2d 355 (3d Cir. 1974) (refusal to certify interlocutory appeal from order decertifying class action, rendered with decision on the merits, held not appealable).

243 "A lawyer shall not recommend employment . . . of himself . . . to a non-lawyer who has not sought his advice regarding employment of a lawyer." ABA Code of Professional Responsibility, D.R. 2-103(A) (1969).

244 458 F.2d 927 (7th Cir. 1972).

245 *Id.* at 931 (citations omitted). *But cf.* Carlisle v. LTV Electrosystems, Inc., 54 F.R.D. 237 (N.D. Tex. 1972).

sibility, which allows an attorney to accept — but not to seek — new parties to a class action.[246] A court, of course, is not confined to the metes and bounds of the Code. Subsections (d)(5) and (e) of Rule 23 provide other ways to halt improper solicitation by giving the judge the discretionary power to oversee any settlement offers and to restrain counsel from contacting absent class members about the litigation without the court's approval.[247]

A second major problem concerns the size of attorneys' fees resulting from class action litigation.[248] Some may argue that substantial fee awards have no place when the recovery by individual class members is small. Yet the absence of counsel fees in this context would virtually dictate that such class actions will not be brought at all.[249] From the fee inducement comes a remedy for the injured, correction of the wrongdoer, and, in many cases, a benefit to society at large. Moreover, defendants could easily wear down a class representative whose only compensation, even if he won, would be reimbursement of litigation expenses.[250] The problems of improper conduct in the fee area are as old as the legal profession itself, and any abuses in the consumer context that may arise can be stopped just as well by close judicial supervision of attorneys' fees as by complete denial of the class action remedy.

In any case, the concept of "private attorney general"[251] suits is neither startling nor new. Indeed, the Supreme Court has held that private treble damage actions constitute one of the surest methods for enforcing the antitrust laws.[252] In these cases, the suc-

246 The ABA CODE, *supra* note 243, D.R. 2-104(A)(5), provides:

> If success in asserting rights or defenses of his client in litigation in the nature of a class action is dependent upon the joinder of others a lawyer may accept, but shall not seek, employment from those contacted for the purpose of obtaining their joinder.

See also H. DRINKER, LEGAL ETHICS 63-66 (1953).

247 *See* Kiser v. Miller, 364 F. Supp. 1311 (D.D.C. 1973) (attorneys' solicited contingent fee agreements declared void and fee determined by court). *See also* 7A WRIGHT & MILLER, *supra* note 36, § 1796 at 225 n.22 (1972).

248 Attorneys' fees amounting to hundreds of thousands of dollars have received judicial approval. *See, e.g.*, Lindy Bros. Builders, Inc. v. American Radiator & Standard Sanitary Corp., 341 F. Supp. 1077 (E.D. Pa. 1972); Philadelphia Elec. Co. v. Anaconda Am. Brass Co., 47 F.R.D. 557 (E.D. Pa. 1969).

249 *See* Dolgow v. Anderson, 43 F.R.D. 472, 495 (E.D.N.Y. 1968).

250 *See generally* Schrag, *Bleak House 1968: A Report on Consumer Test Litigation*, 44 N.Y.U.L. REV. 115 (1969).

251 Associated Indus. v. Ickes, 134 F.2d 694, 708 (2d Cir. 1943) (Frank, J.).

252 *See, e.g.*, Zenith Radio Corp. v. Hazeltine Research Inc., 395 U.S. 100 (1969).

cessful plaintiff is entitled to recover more than three times his damages. He also recovers reasonable attorneys' fees and costs of suit, paid directly by the defendant.[253] Similar provisions for counsel fees — sometimes awarded at the court's discretion — exist in the securities[254] and copyright fields[255] and in the Truth in Lending Act.[256] But since *Aleyeska Pipeline Service Co. v. Wilderness Society*,[257] which forbade the award of attorneys' fees directly from defendants on a "private attorney general" theory in cases where there was no statutory authorization, the only way to insure that a court will have the authority to grant fee shifting in a consumer class action is by express provision in the statute.[258]

A court should have a two-fold duty once the suit terminates through settlement or through litigation of liability and damages. First of all, it must determine what is to be a reasonable fee for the class attorneys. It should consider, *inter alia:* the complexity of issues presented by the case; the actual time and effort put into the litigation of the case, as reflected by the attorneys' time sheets; and the proportion of the fees to the fund produced by the litigation. The last consideration is of importance primarily to ensure that the attorneys' fees provide a reasonable incentive. The problem here, simply stated, is that class action litigation is enormously expensive. It is quite possible for attorneys' fees and litigation

---

253 Clayton Act § 4; 15 U.S.G. § 15 (1971).

254 *See, e.g.,* Securities Act of 1933, § 11(e), 15 U.S.C. § 77k(e) (1970) (reasonable attorneys' fees if the court believes the suit or defense to have been without merit). *See also* Mills v. Electric Auto-Lite Co., 396 U.S. 375, 389-97 (1969) (shareholders are entitled to interim award of attorneys' fees, despite absence of "common fund," upon proof of violation of the proxy rules by their corporation).

255 17 U.S.C. § 116 (1971).

256 15 U.S.C. §§ 1640(a), 1681(n), (o) (1971).

257 421 U.S. 240 (1975). *See generally The Supreme Court, 1974 Term,* 89 HARV. L. REV. 47, 170-82 (1975); Dawson, *Lawyers and Involuntary Clients in Public Interest Litigation,* 88 HARV. L. REV. 849, 889 (1975).

258 Charging losers with the winners' counsel fees in cases where there was no statutory grant, said the *Alyeska* Court, violated the common-law "American Rule" that a party is to pay his own costs of litigation:

> Since the approach taken by Congress to this issue has been to carve out specific exceptions to a general rule that federal courts cannot award attorneys' fees beyond the limit of 28 U.S.C. § 1923, those courts are not free to fashion drastic new rules . . . or to pick and choose among plaintiffs and statutes under which they sue and to award fees in some cases but not in others. . . .

421 U.S. at 269.

costs to consume a large portion of any damage award.[259] Yet a statutory grant of a flat percentage of the fund created in all suits would probably cause attorneys to seek out only the richest cases and impede broad enforcement of the statute. It follows that a court must work to prevent both unjust enrichment by attorneys and the frustration of self-enforcement of the statute by making a careful assessment of the percentage of counsel fees from the suit's recovery.

Class members in the smaller or more difficult cases, however, need not be denied all compensation after the court performs its second duty, which is to determine whether a portion of the attorneys' fees should be borne by the defendants. Historically based on the broad discretionary power of the federal courts in equity cases,[260] fee shifting serves a highly important function within a private enforcement model because the defendant pays representative's attorney for the service he renders to society as a whole, thus augmenting the class action's deterrent effect. To be sure, the fee shifting power should be exercised with restraint in light of its quasi-punitive nature.[261] But in situations where the costs of the suit are large and effectively deplete the fund created, or where a defendant's conduct of the suit has been rife with delay or bad faith,[262] the court should have the power to do equity by shifting all or part of the cost.

259 *See, e.g.,* Trans World Airlines, Inc. v. Hughes, 312 F. Supp. 478, 485 (S.D.N.Y. 1970) *aff'd,* 449 F.2d 51 (2d Cir. 1971), *rev'd on other grounds,* 409 U.S. 363 (1972) (attorneys' fees alone of $7.5 million). In a proposed settlement of a class action against Master Charge, where the damages sought exceeded 10 million dollars, the terms provided for a change in the billing format but allowed no compensation to class members. Instead, the total settlement fund was used to pay attorneys' fees and notice costs. Wall Street Journal, Dec. 8, 1971, at 13, col. 1.

260 *See* Sprague v. Taconic National Bank, 307 U.S. 161, 164-67 (1939); *but cf.* Dawson, *supra* note 257 at 889-906.

261 In addition, the possible economic destruction of the defendant resulting from shifting the classes' costs and attorneys' fees might well be considered by the court. *Cf.* Ratner v. Chemical Bank N.Y. Trust Co., 54 F.R.D. 412, 416 (S.D.N.Y. 1972).

262 A number of cases have developed the doctrine of fee shifting as a punishment for vexatious or "bad faith" litigants, following Toledo Scale Co. v. Computing Scale Co., 261 U.S. 399 (1923).

This sanction has been used in situations involving persistent delaying tactics or abuse of procedure. *See, e.g.,* First Nat'l Bank v. Dunham, 471 F.2d 712, 713 (8th Cir. 1973) (fraudulent transfers to prevent execution of a money judgment, attempted bribery of a witness and falsification of evidence); Undersea Eng'r & Constr. Co. v. International Tel. & Tel. Corp., 429 F.2d 543, 545 (9th Cir. 1970)

Conclusion

The costs of consumer fraud are staggering, and the proposed Consumer Protection Agency bills represent a step in the right direction by providing an advocate for a public badly in need of protection. Indeed, the mere presence of such an agency may well serve to educate the public and enable it to make better informed choices in the marketplace. Yet no amount of education alone will serve to compensate those who have been harmed or deter wrongdoers from attempting the same kinds of schemes in the future. Given the national interest in consumer protection reflected in current proposals, it is all the more curious that Congress should go only part of the necessary distance.

A private right of action, though a necessary supplement to the limited resources and powers of the proposed agency, is currently stymied by *Snyder, Eisen,* and their progeny and requires legislative resuscitation. Without such legislation, any consumer victory in the establishment of a public advocate will be a shallow one indeed.

A FEDERAL CONSUMER CLASS ACTION ACT

TABLE OF CONTENTS

Section 1. Short Title
Section 2. Statement of Findings
Section 3. Definitions
Section 4. Prohibited Practices
Section 5. Prerequisites to Actions
Section 6. Actions and Relief; Jurisdiction and Venue
Section 7. Pleading; Misrepresentations and Reliance; Burdens
Section 8. Procedural Rules
Section 9. Damages
Section 10. Attorneys' Fees

("frivolous, if not groundless and vexatious" motion to disqualify the trial judge); Cleveland v. Second Nat'l Bank & Trust Co., 149 F.2d 466, 469 (6th Cir. 1945), *cert. denied,* 326 U.S. 775 (1945), ("persistent, harassing, and futile litigation" of charges of trust mismanagement); *In re* Swartz, 130 F.2d 229, 231 (7th Cir. 1942) ("unnecessary, groundless, vexatious and oppressive petitions [and] motions").

Section 11. Limitation of Actions
Section 12. Cumulative Remedies
Section 13. Saving Clause

Section 1. *Short Title*

This Act may be cited as "The Federal Consumer Class Action Act of 1976."

Section 2. *Statement of Findings*

(a) Congress hereby finds and declares that:

(1) the interests of consumers are inadequately represented and protected within the federal government;

(2) vigorous protection of the interests of consumers is essential to the fair and efficient functioning of a free market economy;

(3) the lack of an effective and efficient remedy in the courts of the United States causes consumers to suffer economic harm;

(4) the protection available under existing federal laws insufficiently prevents unfair and deceptive acts against consumers;

(5) the protection afforded consumers under substantive state law is ineffective in redressing large-scale consumer abuses; and

(6) deceptive and unfair acts and practices presently perpetrated against consumers affect interstate commerce and present a problem of such magnitude that Congress should act to provide an effective remedy.

(b) Congress therefore declares that the lack of an effective remedy impairs the free flow of consumer goods in interstate commerce.

(c) Congress further declares that class actions are an effective and favored means for redress of wrongs suffered by consumers.

COMMENT: This recital of finding has ample support in the current literature,[263] and sets forth a declaration of national policy that should guide the courts in making a liberal construction of the Act.

---

263 See text accompanying notes 1-108 *supra*.

Section 3. *Definitions*

For purposes of this Act:

(a) "Goods" means any tangible chattels;

(b) "Services" means work and labor furnished for personal and commercial uses, including, but not confined to, work and labor furnished in connection with the sale or repair of goods;

(c) "Person" means an individual, partnership, firm, corporation, or other legal entity;

(d) "Consumer" means any person who seeks or acquires, by purchase or lease, any goods or services for personal, family, or business purposes;

(e) "Transaction" means any agreement between a consumer and any other person, and includes the making of, and the performance pursuant to, that agreement.

COMMENT: The language of this section is designed to protect every person, including partnership and corporate entities, who makes a purchase or receives a service of any kind. Purchases and leases of real property, however, are not included.

A class of relatively large businesses could bring an action against a supplier under the Act. Although these entities would probably sue alone in any case, the need for economy in adjudicating common issues argues for the use of the class device in this context as well.[264] In the case of individuals, a more narrowly drafted provision would fail to meet the goals of compensation and deterrence because consumers excluded from the statute would be forced to seek redress through the present unworkable means. Each term should be given as broad an interpretation as practical to insure that no consumer is excluded from the statute's purpose.

Section 4. *Prohibited Practices*

(a) The following acts or practices undertaken by any person in any transaction of sale or lease of goods or services to any consumer are unlawful:

264 See text accompanying note 79 *supra*.

(1) Advertising goods or services with intent not to sell or lease them as advertised;

(2) Advertising goods or services with intent not to supply reasonably expectable demand, unless the advertisement discloses a limitation of quantity;

(3) Representing that a part, replacement, or repair service is needed when it is not;

(4) Representing that the consumer is legally obligated to pay for, safeguard or return unsolicited goods when the consumer is not;

(5) Representing that the consumer will obtain any rights, privileges, or remedies, when the consumer will not;

(6) Representing that goods are original or new when they have deteriorated unreasonably or are altered, reconditioned, reclaimed, used, or secondhand;

(7) Representing that goods or services are of a particular standard, quality, or grade, or that goods are of a particular style or model, when they are not;

(8) Making false statements of fact concerning:

(i) the reason for the existence of, or the amount of, price reductions, or

(ii) savings in comparison to prices of competitors or one's own regular or former price;

(9) Misrepresenting an affiliation, connection, or association with, or certification by, another person or governmental agency;

(10) Failing to return or refund a deposit or advance payment for goods not delivered or services not rendered, when no default or further obligation of the person making such deposit or advance payment exists; *provided,* that nothing in this provision shall affect laws existing in any state covering deposits and refunds on the rental or sale of real property;

(11) Representing that goods or services have origin, characteristics of safety or performance, ingredients, use, benefits, or quantities that they do not have;

(12) Representing that a transaction confers or involves obligations which it does not confer or involve, or which are prohibited by state or federal law;

(13) Misrepresenting the authority of a salesman, representative, or agent to negotiate the final terms of a transaction with a consumer; *provided,* that this subdivision shall not be construed in a manner inconsistent with the law of agency for a partially disclosed or undisclosed principal;

(14) Any affirmation of the value of goods or services, or statement of opinion of the goods or services, which tends to take unfair advantage of the level of knowledge, ability, or experience of the consumer; and

[(15) Any act or practice found to be oppressive or otherwise unconscionable.]

[(b) The Administrator of the Consumer Protection Agency shall have the power to make rules and regulations to specify acts or practices comprehended within the purpose of this Act, and to exempt acts and practices not comprehended within the purpose of this Act.]

COMMENT: The "types" approach used in this section is based on Section 3.201 of the National Consumer Act[265] and on similar provisions in the California Consumers' Legal Remedies Act.[266] The list of proscribed practices, while extensive, is not intended to be exclusive; other federal statutes, such as the antitrust laws, can serve as a basis for a class action.[267] Subsections (a)(15) and (b), since they may be thought to grant an excessive amount of regulatory power contrary to the intent of the drafters of the Consumer Protection Agency proposals[268] and cause the CPA to become a competitor of the Federal Trade Commission, are optional. But the need for administrative authority to further clarify — and make exemptions from — the "types" seems essential if the Act is to effectuate the statutory purpose.[269]

An actual agreement is required to make the prohibitions actionable. As in securities fraud cases, a person who has been induced by misrepresentations to refrain from a purchase will not have a right of action. Undoubtedly, many harms will thus escape the reach of the Act, but the problems of individual proof would probably be so great that a class action would become unmanageable.[270]

The language of two subdivisions, although of needed breadth, should not be used to alter certain common law doctrines which

265 NATIONAL CONSUMER LAW CENTER, BOSTON COLLEGE LAW SCHOOL, NATIONAL CONSUMER ACT (First Final Draft 1970).

266 CAL. CIV. CODE § 1770 (West 1973), *as amended,* (West Supp. 1976).

267 See § 12 *infra.*

268 See text accompanying note 25 *supra.*

269 See text accompanying notes 89-90 *supra.*

270 See text accompanying notes 81-86 *supra.*

have strong policy bases of their own. Therefore subdivision (10) does not reach liquidated damage provisions in real estate sales contracts or security deposits in leases of realty. Subdivision (13) leaves agency for partially disclosed and undisclosed principals to current law.[271]

The standard of knowledge required for misrepresentation is not intended to include simple negligence,[272] but should be construed to include statements made with reckless disregard of whether they were true or false.[273] The defendant's state of mind, as in many other areas of the law, will be a question of fact.[274]

Subsections (7) and (12) will probably be used extensively. The former looks to breach of warranty, express or implied, while the latter will encompass such situations as existed in the *Daar*[275] case.

Section 5. *Prerequisites to Actions*

No person entitled to bring an action under this Act shall be required to initiate or pursue any other judicial or administrative remedy established by local, state, or federal law; *provided,* that failure to exhaust pending state or federal remedies shall be a bar to any action under this Act.

COMMENT: Since this Act provides a remedy in addition to any others that may exist,[276] the only proper bars to any action by a representative are lis pendens by virtue of any action he has already brought, or res judicata through his own or another's prior action.

---

271 *See generally* RESTATEMENT (SECOND) OF AGENCY §§ 4, 144-211 (1957).

272 *See* Ultramares Corp. v. Touche, Niven, & Co., 255 N.Y. 170, 174 N.E. 441 (1931); Ernst & Ernst v. Hochfelder, 96 S. Ct. 1375 (1976), *rev'g* 503 F.2d 1100 (7th Cir. 1974).

273 RESTATEMENT (SECOND) OF TORTS § 310 (1965); Lanza v. Drexel & Co., 479 F.2d 1277, 1306 (2d Cir. 1973); *cf.* RESTATEMENT (SECOND) OF TORTS § 496D, comment *d* (1965); *see generally* Ernst & Ernst v. Hochfelder, 96 S. Ct. 1375, 1381 n.12 (1976).

274 *See* W. PROSSER, HANDBOOK OF THE LAW OF TORTS § 109, at 728 (4th ed. 1971); *cf.* Edgington v. Fitzmaurice, L.R. 29 Ch. D. 359 (1882) (Bowen, L.J.) ("the state of a man's mind is as much a fact as the state of his digestion").

275 *Supra* notes 29, 190.

276 See § 12 *infra.*

Section 6. *Actions and Relief; Jurisdiction and Venue*

(a) Any consumer who suffers an injury as a result of an act or practice declared unlawful by Section 4 of this Act may bring an action as representative for a class of similarly affected consumers against any person engaging in such practice, for any or all of the following:

(1) actual damages for members of the class;
(2) an order enjoining such acts or practices;
(3) court costs; and
(4) any other relief the court may deem proper.

(b) The district courts of the United States shall have original jurisdiction, exclusive of the courts of the states, of all actions under this Act where the aggregate of injury to the class exceeds the sum of $50,000, exclusive of interest and costs, and regardless of the amount of any individual claims.

(c) Any person entitled to sue under this Act may bring an action in any district court of the United States in which any of the defendants is incorporated, is licensed to do business, is doing business, has an agent, or is found.

(d) The district court shall have power, for the convenience of parties and witnesses, or in the interests of justice, to transfer any action to any other district or division.

(e) All actions shall be subject to the provisions of § 1407 of Title 28, United States Code.

(f) The Administrator of the Consumer Protection Agency shall be notified of all actions brought under this Act.

COMMENT: This section contains several variations from current practice, not the least of which is its repudiation of *Snyder v. Harris*.[277] Subsection (c), an amalgam of provisions of § 4 of the Clayton Act[278] and the federal question venue statute,[279] is designed to accord the widest possible service of process and venue. It broadens venue in situations involving multiple defendants, subject only to the limitations of jurisdiction over the person. Under subsection (d), the court retains the discretion to trans-

277 394 U.S. 332 (1969); see text accompanying notes 41-47 *supra*.
278 15 U.S.C. § 15 (1973).
279 28 U.S.C. § 1391(c) (1962).

fer the case,[280] but, consistent with the statute's intent to be broad in procedure as well as in remedy, removes the crippling effect of *Hoffman v Blaski.*[281] Subsection (e) explicitly authorizes action by the Judicial Panel on Multidistrict Litigation; such action could well be necessary in cases where a large-scale harm precipitates multiple suits.[282] The Administrator of the Consumer Protection Agency, under subsection (f), receives notice of his opportunity to intervene or otherwise participate in actions.[283]

Section 7. *Pleading; Misrepresentations and Reliance; Burdens*

(a) In all actions brought under this Act, the plaintiff's complaint shall set forth with particularity the circumstances averred to support the allegation of a violation of Section 4 of this Act.

(b) In any class action pursuant to this Act, the representative shall have maintained his burden regarding misrepresentation as a common question if he produces such evidence as could show, by a preponderance of the evidence, that the person or persons alleged to have engaged in acts or practices declared unlawful under this Act have engaged in a continuous and common scheme or plan against the class or subclass of consumers.

(c) A showing that the class or subclass of consumers leased or purchased the goods or services involved shall, at all stages of the proceedings, be prima facie evidence of reasonable reliance on the misrepresentation.

(d) If the provisions of subsections (b) and (c) of this Section are satisfied by the representative, the burden shall be on the defendant to persuade the trier of fact, by a preponderance of the evidence, that he has not made such misrepresentations or that there has been no reliance thereon.

COMMENT: Consumer class actions typically involve such complex

---

280 *Cf.* 28 U.S.C. § 1404(a) (1962).

281 363 U.S. 335 (1960) (action cannot be transferred to forum where venue would have originally been improper).

282 See text accompanying notes 72-76 *supra.*

283 See text accompanying notes 20, 24 *supra.*

questions of fact and law that the notice pleading provided by Rule 8(a)(2) of the Federal Rules of Civil Procedure seems clearly unworkable. Therefore the allegations should be pleaded with the same specificity as is required under current federal practice for averments of fraud and other "special matters."[284]

Subsections (b) and (c) are to be used both at any preliminary hearing on certification, if ordered, and at trial. At the preliminary hearing the court's task will be to determine whether, on the basis of an outline of evidence drawn from such limited discovery as the court may direct,[285] a subsequent trier of fact could reasonably find that it is more likely than not that a common course of conduct has been employed by the defendant. The standard here should work out in practice to be higher than that on a motion to dismiss,[286] but, since pre-hearing discovery will be necessarily limited, it should be somewhat lower than that on a motion for summary judgment.[287] The burden of production, of course, will be on the plaintiff, but given the prima facie reliance standard in subsection (c), the defendant will most likely have difficulty in showing that no reasonable fact-finder could accept the plaintiff's evidence.

At trial, subsection (d) becomes operative. The burden of production as to misrepresentation and reliance will remain on the plaintiff, but he will not have to demonstrate to the court that his evidence to prove fraud is clear and convincing, as is now the standard in many jurisdictions.[288] Instead, the standard will be whether a reasonable trier of fact could find that the necessary scienter existed. Again, proof of the making of a transaction by members of the class will satisfy the reliance standard. Once the elements are so made out, the risk of non-persuasion as to all the elements, on a "more likely than not" standard, will rest on the defendant.[289]

---

284 FED. R. CIV. P. 9(b).

285 See text accompanying notes 144-49 *supra*.

286 FED. R. CIV. P. 12(b)(6); *cf.* Dioguardi v. Durning, 139 F.2d 774 (2d Cir. 1944).

287 FED. R. CIV. P. 56(c); *cf.* Arnstein v. Porter, 154 F.2d 464 (2d Cir. 1946); Cross v. United States, 336 F.2d 431 (2d Cir. 1964).

288 *See, e.g.,* Yoo Hoo Bottling Co. v. Leibowitz, 432 Pa. 117, 242 A.2d 469 (1968).

289 See text accompanying notes 93-96 *supra*.

Section 8. *Procedural Rules*

(a) *Requirements.* An action brought under this Act shall be maintained as a class action if the court finds that:

(1) the class is so numerous that joinder of all members is impracticable under the Federal Rules of Civil Procedure;

(2) the questions of law or fact common to the members of the class are substantial when compared to questions affecting only individual members;

(3) the claims of the representative party are typical of the claims of the class represented; and

(4) the representative party will fairly and adequately protect the interests of the class.

COMMENT: The language of this subsection is similar to that of present Rules 23(a) and 23(b)(3). There are, however, certain important changes, reflecting the class action's status as a favored remedy. Subsection (2) requires only that common questions be "substantial," not "predominant." Although the provisions of Section 7 on substantive common questions and the authorization given in subsection (c)(1) of this Section for common trials on damages should aid in making common issues predominant, the desirability of res judicata on as many questions as possible makes a substantiality requirement more useful for all parties. Given the class action's favored status in the Act, the present requirement of Rule 23(b)(3) that the class action be "superior to other available methods for the fair and efficient adjudication of the controversy" is surplusage at best and an impediment at worst. Subsection (3) should be given as broad a reading as is consistent with the doctrine of standing to ensure that the existence of only slightly diverging interests does not cause the class action to be aborted; the primary emphasis should be on subsection (4)'s requirement of adequate representation.[290]

(b) *Certification and Notice*

(1) As soon as practicable after commencement of an action under this Act, the court shall determine by order whether it is to be maintained as a class action. An order under this subsection,

290 See text accompanying notes 119-29 *supra*.

which may be conditional and may be amended or altered at any time before decision on the merits, may be made on the pleadings, or after notice to absent members, limited discovery, preliminary hearing on the merits, or by any other means the court shall deem necessary and expedient. The court shall have power to restrict the class action to certain issues, and to divide the class into subclasses and treat each subclass as a class.

(2) In making its order under subsection (1) of this subsection, the court may provide that the representative, the defendant, or both give notice to the absent members of the class by any means found capable of reaching a significant percentage of the absent members of the class. In determining the means to be employed, the court shall consider the following factors, among others:

(A) whether the nature of the class interests asserted raises a significant probability that conflicting interests will emerge during the course of the litigation;

(B) the apparent adequacy of representation of the class;

(C) the probability of any single absent member's intervening in or withdrawing from the litigation;

(D) the size of the class or subclass asserted;

(E) the extent to which the identities of class members are readily ascertainable;

(F) the probable recovery per member;

(G) whether, absent the class action, the class would be without an adequate means of redressing its grievances; and

(H) the extent to which the costs of any suggested notification plan would significantly deplete the possible recovery or place an unreasonable burden on the party or parties bearing the cost of notice.

(3) Any notice ordered under this subsection shall advise the member that:

(A) an action has begun in which he is an asserted class member;

(B) the court will exclude him from the class if he so requests by a specified date;

(C) the judgment, whether favorable or not, will include all members who do not request exclusion; and

(D) any member who does not request exclusion may, if he desires, enter an appearance through his counsel.

(4) The costs of such notice as the court shall direct under this subsection shall be borne by the representative, the defendant, or

both, at the court's discretion. The court may hold such preliminary hearing as it deems necessary to determine, in view of the factors set forth in subsection (2) of this subsection and the probability that the plaintiff will prevail on the merits, the extent to which each party shall be responsible for the cost of notice.

COMMENT: This subsection recognizes that certification, notice, and manageability concerns are inextricably joined, and attempts to free class actions from the restrictions of present Rule 23 and *Eisen IV*.[291] Although subsection (1) follows present practice under Rules 23(c)(1) and 23(c)(4) in permitting the class order to be provisional and authorizing the court to divide the class or restrict it to certain issues, it significantly alters present practice by giving the court maximum flexibility in the manner in which it determines its certification order. The essential purpose of this subsection and subsection (2) is to ensure an informed certification order that ensures adequate representation.[292]

Subsection (2) explicitly recognizes that the due process concerns of notice are best served through a balancing of various considerations. Although an action for injunctive relief may require no more notice than is currently necessary in present practice under Rules 23(b)(1) and (b)(2),[293] notice under what would presently be a 23(b)(3) action need not necessarily include individual notice to all ascertainable class members. Instead, the court should attempt to tailor its notice to discover conflicting interests and determine the adequacy of representation to ensure the practicable due process envisioned by *Mullane*.[294] The form of notice provided by subsection (3) is substantially similar to that required by Rule 23(c)(2)(A)-(C).

Subsection (4) explicitly authorizes the holding of a preliminary hearing by the court to allocate notice costs, as was done by Judge Tyler in *Eisen*.[295] The court should make a preliminary inquiry, after such discovery as it deems necessary, into the practical con-

291 See text accompanying notes 56-62 *supra*.
292 See text accompanying notes 125-29 *supra*.
293 *Compare* FED. R. CIV. P. 23(d)(2) *with* FED. R. CIV. P. 23(c)(2).
294 See text accompanying notes 109-18 *supra*.
295 *See* Eisen v. Carlisle & Jacquelin, 52 F.R.D. 253, 271-72 (S.D.N.Y. 1971); *id.*, 54 F.R.D. 565 (S.D.N.Y. 1972). See also text accompanying notes 145-49 *supra*.

siderations set forth in subsection (2) and the probability that the plaintiff will ultimately prevail on the merits. The court, however, need not be bound by a party's request for a hearing, and it may require that a party furnish security similar to an injunction bond.

(c) *Additional Orders in Conduct of Actions.* In the conduct of actions under this Act, the court may issue further orders:

(1) establishing a plan by which common issues, including, if feasible, damages, shall be tried together and individual questions determined in subsequent proceedings;

(2) prescribing measures to prevent undue repetition or complication in the presentation of evidence or argument;

(3) imposing conditions on representative parties or intervenors;

(4) requiring, for the protection of the members of the class or otherwise for the fair conduct of the action, that notice be given in such manner as the court may direct to some or all of the members of the class, informing them:

(i) of the opportunity of members to signify whether they consider the representation fair and adequate, and to intervene or otherwise participate in the action; or

(ii) of the nature and extent of the judgment or settlement and the means by which they may present their individual claims;

(5) dealing with other procedural matters.

Such orders may be combined with a pretrial order under Rule 16 of the Federal Rules of Civil Procedure, and may be altered or amended as may be desirable from time to time.

COMMENT: This subsection largely incorporates the provisions of Rule 23(d), but also makes explicit the court's power to inquire into the feasibility of a trial on class-wide damages[296] and to establish the notice procedure for the filing of individual claims.[297]

(d) *Settlement.* A class action under this Act shall not be dismissed or compromised without the approval of the court. Notice of the proposed dismissal or compromise shall be given to all members of the class in such manner as the court directs, and the court shall have

296 See text accompanying notes 172-82 *supra*.
297 See text accompanying notes 183-84 *supra*.

power to order a hearing on the propriety of the proposed dismissal or compromise.

(e) *Appeals.* Any judgment or order rendered in any action under this Act shall be subject to review as provided in §§ 1254, 1291, 1292, 1294, and 1651 of Title 28, United States Code; *provided,* that an order denying certification of an asserted class or subclass, which terminates the representative's pursuance of relief under this Act, may be appealed as of right.

COMMENT: Subsection (d) adds to present Rule 23(e) by expressly granting power to the court to hold a hearing on the propriety of any proposed settlement.[298]

The only change from present appellate practice under subsection (e) is that any order of the court which effectively puts a party "out of court" is appealable as of right. The precise scope of this "death knell" doctrine, however, will depend on appellate court decisions. All other current statutory and judge-made exceptions to the final judgment rule are preserved.[299]

Section 9. *Damages*

(a) Any defendant who makes a written tender of settlement which is rejected by the representative may file the tender and an affidavit concerning its rejection with the court together with a petition for a remittitur. The court may grant such remittitur if it finds that the relief tendered was reasonable in relation to the injury actually suffered by the class and that the jury's award of damages was excessive.

(b) The trier of fact shall have the power to grant to the class additional damages in an amount up to the actually incurred damages awarded if it finds by clear and convincing evidence that the defendant acted with knowledge that the act or practice complained of violated Section 4 of this Act.

COMMENT: Subsection (a) is an express grant of remittitur power to the trial court which permits it to order the plaintiff to accept either an amount previously offered as a settlement or a new trial. The power to order remittitur has long been recognized,[300] but

298 See text accompanying notes 151-63 *supra.*

299 See text accompanying notes 191-242 *supra.*

300 Blunt v. Little, 3 F. Cas. 760 (No. 1578) (C.C.D. Mass. 1822) (Story, J.); New York, L.E. & W.R.R. v. Estill, 147 U.S. 591 (1893). The Supreme Court has held, however, in a sharply disputed opinion, that additur, by which a defendant is compelled to choose between an increased damage award and a new trial,

it is highly doubtful that a court could constitutionally substitute its own damage award for that of the jury as a final judgment.[301]

The subsection does not, however, lock the court into a single remittitur figure. Although a court should not follow the discredited practice of directing a remittitur at the lowest figure that a jury could have awarded,[302] it should feel free to adopt what it finds to be a reasonable figure.[303] Indeed, a court could well conclude that the jury was inclined to award the maximum permissible damages and set the remittitur at the highest figure it finds proper.[304]

Just as remittitur encourages the plaintiff to avoid reaching for unreasonable damages and to accept a reasonable settlement offer, subsection (b)'s provision for punitive damages should encourage defendants to avoid vexatious litigation. The grant of these damages, however, is not automatic;[305] the trier of fact must find, under a high evidentiary standard, that the defendant knowingly violated the substantive provisions of the Act.[306] "Runaway" juries are thus subject to considerable control by the court.

Section 10. *Attorneys' Fees*

(a) If the representative prevails on the merits, or a compromise or settlement is effected, the court shall determine a reasonable attorneys' fee based on the following considerations, among others:

(1) the complexity of the issues presented;

---

violates the seventh amendment's right to jury trial. Dimick v. Schiedt, 293 U.S. 474 (1935). *But see id.* at 488 (Stone, J. dissenting); Fisch v. Manger, 24 N.J. 66, 130 A.2d 815 (1957).

301 As the Court stated in Kennon v. Gilmer, 131 U.S. 22, 29 (1889): "No court . . . is authorized, according to its own estimate of the amount of damages which the plaintiff ought to have recovered, to enter an absolute judgment for any other sum than that addressed by the jury." *See also* Brewer v. Uniroyal, Inc., 498 F.2d 973, 976 (6th Cir. 1974).

302 *See* Powers v. Allstate Ins. Co., 10 Wis. 2d 78, 102 N.W.2d 393 (1960), *overruling* Heimlich v. Tabor, 123 Wis. 565, 102 N.W. 10 (1905).

303 Meehan v. Central R.R., 181 F. Supp. 594, 608 (S.D.N.Y. 1960). *See generally* C. McCormick, Handbook of the Law of Damages, § 19, at 82 (1935).

304 *Cf.* Gorsalitz v. Olin Mathieson Chem. Corp., 429 F.2d 1033, 1047 (5th Cir. 1970); Glazer v. Glazer, 278 F. Supp. 476, 481-82 (E.D. La. 1968).

305 The statute thus differs from the antitrust laws. Clayton Act § 4, 15 U.S.C. § 15 (1973).

306 *See generally* Walther & Plein, *Punitive Damages: A Critical Analysis: Kirk v. Combs,* 49 Marq. L. Rev. 369 (1965) (discussing Wisconsin cases in large part).

(2) the amount of the representative's counsel's time and effort involved;

(3) the relative size of the fund produced;

(4) the conduct of counsel during the course of the litigation; and

(5) the amount necessary to provide an incentive for the further enforcement of the Act.

(b) The court may, in its discretion, hold a hearing on such considerations, and may require that all or part of such fees be paid directly by the defendant.

COMMENT: This section, which mandates a determination of counsel fees by the court, grants considerable discretion in the setting of the fees. Wisely applied, it is a central means of providing incentive for meritorious claims and deterring strike suits. A Congressional grant of this power is now essential under current decisional law.[307]

If the court finds that properly allowable fees, given the fund created, would deprive the class of significant compensation or fail to provide an adequate incentive for future enforcement, it may properly require the defendant to pay counsel fees in addition to the fund created by the litigation. It may also do so if the defendant's conduct during the suit has been vexatious or in bad faith.[308]

Similarly, the court may limit the contingent fee to be taken by the representative's attorney from recovery or, if it finds that counsel has improperly conducted the litigation (*e.g.*, engaged in harassment or refused an adequate settlement offer), may properly find that counsel is entitled to no fee whatever. This power should thus create a strong incentive for both parties to act in good faith. In any case, the court is required by this section to consider all the circumstances of the litigation in setting its fee award.[309] Ordinarily, a written opinion on the matter will be desirable, given the importance of the fee award to all the parties and to the bar and general public. The court may even wish to hold a hear-

307 See text accompanying notes 257-58 *supra*.

308 See note 262 *supra*.

309 See text accompanying note 259 *supra*.

ing on the matter. Any abuse of discretion by the court can be dealt with by appellate review.[310]

Section 11. *Limitation of Actions*

(a) Any action under this Act shall be brought within one year of the time the act or practice complained of by the representative became known or should have become known by him.

(b) Only those persons who suffered the alleged harm within a period of five years prior to the time of the alleged injury to the class representative may be included in any alleged class.

COMMENT: This statute of limitations is an effort to effect a workable balance between the problems of stale claims and concealed fraud. It provides for a maximum period of six years from the time of the injury to a class member to the time of the suit. Thus if the alleged unlawful practice should have been discovered by all members immediately, an action must be brought within one year of that time; but if the act is hidden, the injured consumer has five years until his claim cannot be cognizable. Should the injury become known within five years of its commission a representative has one year from the date of actual or constructive discovery in which to bring his claim. The commencement of the action, as in current practice, tolls the statute for all members of the class;[311] and should the representative fail to qualify, any member whose actual or constructive discovery of the injury occurred within one year of the bringing of the first suit may begin another action. If a representative qualifies and prevails on the merits, however, it is irrelevant that any member of the class has actually or constructively discovered the injury to himself more than one year prior to the commencement of the successful suit. The only restriction on qualification for recovery is that a member must have been injured within five years of the time of the injury to the representative.

If the injury remains undetected for more than five years after

---

310 *Cf.* § 8(e) *supra* and comments thereto.

311 American Pipe & Constr. Co. v. Utah, 414 U.S. 538 (1974).

its commission, the policies against aged, unreliable evidence and manufactured claims prevail over the policy of redress. Whether an alleged unlawful act was known or should have become known at or before the time alleged will often be a question for the trier of fact.[312]

Section 12. *Cumulative Remedies*

The remedies provided in this Act shall be in addition to any other remedies provided in any other state or federal laws.

Section 13. *Saving Clause*

The provisions of this Act shall be severable, and if the application of any provision of this Act is held to be unconstitutional, the remaining provisions of this Act shall be given full force and effect, as completely as if the part or parts held unconstitutional had not been included herein.

COMMENT: Particularly in its provisions on notice and fluid recovery, this Act goes to the verge of existing constitutional law, and it is unclear whether the Supreme Court might adopt the due process arguments of Judge Medina in *Eisen III*.[313] Nevertheless, the other provisions of the Act are of such potential usefulness that even should part of the approach be held invalid, the rest should remain available.

312 Bertha Bldg. Corp. v. National Theatres Corp., 248 F.2d 833, 835-36 (2d Cir. 1957), *cert. denied* 356 U.S. 936 (1958); Chittenholm v. Griffin, 357 Pa. 616, 620, 55 A.2d 324, 326 (1947).

313 *See* 479 F.2d at 1017 & n.21, 1018. *But see id.* at 1024 (Oakes, J., dissenting from denial of rehearing en banc). See text accompanying notes 58-62, 109-18, 175, 177-82 *supra*.

NOTE

THE TAX TREATMENT OF ENTERTAINMENT EXPENSES UNDER § 274: RESULTS REVIEWED & REVISIONS RECOMMENDED

STEVEN KNOWLES*

Section 274 of the Internal Revenue Code limits the deductions for travel and entertainment expenses allowed under §§ 162 and 212 for ordinary and necessary expenses incurred in the pursuit of income. As presently written and administered, this statutory scheme has produced arbitrary and unfair results. Mr. Knowles criticizes this approach and compares it with the approaches taken by Great Britain and West Germany. He proposes a sweeping reform of the tax treatment of entertainment and certain travel expenses in order to bring the treatment of these expenses in line with modern conceptions of net income.

Introduction

"Now instead of telling us lies, taxpayers are just putting their lies in writing."[1] With that caustic remark, an official of the Internal Revenue Service describes the reaction of many agents to § 274 of the Internal Revenue Code.[2] Enacted in 1962, this section was designed to curb abuses of "expense account living"[3]

*Member of the Class of 1976 at Harvard Law School. The author wishes to thank Greyson Bryan, Bruce Judelson, Jerome Kurtz, and Tom Mooney for their contributions to this Note.

1 Interview with an attorney with the Internal Revenue Service (who wished to remain anonymous), in Boston, March 21, 1976.

2 INT. REV. CODE OF 1954, § 274. This section was enacted as part of the Revenue Act of 1962, 26 U.S.C. § 274 (1962), generally designed to stimulate a lagging economy while closing some loopholes.

3 The notion of "expense account living" seems to have been popularized and criticized by the press beginning in the mid-1950's. *See, e.g., Expense Accounts: A $5 Billion Tax Deduction, and Growing,* U.S. NEWS AND WORLD REPORT, Aug. 16, 1957, at 83; Barcella, *New Accounting of Expense Accounts,* N.Y. TIMES MAGAZINE, Apr. 20, 1958, at 48; *I'll Just Sign, Those Big-Time Expense Accounts,* NEWSWEEK, May 20, 1957, at 87; *The Expense-Account Society,* TIME, Aug. 29, 1960, at 58. President Kennedy's proposals transmitted to the 87th Congress for limitations on travel and entertainment deductions ultimately were phrased as an attack on abuses in the use of "the expense account."

by limiting the deductibility of certain travel and entertainment expenses.[4] Today, many doubt whether the section is accomplishing its purpose.[5] This Note analyzes whether such doubts are justified.

A comparative approach is used. Cases dealing with travel and entertainment[6] deductions before the enactment and enforcement of § 274 are compared with the cases dealing with these deductions under the new rules to determine where the section has been effective and where it has failed.[7] A comparative examination of the tax treatment of travel and entertainment expenses by other major Western industrial nations — Great Britain and Germany — is presented in an effort to learn from the experience of others. Finally, conclusions about the present law and recommendations for change are presented.

I. History of Section 274

A. *The Problem*

The great expense account charade really began as a backlash to the excess profits taxes of World War II.[8] Wartime businessmen, anxious to avoid paying taxes on large profits, fathered a large family of new facilities which grew and prospered because of the travel and entertainment indulgences of the tax code of the

---

4 Section 274 is restrictive in nature, in that the only purpose of the section is to disallow deductions in certain cases which might otherwise be allowed under § 162. *See* H.R. Rep. No. 1447, 87th Cong., 2d Sess. 19-20 (1962) [hereinafter cited as H.R. Rep.].

5 *See, e.g.,* Halperin, *Business Deduction for Personal Living Expenses: A Uniform Approach to an Unsolved Problem,* 122 U. Pa. L. Rev. 859, 929-31 (1974) [hereinafter cited as Halperin]; R. Goode, The Individual Income Tax 97-98 (1964).

6 The emphasis of this article is on "entertainment," including activities and facilities, because the limitations in travel expenses in the original law were revised to apply to only *foreign* travel by the Revenue Act of 1964. See text accompanying notes 32-46 *infra.*

7 The number of cases dealing with § 274 deductions necessarily is limited, particularly on the appellate level. The Supreme Court has yet to decide any § 274 case (although certiorari has been denied in a few). However, the body of case law is by this time sufficiently large to provide a basis for arriving at definite conclusions as to the "before and after" question.

8 *The Age of the Expense Account,* 45 Mgmt. Rev. 283 (1956); Gehman, *Expense Accounts,* Cosmopolitan, March 1957, at 44-47.

era. The post-war years witnessed the growth of this "T & E" industry — new hotels were built and old ones expanded, restaurants and night clubs boomed, sports for the masses caught on as new country clubs and recreational resorts sprang up. By the time a new generation had moved up through the business ranks, the business life-style was established.[9]

By the mid-1950's, expense account spending was estimated variously at between five billion and ten billion dollars a year[10] — resulting in a tax loss of one to two billion dollars.[11] But then the press turned sour. The whole expense account scheme turned from an "elbow dig"[12] to an "uneasy national joke"[13] all the way to an outright "scandal."[14] The Internal Revenue Service warned taxpayers that there would be more rigorous auditing of claimed deductions for travel and entertainment.[15] When the Kennedy administration took over in Washington, action seemed inevitable. It was not long in coming.

B. *The Proposal*

In his first tax message to Congress, President Kennedy declared war on the expense account. He stated:[16]

> This is a matter of national concern, affecting not only our public revenues, our sense of fairness, and our respect for the

9 *See Expense Accounts: A $5 Billion Tax Deduction, and Growing, supra* note 3, at 86-87; THE INVESTOR, Oct. 1957, at 11 (advising investment in Diners' Club because "living on [the] expense account is one of the ways smart people have mitigated taxes").

10 *Expense Accounts: A $5 Billion Tax Deduction, and Growing, supra* note 3, at 83.

11 Rothschild and Sobernheim, *Expense Accounts for Executives*, 67 YALE L.J. 1363 (1958).

12 *I'll Just Sign, Those Big-Figure Expense Accounts, supra* note 3, at 90.

13 Lynes, *Visit to the World of Expense Accounts*, N.Y. TIMES MAGAZINE, Feb. 24, 1975, at 17.

14 *Expense Account Scandal*, U.S. NEWS AND WORLD REPORT, Jan. 25, 1960, at 50.

15 4A MERTENS, THE LAW OF FEDERAL INCOME TAXATION § 25.90, at 326 (Rev. ed. 1966). *See* Rev. Rul. 54-195, 1954-1 CUM. BULL. 47 (requiring fuller documentation for T & E deductions, including use of "secondary sources" in lieu of direct evidence); 1957 INT. REV. BULL. No. 48, at 83 (informing individuals who incur expenses in connection with their employment that for 1958 onward they must keep personal records of expense account information).

16 *Hearings on Tax Recommendations of the President Contained in His Message Transmitted to Congress, April 20, 1961, Before the House Comm. on Ways and Means*, 87th Cong., 1st Sess. 12-13 (1961) [hereinafter cited as *Hearings*].

tax system but our moral and business practices as well. This widespread distortion of our business and social structure is largely a creature of the tax system, and the time has come when our tax laws should cease their encouragement of luxury spending as a charge on the Federal Treasury. The slogan — "It's deductible" — should pass from our scene.

Secretary of the Treasury Dillon emphasized that disallowance of certain deductions was justified because "they confer substantial personal benefits which are in large measure a substitute for personal living expenses."[17]

The recognition that entertainment and most travel represent a mixture of personal and business activities is crucial to an understanding of why these deductions have been singled out in most countries for special treatment. For tax purposes, the individual has two personalities: one is a seeker of profit who can deduct expenses incurred in the search; the other is a person satisfying his own needs through consumption.[18] Most travel and entertainment is undertaken for both reasons, though one or the other motivation may predominate on any single occasion.[19]

Business expenses are allowed deduction under § 162,[20] while personal expenses are disallowed under § 262.[21] In a perfect world, T & E expenses would be allocated accordingly. But in a tax system, such line-drawing is virtually impossible.[22]

The President, with his original proposals, confronted the mixed business-personal expense problem directly. He recognized that "tax expenditures,"[23] or indirect subsidies, should not be

---

17 *Id.* at 43. *See also* Note, *The Additional Expense Test: A Proposal to Help Solve the Dilemma of Mixed Business and Personal Expenses*, 1974 DUKE L.J. 636 (1974).

18 S. SURREY, W. WARREN, P. MCDANIEL & H. ANET, FEDERAL INCOME TAXATION, CASES AND MATERIALS 496 (1972) [hereinafter cited as SURREY].

19 *See* Halperin, *supra* note 5, at 866-67.

20 INT. REV. CODE OF 1954, § 162(a).

21 INT. REV. CODE OF 1954, § 212.

22 The President recognized this fact in his proposals. Deduction for all entertainment facilities would have been eliminated. *Hearings*, *supra* note 16, at 43-44. Business gift deductions would have been limited to ten dollars per person per year. *Id.* at 44. The President's proposals also called for eliminating deductions for any personal travel expenses by requiring allocation between business and personal purposes, and establishing an allowance for the cost of meals and lodging of the traveler to twice the per diem rate authorized to be paid by the federal government. *Id.* at 44. *See generally* Halperin, *supra* note 5, at 928-29.

23 The term "tax expenditure" as used in this Note refers to the indirect gov-

made for the amusement, recreation, or similar pleasurable diversion of any individual. He also appreciated the administrative burden of attempting to allocate between business and personal elements in entertainment. The sweeping nature of the proposals suggests that a decision was made not to permit deductions at all for such mixed expenditures.[24]

The difficulty of separating business from personal elements in travel and entertainment is obvious. While one taxpayer may enjoy heartily an elegant dinner with a client, another may consider the occasion a painful chore. Thus, any rule will produce arbitrary and unfair effects. Moreover, anything less than a per se rule will end in an evidentiary miasma — exactly where the courts often seem to find themselves today.

C. *The Conflict*

A stormy debate followed.[25] Neither the House nor the Senate agreed with the President's idea of complete disallowance of entertainment deductions.[26] The Senate Finance Committee Report expressed the fear that complete disallowance would both discourage business transactions and result in unemployment in the entertainment industry, thus reducing total tax revenues.[27] The Congress believed that abuses could be curbed by eliminating the so-called *Cohan* rule,[28] which permitted the taxpayer to receive

---

ernment subsidization of an activity through special tax provisions — exclusions, credits, exemptions, deferrals, preferential rates, and deductions. *See* SURREY, *supra* note 18, at 240-57. *See generally* Surrey, *Tax Incentives as a Device for Implementing Government Policy: A Comparison With Direct Government Expenditures*, 83 HARV. L. REV. 705 (1970).

24 The generally accepted notion that our concept of net income includes mixed business and personal expenditure would have been changed. Certain common business practices, such as entertainment or mixed business trips and vacations, would no longer be considered a § 162 "cost of doing business." *Cf.* Huffater, *New T & E Regulations Raise Basic Questions Concerning Gross Income*, 18 J. TAXATION 90 (1963).

25 For a discussion of the complicated legislative history of section 274, *see* Lipoff, *Entertainment and Related Expenses under Legislative Attack*, 17 TAX L. REV. 183 (1962).

26 *Cf.* H. CONF. REP. No. 2508, 87th Cong., 2d Sess. (1962), *reprinted in* 1962 U.S. CODE CONG. & AD. NEWS 3732, 3735-36 [hereinafter cited as H. CONF. REP.].

27 S. REP. No. 1881, 87th Cong., 2d Sess. 24-25 (1962) [hereinafter cited as S. REP.].

28 The rule was announced in Cohan v. Comm'r, 39 F.2d 540 (2d Cir. 1930).

a deduction based on the court's approximation of the taxpayer's expenses.[29] The President's original proposal said nothing about substantiation, perhaps in the recognition that such procedural rules would be difficult to enforce in a full and fair way.[30]

The Senate Report also indicates a more fundamental disagreement between the President and Congress — the former believing that "tax expenditures" should not subsidize activity which is essentially personal in nature, the latter upholding the orthodox view that mixed personal and business expenses should be deductible if made in pursuit of income. Congressional confrontation of these conflicts led to the enactment of § 274.

II. The Rules of Section 274

A. *Introduction*

Section 274 provides for limitations on, and disallowances of, business gift, entertainment, and certain travel deductions. It also provides stricter substantiation requirements for such deductions. Section 274(b), establishing a maximum deduction of twenty-five dollars per recipient per year for business gifts, has been of minimal significance and therefore will not be discussed.[31] Simi-

---

There, the Second Circuit allowed George M. Cohan a deduction for his freehanded entertaining of actors, assistants, and critics based on estimates. The court concluded that "it is not fatal that the result will inevitably be speculative; many important decisions must be such." *Id.* at 544.

29 *See* H.R. REP., *supra* note 4, at 23; S. REP., *supra* note 27, at 34-35. The abolition of this procedural rule and the imposition of substantiation requirements under § 274(d) became a leading feature of the new law, and undoubtedly has had more effect than the other rules. *See generally* Comment, *Tax Law: Substantiation of Business Related Entertainment Expenses,* 54 MARQ. L. REV. 346 (1971); BNA Tax Management Portfolio No. 180, *Travel and Entertainment Expenses — Recording and Reporting* (1973).

30 See text accompanying notes 100-45 *infra.*

31 No reported case has turned on the application of § 274(b). Indeed, very few cases under the travel and entertainment heading have even involved claimed business gift deductions. *See* BNA Tax Management Portfolio No. 26, *Travel and Entertainment Deductions B-79* (1967 with annual supplements).

No doubt, the "bright line" rule under the section has contributed to the lack of litigation in this area, but the tax problem of business gift-giving appears to have been on the wane since the mid-1950s, largely in response to the realization among those engaged in the practice that it was an ineffective and often counterproductive method of promotion. *See* BUS. REC., Nov. 1959, at 503 (reporting that larly, the limitation of travel expenses under § 274(c) has had little effect because of its narrow reach.[32] Travel expenses thus will be discussed only briefly before the more thorough analysis of the entertainment deduction problem.

B. *Travel*

President Kennedy believed that too often personal and luxury expenditures were being charged to the federal purse.[33] He therefore proposed to Congress that all transportation expenses be subject to allocation based on the amount of business versus personal time spent on the trip,[34] and that a maximum limitation on deductions for meals and lodging while traveling be established.[35] These proposals seemed reasonable and workable.[36]

Congress finally passed a watered-down version of the Kennedy proposal.[37] The 1962 version of § 274(c) required taxpayers to

---

291 large manufacturing firms had eased the practice of giving business gifts); *Executives Express Misgivings on Christmas Giving*, CONF. BOARD BUS. REC., Dec. 1957, at 560. In any event, a complete analysis of the problem of business gifts is beyond the scope of this Note. *See generally* Reid, *Effects of Business Gifts on Payor and Payee*, 12 TUL. TAX. INST. 624 (1963); Rev. Rul. 63-144, 1963-2 CUM. BULL. 129, 142-44 (questions 87-93).

32 This provision was amended in 1964 to restrict its application to foreign travel. Revenue Act of 1964, § 217, 78 Stat. 56. See text accompanying notes 41-43 *infra*.

33 *Hearings*, *supra* note 16, at 12-13.

34 *Id.* at 40 (Statement of Secretary Dillon).

35 *Id.* It was recommended that deductions be allowed for meals and lodging while traveling only to the extent of two hundred percent of the maximum per diem rates for federal government employees. The current maximum per diem rate for the continental United States is $35.00. 3 CCH 1976 STAND. FED. TAX REP. ¶ 2296.625.

36 The concept of uniform application of maximum limitations for all taxpayers was criticized harshly as based on the erroneous conception of uniformity among taxpayers and their needs. Differences in size, nature of the business, and the level of the executive were emphasized. *See Hearings*, *supra* note 20, at 1610 (Statement of Thomas W. Power) and 1698 (Statement of Leslie Mills). *See also* Emmanuel and Lipoff, *Travel and Entertainment: The New World of Section 274*, 18 TAX L. REV. 487, 515 (1963) [hereinafter cited as Emmanuel and Lipoff]. This criticism fails to recognize two points. First, the allowances were set at a very high rate — recognizing the differences in taxpayers as argued. Second, any deductions for food and any but a reasonable deduction for lodgings are a concession to what in theory should be disallowed anyway. While the cost of transportation may be attributable solely to business, the cost of living at the destination rarely is more than what a taxpayer would be paying for personal consumption anyway. *See* Halperin, *supra* note 5, at 921-28.

37 Revenue Act of 1962 § 4, Pub. L. No. 87-834. *See* H. CONF. REP., *supra* note 26, at 3737. The legislative history of the present § 274(c) spans over two years and

allocate traveling expenses, including meals and lodging, between the portion of the trip which is for business purposes and that portion which is for pleasure.[38] Exceptions were provided where the period of travel did not exceed a week or where less than twenty-five percent of the trip was personal.[39] The stated purpose of the allocation requirement was to eliminate sole reliance on the "primary purpose" test, which permitted deductions of the *entire* cost of a trip if it was undertaken principally for business reasons.[40]

But what Congress giveth, Congress taketh away. The Revenue Act of 1964 contained a provision repealing the 1962 section and thus reinstating the prior law, except as to travel "outside the United States."[41]

The Senate Report explains that, under the regulations, "the area of application of the provision is so restricted, since it applies only to self-employed persons and to employees who are managing executives or related to the employer . . . that the provision in its present form served little purpose."[42] But it is an odd governmental phenomenon indeed when the legislative branch repeals a law because the administrative rules adopted only by legislative grace happen to restrict the operation of the law unduly. Section 274(c) should have been *strengthened,* not repealed.[43]

---

involves some inexplicable omissions on the part of Congress. This area is again the subject of proposed reform. *See* Patterson, *Kennedy Pushes for Tax Reform,* The Boston Globe, Mar. 19, 1976, at 25.

38 *See* S. REP., *supra* note 27, at 34-35.

39 INT. REV. CODE OF 1954, § 274(c). The regulations set forth a definite standard for determining whether travel exceeds one week, and for applying the 25 percent rule. *See* Treas. Reg. §§ 1.274-4(c), 1.274-4(d)(2) (1964).

40 The "primary purpose" test has been criticized as being too subjective to be workable. *See* Klein, *The Deductibility of Transportation Expenses of a Combination Business and Pleasure Trip — A Conceptual Analysis,* 18 STAN. L. REV. 1099, 1104-06 (1966).

41 Revenue Act of 1964, § 217, 78 Stat. 56, *amending* INT. REV. CODE OF 1954, § 274(c). The Senate Finance Committee, the source of the 1962 law, wanted to abolish the section entirely, but compromised with the House on keeping the allocation rule applicable to foreign travel. H. CONF. REP. No. 1149, 88th Cong., 2d Sess. (1964), *reprinted in* U.S. CODE CONG. AND AD. NEWS 1940, 1969-70.

42 S. REP. No. 870, 88th Cong., 2d Sess. (1964), *reprinted in* 1964 U.S. CODE CONG. & AD. NEWS 1754. The regulations provide that allocation is unnecessary if "the individual incurring such [travel] expenses did not have substantial control over the arranging of the trip." Treas. Reg. § 1.274-4(f)(5)(i) (1964). This provision virtually eliminated the necessity for allocation for all employees who are not at the upper echelons of a business hierarchy.

43 Congressional repeal seems to have been overly hasty, at the very least, in light of the fact that the regulations were not adopted until June 21, 1963, and

Present § 274(c) has not been employed by the courts either to limit or to allocate the costs of foreign travel. The sweeping regulations which were the basis of the 1964 amendment probably are to blame here.[44] Some courts have shown a remarkable willingness to allow deductions for trips abroad which are substantially personal in nature.[45] Because some trips may be taken exclusively for business reasons, transportation costs should be treated more liberally than entertainment expenditures. But there is no justification for not demanding allocation of a mixed business and personal trip and for setting limits to the amounts deductible for meals and lodging even while on an exclusively business trip, since these are clearly personal in nature.[46] The original Kennedy proposal would have this effect, and should be enacted.

C. *Entertainment*

Section 274(a) provides separate tests for entertainment activities and facilities.

Section 274(a)(1)(A) limits the deductibility of expenditures for entertainment *activities;* § 274(a)(1)(B) further limits the deductibility of expenses connected with entertainment *facilities.*[47]

The former establishes the general rule that no deduction shall be allowed for expenses stemming from entertainment activities

---

the Senate Report was dated January 28, 1964. *See* T.D. 6659, 1963-2 CUM. BULL. 113; S. REP. No. 830, 88th Cong., 2d Sess. 1313 (1964), *reprinted in* 1964 U.S. CODE CONG. & AD. NEWS 1673.

44 See note 42 *supra*. Under the regulations, a day is deemed entirely for business even though only a few hours were spent on business pursuits. And even if this condition is not met, it is still a business day — allowing full deductions — if the taxpayer can show that his presence outside the United States was required for a specific and bona fide business purpose. Treas. Reg. §§ 1.274-4(d)(2)(ii), (iii) (1964). Thus, if a person can show that he left for Europe on Thursday and spent five hours on business on Friday and another five on Monday, the entire trip would be considered for business purposes and all traveling expenses deductible accordingly.

45 *See, e.g.*, Wilkins v. United States, 348 F. Supp. 1282 (D. Neb. 1972), *aff'd*, 486 F.2d 1407 (8th Cir. 1973) (full deduction allowed for wife's traveling expenses); John C. Ford, 56 T.C. 1300 (1971), *aff'd per curiam*, 487 F.2d 1025 (9th Cir. 1973) (full deduction allowed for traveling expenses connected with enrollment in Ph. D. program in Oslo, Norway); Stanley Marlin, 54 T.C. 560 (1970) (teacher of Western Civilization allowed full deduction for trip to France).

46 See note 36 *supra*.

47 INT. REV. CODE OF 1954, §§ 274(a)(1)(A), (B). The term "facility" was new to the Code and caused some initial confusion. *See, e.g.*, Emmanuel and Lipoff, *supra* note 36 at 488; Skinner, *1962 Act: Burdens Imposed by the T & E Rules Require New Expense Account Practices*, 17 J. TAXATION 360, 361 (1962).

unless the taxpayer proves that such expenses are either "directly related to" or "associated with" the active conduct of a trade or business.[48] The "associated with" test was added by the Senate to permit the deduction of certain expenses related to goodwill entertainment.[49] However, this test applies only to activities *"directly preceding or following a substantial and bona fide business discussion"* (emphasis added).[50] Because of a very narrow interpretation by the courts, it rarely is a basis for an entertainment deduction not otherwise allowed.[51]

Under the "directly related" test, the taxpayer supposedly must show a greater degree of proximate relation between the entertainment expenditure and his trade or business than is required under § 162.[52]

Entertainment also will automatically be considered directly related if it occurs in a "clear business setting."[53] Under these

48 INT. REV. CODE OF 1954, § 274(a)(1)(A). "Entertainment, amusement, or recreation" as delineated in the Code is reduced to "entertainment" by Treas. Reg. 1.274-2(b)(1)(i), and then *circularly defined* to include "an activity which is of a type generally considered to constitute entertainment, amusement, or recreation. . . ." Fortunately, there are examples. Treas. Reg. § 1.274-2(b)(1)(i) (1963).

49 The Senate Finance Committee explained:

> Goodwill has long been recognized as a legitimate objective of business entertaining and where the purpose of the expense and its clear relationship to a business is firmly established, the expense ordinarily will continue to be deductible.

S. REP., *supra* note 27, at 28.

50 INT. REV. CODE OF 1954, § 274(a)(1)(A). For a discussion of the demands of the "associated with" test, see Israelton v. United States, 367 F. Supp. 1104, 1107-08 (D. Md. 1973), *aff'd*, 508 F.2d 838 (4th Cir. 1974) (vague possibility of goodwill inadequate to allow deductions for party at country club); St. Petersburg Bank & Trust Co. v. United States, 362 F. Supp. 674, 680-81 (M.D. Fla. 1973) (denial of deduction for parties because of lack of any substantial business discussions).

51 Few cases have turned on the question of whether entertainment was "associated with" a trade or business. However, most instances where such a claim might arise have been denied on the basis of failure under the threshold requirement of § 162 (indicating, as we shall often see, tight enforcement of the "ordinary and necessary" requirement) or because of failure to substantiate the claim as required by § 274(d). See text accompanying notes 89-97 *infra*.

52 The regulations set forth two principal requirements under this test: (1) The taxpayer must have more than a general expectation of deriving income or some business benefit, other than good will, at some indefinite future time from the making of the expenditure; and (2) The taxpayer must actively engage in a meeting, discussion, negotiation, or other bona fide business transaction, other than entertainment, for the purpose of obtaining income or some other specific business benefit. Treas. Reg. § 1.274-2(c)(3)(i), (ii) (1963). *See also* H.R. REP., *supra* note 4, at 19.

53 Treas. Reg. § 1.274-2(c)(4) (1963).

amorphous additions to the basic requirements, a taxpayer must "clearly [establish] that any recipient of the entertainment would have reasonably known that taxpayer had no significant motive, in incurring the expenditure, other than directly furthering his trade or business."[54] The only thing that seems clear from this language is that a court may do just about what it wishes in interpreting it.[55]

Section 274(e) sets forth nine exceptions to the general disallowance provisions of § 274(a).[56] Most of these involve items which would not be contested in any event, and it therefore is not surprising that there has been very little litigation involving defenses based on the exceptions.[57] All of the items listed in § 274(e) supposedly are governed by prior law,[58] although some of the excepted items are subject to the substantiation requirements of § 274(d).[59]

Entertainment facilities under the Code include any item of personal or real property owned or rented by the taxpayer.[60] They

---

54 *Id. See also* Rev. Rul. 63-144, 1963-2 CUM. BULL. 129, 136 (questions 38-40).

55 As under all of the substantive rules of § 274(a), however, few courts have had either the need or desire to speculate on what is a "clear business setting." An exception is D.A. Foster Trenching Company, Inc. v. United States, 473 F.2d 1398 (Ct. Cl. 1973), where the court said that in order for a "clear business setting" to be established, the recipients of the entertainment must *know* that the entertainment was clearly business motivated.

56 INT. REV. CODE OF 1954, § 274(e). Section 274(e) provides that subsection (a) shall not apply to the following:

(1) Business meals;
(2) Food and beverages for employees;
(3) Expenses treated as compensation;
(4) Reimbursed expenses;
(5) Recreational, etc., expenses for employees;
(6) Employee, stockholder, etc., business meetings;
(7) Meetings of business leagues, etc.;
(8) Items available to public;
(9) Entertainment sold to customers.

57 The only case where the court invoked these exceptions is Richard Haman, CCH Tax Ct. Mem. 466 (1972), where both § 274(e)(1) and § 274(e)(5) were cited to allow deductions for meals purchased and Christmas parties given for a boatman's crew.

58 H.R. REP., *supra* note 4, at 24; S. REP., *supra* note 27, at 36; Rev. Rul. 63-144, 1963-2 CUM. BULL. 129, 131 (question 6).

59 Treas. Reg. §§ 1.274-5(a)(2), (c)(7) (1962). *See, e.g.*, Steel v. Comm'r, 437 F.2d 71 (2d Cir. 1971) (exception for business meals applies only to subsection (a) and not to subsection (d)'s substantiation requirements, which the taxpayer failed to meet).

60 Treas. Reg. § 1.274-2(e)(2)(i) (1963). The House and Senate Reports give as

also include dues and fees paid to social and athletic clubs or other organizations.[61] In addition to meeting the "ordinary and necessary" test of § 162, the facility must be "used primarily for the furtherance of the taxpayer's trade or business."[62] Once this test has been met, deductions for the costs of the facility will be allowed to the extent that the activities in the facility pass the "directly related" test discussed above.[63] The "associated with" test is *not* applicable to facilities.[64]

The requirement that the facility be used "primarily" for business contemplates that it be used for *"ordinary and necessary"* entertainment more than one-half of the time that it is used.[65] Failure to meet the "primary purpose" test does not mean that the expenditures for entertainment *activities* in the facility which meet the "directly related" test are not deductible.[66] It only means

examples a yacht, hunting lodge, fishing camp, swimming pool, tennis court, bowling alley, automobile, airplane, apartment, hotel suit, home in vacation resort, dining room, and cafeteria. H.R. REP., *supra* note 4, at 21; S. REP., *supra* note 27, at 31.

61 Treas. Reg. § 1.274-2(e)(3)(ii) (1963). Dues or fees paid to any "social, athletic, or sporting club or organization" are included. The regulations provide: "Generally, the phrase 'social, athletic, or sporting club or organization' has the same meaning for purposes of this section as it has in part II of Chapter 33 of the Code, and the regulations thereunder. . . ." Unfortunately, part II of Chapter 33 was repealed in 1965. Pub. L. No. 89-44, § 301, 79 Stat. 145 (1965). The regulations thereunder are no longer published.

The more rigorous restraint of this category of deductions undoubtedly stems from the Treasury Department's 1960 Study on Entertainment Expenses, which indicated that numerically the claims for club dues and fees was the largest category of travel and entertainment deductions. INTERNAL REVENUE SERVICE (Dep't of the Treasury), STUDY ON ENTERTAINMENT EXPENSES (April 1961), appearing as Exhibit V, *Hearings*, *supra* note 20, at 133 [hereinafter cited as TREAS. STUDY]. Out of a total of 25,032 returns examined, 5,384 of them claimed deductions for club dues or fees. *Id.* at 193 (Table B).

This survey was undertaken pursuant to a Congressional request to the Secretary of the Treasury in 1960. Beginning in September, 1960, the I.R.S. conducted a three-month audit survey of travel and entertainment deductions. In all, approximately 38,000 returns were examined including T & E deductions. The results of the survey then were summarized in the Treasury Study and presented to Congress by the President. It is the only such survey in this area ever done. *See* Hannon, *The New Travel and Entertainment Rules: A Perspective*, 117 J. ACCT. 20 (Feb. 1964) [hereinafter cited as Hannon].

62 INT. REV. CODE OF 1954, § 274(a)(1)(B).

63 *Id. See* H.R. REP., *supra* note 4, at 21. See text accompanying notes 52-55 *supra*.

64 H. CONF. REP., *supra* note 26, at 3736. *Cf.* S. REP., *supra* note 27, at 31.

65 Rev. Rul. 63-144, 1963-2 CUM. BULL. 129, 137 (question 46). *See* H.R. REP., *supra* note 4, at 22; S. REP., *supra* note 27, at 31.

66 Rev. Rul. 63-144, 1963-2 CUM. BULL. 129, 137 (question 48).

that no portion of the costs of maintaining the facility may be deducted.

III. Entertainment Facilities: The Goodwill Limitation

Entertainment facilities such as yachts and resort lodges were the primary target of the disallowance provisions of § 274.[67] This emphasis stems from the Treasury Department's Study, which reported a large number of deductions claimed and disallowed in this category.[68] The Study's analysis of "abuse cases" decided by the Tax Court heavily emphasized facilities rather than activities, travel, or gifts — forty-one of the fifty cases dealt with facilities.[69] The President,[70] the Secretary of the Treasury,[71] the House[72] and the Senate[73] all listed facilities as the prime example of abuse, the deductions for which required restraint. Was this attack necessary? Was it successful?

Allowance of full deductions for expenses related to entertainment facilities always has been rare. Of sixty-six cases analyzed in this category, only nine conceded the entire claimed deduction to the taxpayer.[74] Some of these cases had factual peculiarities

67 See note 47 *supra.*

68 *See* Treas. Study, *supra* note 61, at 193 (Table B). The Study indicates startlingly high disallowance rates for certain facilities, such as 60 percent for yachts and 55 percent for resort properties. Dues and fees paid to clubs represented the largest category of deductions, but with a comparatively low adjustment rate of 34 percent. *Id.* at 192-93 (Tables B and C). These statistics obviously impressed the Treasury and Congress. They also must have had an impact on the courts since § 162's ordinary and necessary test has been stricter since 1962, often without reference to the rules of § 274.

69 Treas. Study, *supra* note 61. It should be noted, though, that only nine out of the 50 examples are identified by case name, all of which were included in the cases analyzed for this Note.

70 *Hearings, supra* note 16, at 13.

71 *Id.* at 42 (Statement of Secretary Dillon).

72 H.R. Rep., *supra* note 4, at 21.

73 S. Rep., *supra* note 27, at 30.

74 *See, e.g.,* McAdams v. United States, 339 F. Supp. 926 (M.D. Tenn. 1971), *rev'd sub nom.* Hippodrome Oldsmobile, Inc. v. United States, 474 F.2d 959 (6th Cir. 1973) (yacht); George Durgom, 33 CCH Tax Ct. Mem. 276 (1974) (resort home); Cleveland-Sandusky Brewing Corp., 30 T.C. 539 (1958) (yacht); E.E. Dickinson, 8 B.T.A. 722 (1927) (yacht).

explaining these results.[75] Others involved the one especially favored group of taxpayers — those in the entertainment industry.[76] But most cases prior to the enactment of § 274 suggest that a "direct relation" was indeed necessary under prior law.[77]

It thus appears that the substantive test under § 274 requiring a "direct relation" did not change prior law in the area of entertainment facilities.[78] For example, in *International Trading Co. v. Commissioner*,[79] a closed corporation claimed a deduction for the depreciation and expense of a lake-front resort.[80] The court denied the claim in full, first noting that "[a]ll deductions, whether with respect to individuals, or corporations, are a matter of legislative grace,"[81] and then defining "ordinary and necessary" under § 162 as requiring that expenses be "directly connected with" the trade or business.[82]

The substantive law concerning organizational dues and fees was fixed by the 1937 case of *Louis Boehm*.[83] The taxpayer, a New York City attorney, claimed business deductions for dues and entertainment expenses of several clubs and lodges.[84] The court denied the deduction in full, stating:[85]

75 *See, e.g.*, E. E. Dickinson, 8 B.T.A. 722 (1927). The taxpayer's yacht had been leased to the government for naval service in World War I and the taxpayer's son put in command of the vessel. The court's opinion details at length the severe conditions under which the boat traveled and the repairs which the taxpayer had to make after its return. 8 B.T.A. at 723.

76 *See, e.g.*, George Durgom, 33 CCH Tax Ct. Mem. 276 (1974) (a well-known Hollywood agent allowed his full claim for expenses relating to a resort home in Palm Springs and a projection room in his Beverly Hills residence despite poor substantiation of business versus personal use); United States v. Disney, 413 F.2d 783 (9th Cir. 1969); Blackmer v. Comm'r, 70 F.2d 255 (2d Cir. 1934); International Artists, Ltd., 55 T.C. 94 (1970); Oswald "Ozzie" G. Nelson, 25 CCH Tax Ct. Mem. 1142 (1966).

In *Durgom*, the court commented that "in show business personal contacts and personal acquaintances are very important" in allowing the deduction. 33 CCH Tax Ct. Mem. at 277. In the famous (George M.) *Cohan* case, the Second Circuit concluded that the taxpayer's business "forced him to be free-handed in entertaining actors, employees and . . . dramatic critics." 39 F.2d at 543.

77 *See, e.g.*, Lanteen Medical Laboratories, Inc., 10 T.C. 279, 290 (1948) (deduction for golf course construction disallowed).

78 The one exception, as we shall see, is in the area of "direct goodwill entertainment." See text accompanying notes 88-97 *infra*.

79 275 F.2d 578 (7th Cir. 1960).

80 *Id.* at 581-83.

81 *Id.* at 584.

82 *Id.* at 585.

83 35 B.T.A. 1106 (1937).

84 *Id.* at 1108-09.

85 *Id.* at 1109-10.

> We do not think the burden of proof is met by the argument that in general, membership in social, political, and fraternal organizations is helpful in obtaining clients through contacts made thereby. . . . No evidence has been introduced to show that any part or all of the expenditures in question were so closely related to show that such expenditures had a *direct relation* to the conduct of a business or the business benefits expected (emphasis added).

This rule has been followed ever since.[86] Thus the same nexus between expenses and the business was demanded under former law as was codified by § 274(a).

The key difference that the substantive rules[87] of the new law have made is in denying deductions for facilities used for promotion of *general business goodwill.* Each of the three cases in this category reaching the appellate courts was decided solely or alternatively on the goodwill issue.[88]

A distinction can be drawn from the cases between what can be called "direct goodwill entertaining," and "indirect goodwill entertaining." Where a taxpayer himself accompanies actual or potential customers in entertainment activities for the specific purpose of generating income from those entertained but discusses business only vaguely, he is engaged in direct goodwill entertaining. Where, however, the taxpayer does not discuss business or is not even present during the entertainment, but the activity nonetheless will provide him with public exposure leading to more business, then he is engaged in indirect goodwill enter-

---

86 The *Boehm* case was cited and followed in Ronald W. Sholund, 50 T.C. 503, 508-09 (1968); Alexander P. Reed, 35 T.C. 199, 202 (1960); Chas. D. Long, 32 T.C. 511 (1959), *aff'd* 277 F.2d 239 (8th Cir. 1960). As the dates of the cases indicate, the *Boehm* rule has been applied even under § 274. The greater generosity of district courts is illustrated by several cases allowing deductions of this type. *See, e.g.,* Holland v. United States, 311 F. Supp. 422 (C.D. Ca. 1970) (Crown Zellerbach's vice-president allowed deduction of one-half of country and luncheon club dues); Johnson v. United States, 45 F. Supp. 377 (S.D. Ca. 1941) (allowed three-fourths of country club dues of attorney on evidence of personal dislike of golf and direct connection to large fees obtained).

87 As the term is used in this Note, "substantive rules" refers to the tighter tests of proximate relationship for activities and facilities established by § 274(a). These should be distinguished from what properly are termed the "procedural" requirements of substantiation under § 274(d).

88 *See* Handelman v. Comm'r, 509 F.2d 1067 (2d Cir. 1975); Hippodrome Oldsmobile, Inc. v. Comm'r, 474 F.2d 959 (6th Cir. 1973); William K. Coors, 60 T.C. 368 (1973), *aff'd without discussion of this point sub nom.* Adolph Coors Co. v. Comm'r, 519 F.2d 1280 (10th Cir. 1975), *cert. denied,* 96 S. Ct. 878 (1976).

taining. It is important to note that deductions for indirect goodwill activity always have been disallowed, as was seen in the *Boehm* case.[89]

Prior to the enactment of § 274, however, deductions for facilities used for direct goodwill promotion were allowed. For example, in *Cleveland-Sandusky Brewing Corp. v. Commissioner*[90] a brewery owned a cabin cruiser on which it entertained customers and distributors, including taking persons out on weekend fishing trips.[91] The case thus is illustrative of direct goodwill promotion. Nonetheless, the court held that the ordinary and necessary test of § 162 was met, and permitted the deduction in full.[92] This result was typical of those reached in other pre-§ 274 cases.[93]

By comparison, in *Hippodrome Oldsmobile, Inc. v. United States,*[94] the Sixth Circuit denied a deduction for a company yacht using the substantive tests of § 274(a). The court quoted the House and Senate reports extensively in concluding that "more than a general expectation of deriving some income at some indefinite future time" is necessary.[95] The Court of Appeals accepted the findings of the District Court (which decided for the taxpayer), that the boat was used to improve business relationships and indeed did have a beneficial effect on business,[96] but concluded that the "soft-sell" approach of the taxpayer was insufficient to

89 *See, e.g.,* Donald W. Barker, 31 CCH Tax Ct. Mem. 234 (1972) (dentist denied deduction under § 162 for racing sailboat on which he entertained persons he hoped to attract as clients); Ben R. Stein, 31 CCH Tax Ct. Mem. 663 (1972) (boat deduction disallowed because of failure to qualify under § 162); Challenge Manufacturing Co., 37 T.C. 650 (1962) (indirect business benefit inadequate); American Lithofold Corp., 55 T.C. 904 (1971) (*Challenge Manufacturing* cited in denying deductions for use of boat deemed for personal entertainment of principal shareholder); Robert Lee Henry, 36 T.C. 879 (1961) (the "1040 Flag Case" denying any deduction to attorney claiming boat's name and flag drew business to his tax practice).

90 30 T.C. 539 (1958).

91 *Id.* at 546.

92 *Id.*

93 *See, e.g.,* United Anilene Co. v. Comm'r, 316 F.2d 701 (1st Cir. 1963) (one half deduction allowed for yacht on which textile company entertained customers); First National Bank of Omaha v. United States, 276 F. Supp. 905 (D. Neb. 1967) (country club expenses allowed bank for entertainment for nurturing of personal friendships); Charles M. Kilborn, 29 T.C. 102 (1957) (one-third deduction allowed for cabin cruiser on which used car salesman entertained customers).

94 474 F.2d 959 (6th Cir. 1973).

95 474 F.2d at 961-64, *quoting* H.R. REP., *supra* note 4, at 20.

96 339 F. Supp. at 828; 474 F.2d at 960.

show the "direct relationship" required between the expense and the deduction.[97]

Section 274 clearly has had a positive effect in limiting deductions for entertainment facilities used for direct goodwill promotion. However, these disallowances are based merely on the nonapplicability of the "associated with" test to entertainment facilities. We are left with a picture that does not evoke much confidence in the new section. The section does not require disallowance of deductions for uses of facilities other than for promotion of goodwill even though there is a major element of personal pleasure and benefit derived from such use. Entertainment facilities, even more than activities, inherently are subject to a large measure of personal versus business use, if only because they are always there to be enjoyed. Accordingly, new nexus tests are not the solution. Rather, a fundamental redefinition of the net income concept is needed to eliminate, once and for all, federal subsidization of personal pleasure. Deductions for entertainment facilities should be disallowed completely.

IV. ENTERTAINMENT ACTIVITIES: THE DEMISE OF *Cohan* AND THE RISE OF SUBSTANTIATION

An express purpose of § 274 was the abolition of the *Cohan* rule.[98] Under this liberal approach, taxpayers were permitted to estimate broadly their expenditures. Many of taxpayers' claims probably failed to meet the required nexus tests of § 162. The

---

97 474 F.2d at 960. The closing paragraph of the opinion states:

> Under the language of § 274(a)(1)(B) and its legislative intent . . . the corporate taxpayer which claims a deduction for entertainment not covered by § 274(a)(1)(A) must in meeting the statutory requirements show a definable business purpose for the occasion (other than the creation of general good will) . . .

See also Handelman v. Comm'r, 509 F.2d 1067 (2d Cir. 1975) (denying deduction for yacht of attorney claiming goodwill use).

The problem of claimed deductions for goodwill is not limited to the category of entertainment facilities, of course. *See, e.g.*, St. Petersburg Bank & Trust Co. v. United States, 362 F. Supp. 674 (M.D. Fla. 1973) *aff'd mem.*, 503 F.2d 1402 (5th Cir. 1974), *cert. denied* 96 S. Ct. 58 (1975) (disallowing deductions for parties aimed at generating goodwill).

98 See note 29 *supra*.

courts, confronted with a net income concept which permitted mixed personal and business expense deductions in the T & E category, retreated behind the *Cohan* rule and inevitably permitted *some* amount of the deduction.[99] Reacting to such allowance, Congress replaced *Cohan's* allocation rule with the strict substantiation requirements of § 274(d). It is doubtful whether this change in procedure, without a more fundamental re-evaluation of the allowance of any T & E deductions, has accomplished very much besides confusion.

The best example of the confusion is in the category of business meals — the first and major exception to the substantive tests of § 274(a).[100] The rules as to what constitutes a "business" meal as opposed to any other kind of meal are more complicated than, and overlap, other regulations under the section.[101] The courts therefore have bypassed completely the problem of defining the

---

99 *See, e.g.*, Berkeley Machine Works Foundry Co., Inc. v. Comm'r, 422 F.2d 362 (4th Cir. 1970), *aff'g* 27 CCH Tax Ct. Mem. 1487 (1968) (one-half of the deduction allowed for fishing lodge despite inadequate evidence); Chas. D. Long, 32 T.C. 511 (1959) (two-thirds of athletic and country club dues and expenses allowed); William T. Stover Co., 27 T.C. 434 (1956) (one-half deduction allowed for boat despite records which "do not even closely indicate the portion of the expenditures which should be attributed to petitioner's business"); Charles M. Kilborn, 29 T.C. 102 (1957) (allowance of one-third of costs of yacht despite evidence based on "generalities"); Richard A. Sutter, 21 T.C. 170, 174 (1953) (allocation of one-fourth of expenses of boat despite indication of personal nature); Norman M. Hussey, 11 CCH Tax Ct. Mem. 141 (1952) (two-thirds of golf club dues and fees allowed).

100 INT. REV. CODE OF 1954, § 274(e)(1); Treas. Reg. § 1.274-2(f)(2)(i) (1963).

101 The circumstances for the meal must be "conducive to business discussion" and the person to whom the meal is furnished "must be such as will reasonably indicate that the food or beverages were furnished for the primary purpose of furthering the taxpayer's trade or business and did not primarily serve a social or personal purpose." All of this sounds strikingly similar to the regulations under § 274(a) which consider entertainment in a "clear business setting" as satisfying the "directly related" test. Thus, it is submitted, there is really very little difference between the "business meal" exception and the "clear business setting" rule.

It could be argued that the real difference is that under the former "[t]here is no requirement that business actually be discussed." Treas. Reg. § 1.274-2(f)(2)(i) (1963) (last sentence). However, there is no indication that business must be discussed in the "clear business setting." Indeed, the examples given would indicate that there is no such requirement. For example, a taxpayer owning a hotel may provide occasional free dinners at the hotel for regular customers, and this will qualify as a "clear business setting." Business representatives may be entertained at the opening of a new hotel in order to obtain new business and also qualify. Treas. Reg. § 1.274-2(c)(4) (1963). Presumably, the I.R.S. does not expect the hotel owners to run up to the old or potential customers and talk business. Therefore, such occasions seem to fit both "exceptions."

scope of the business meal exception, and have instead zeroed in on the substantiation requirements of § 274(d) as the basis for disallowing deductions for meals.[102]

Prior to the adoption of the new law, allocation was the predominant result of litigated cases in this category.[103] The so-called "*Sutter* rule," disallowing meals as too personal in nature, usually was the basis for most of the disallowed portions of the claims.[104]

Yet the *Sutter* rule has been limited under § 274 to "abuse cases."[105] And, indeed, the *Sutter* rule has not been the basis for any disallowance under § 274. This limitation of the *Sutter* rule seems inconsistent with the goal of the section: eliminating tax support for personal expenses in the T & E area. Practically speaking, however, the cut-back on the *Sutter* rule has had little effect. The courts need not find "abuse"; they only have to *not* find adequate substantiation.

An example is found in the first stab of the appellate courts at § 274(d), *Sanford v. Commissioner*.[106] During 1963 the taxpayer maintained a diary in which he recorded his business entertainment expenses, mostly luncheons and dinners. The Commissioner disallowed all expenditures of twenty-five dollars or more because the taxpayer failed to substantiate his expenses with receipts as required by § 274(d) and the accompanying regulations.[107] The Tax Court sustained the Commissioner's disallowance, holding that on these facts, a diary does not meet the "adequate records" or "other sufficient evidence" requirement of § 274(d), and that the regulations supporting the Commissioner's position are valid.[108]

102 Of 11 cases analyzed under the new law, ten were decided on the basis of § 274(d).

103 *See, e.g.*, Lewis F. Jacobson, 6 T.C. 1048 (1948) (one-fourth of unsubstantiated claim for meals allowed); John R. Guglielmetti, 35 T.C. 668, 672 (1961) (200 dollars allowed out of claim of 2500 dollars); Donald G. Teeling, 42 T.C. 671 (1964) (three-fourths of meals allowed despite "vague" evidence and personal nature).

104 *See* Richard A. Sutter, 21 T.C. 170 (1953), *acquiesced in*, 1954-1 CUM. BULL. 6 (complete disallowance of a variety of travel and entertainment expenses as personal and nondeductible under § 262). It is paradoxical that the *Sutter* rule resulted in allocations when applied, especially in light of the language and the outcome of *Sutter* barring any deductions for such personal expenditures.

105 Rev. Rul. 63-144, 1963-2 Cum. Bull. 129, 135 (question 31).

106 412 F.2d 201 (2d Cir.), *cert. denied*, 396 U.S. 841 (1969), *noted in* 35 Mo. L. REV. 70 (1970).

107 *See* Treas. Reg. § 1.274-5(c) (1962).

108 412 F.2d at 202.

The Second Circuit affirmed. Thus, a well-kept personal record — even under § 274 believed sufficient to substantiate expenses[109] — failed the rigid test of § 274(d).[110]

It seemed as if the strict disallowances foreseen under § 274 were to be accomplished — even if by the circuitous route of denying deductions because of failure to substantiate, rather than because of the personal element of the expense. But a reverse conclusion was prompted by the same Circuit a few months later in *LaForge v. Commissioner*.[111] While *Sanford* addressed the question of what constitutes "adequate records" under § 274(d), *LaForge* dealt with the alternative requirement of "other sufficient evidence" included in the statute.[112] Under the regulations as issued in 1962, the "other sufficient evidence" required an itemized written statement for each expenditure.[113]

The taxpayer in *LaForge* was a surgeon who regularly purchased luncheons for residents and interns who assisted him. The luncheons were bought in a hospital cafeteria that issued no receipts and LaForge kept no records of the expenditures. Despite the lack of substantiation or record, LaForge treated these expenditures as business entertainment expenses and deducted two dollars per day for each day he worked in the hospital.[114] The Internal Revenue Service conceded that the deduction met the substantive requirements of §§ 162 and 274 but contended that the testimony of the taxpayer and of the cafeteria cashier failed to satisfy the "other sufficient evidence" requirements. The deductions thus were disallowed completely.[115]

The taxpayer contended that the requirement of the regulations demanding a written statement of claimed business related entertainment expenditures exceeded the statutory requirements. The

109 The taxpayer's attorney advised him that documentary evidence to the extent stated in the regulations was unnecessary to support his personal record. 50 T.C. 823, 830 (1968).

110 *Accord,* Robinson v. Comm'r, 422 F.2d 873 (9th Cir. 1970) (disallowing deductions evidenced only by personal records); John L. Ashby, 50 T.C. 409 (1968) (disallowing deductions for a boat substantiated by log and weekly expense reports).

111 434 F.2d 370 (2d Cir. 1970).

112 INT. REV. CODE OF 1954, § 274(d).

113 Treas. Reg. § 1.274-5(c)(3) (1962), *as amended* Treas. Reg. § 1.274-5(c)(3)(i) (1974).

114 434 F.2d at 371.

115 *Id.*

court agreed. It believed that to require a written itemized statement is, in effect, to require "adequate records" — thereby abolishing the disjunctive substantiation requirement of the section.[116] The court quoted the Senate Report and concluded that the statute contemplated substantiation by oral testimony.[117] Accordingly, the regulation was held invalid.[118]

The ink was barely dry on the *LaForge* opinion before the Second Circuit started narrowing its holding. In *Steel v. Commissioner*,[119] a taxpayer claimed a deduction for business meals. The only evidence introduced was a few cancelled checks and the taxpayer's *own* testimony. The court distinguished *LaForge* and disallowed the deduction squarely on § 274(d).[120]

In *Hughes v. Commissioner*,[121] decided just a few weeks after *Steel*, the taxpayer was a stage manager who frequently purchased minor food items for his crew and claimed a deduction for them.[122] The taxpayer produced a few receipts in support of his claim, but relied principally on the corroborative testimony of two witnesses who verified that he did supply his crew with snacks and that this practice was usual for the industry.[123] The court nonetheless disallowed the deduction in full.[124]

The Second Circuit struggled with distinguishing *LaForge*.[125] Hughes' evidence was condemned because there were no "written records, receipts, or other [documentation]."[126] Neither did LaForge present any "written" substantiation. It is difficult to see much difference between Hughes buying morning donuts and afternoon snacks for his stagehands and LaForge purchasing cafeteria lunches for his medical assistants. Indeed, the evidence sug-

116 *Id.* at 372.

117 *Id.*, quoting S. REP., *supra* note 27, at 35.

118 434 F.2d at 373.

119 437 F.2d 71 (2d Cir. 1971).

120 *Id.* at 72. The taxpayer was a New York City attorney, and the business meals were all with clients. Many of them were provided in the taxpayer's home, which explained in part the lack of documentation. The court held that the same recording requirements apply to entertainment in the home.

121 451 F.2d 975 (2d Cir. 1971).

122 *Id.* at 976.

123 *Id.* at 977-78.

124 *Id.* at 978.

125 As have other courts. *LaForge* has been discussed and distinguished in at least 25 Tax Court opinions since 1970.

126 451 F.2d at 978.

gested that Hughes was following the expectation and practice of the industry while Dr. LaForge simply was being understandably but unnecessarily generous.[127] The conclusion is inescapable that the court was searching for an opportunity to narrow *LaForge*, perhaps realizing its holding was too broad. The Second Circuit's problems with the statute reveal the tension in § 274 itself between the concept and the rule. Further problems seemed inevitable.

The most recent case continues this pattern of uncertain application of § 274(d). In *Dowell v. United States*,[128] the taxpayer was the president of a bank, and took a deduction for the entertainment[129] of potential customers. This was his principal tool for cultivating new accounts.[130] The District Court allowed all of the claims,[131] while the Court of Appeals was unsatisfied.[132] The key to both decisions was whether the taxpayer's evidence satisfied § 274(d).

The District Court found that the times, places and amounts of expenditures clearly had been provided by "a virtual blizzard of bills, chits and other papers relating to Dowell's meals with other persons."[133] The court also determined that those elements for which adequate records had not been provided, namely, the business purpose and business relationship of those entertained, had been adequately substantiated by Dowell's own testimony as corroborated by the oral testimony of some 20 witnesses.[134] The *LaForge* case was cited and followed in the District Court's holding that the taxpayer had proven his claim beyond doubt by "other sufficient evidence." The court believed that direct substantiation

---

127 *Id.* at 977.

128 522 F.2d 708 (5th Cir. 1975), *rev'g* 370 F. Supp. 69 (N.D. Tex. 1974).

129 Dowell's entertaining consisted primarily of lunches and dinners — "business meals" — with well-heeled members of the business community. He also deducted the cost of travel from his suburban bank to the downtown business district for these occasions.

130 There was no question that the entertainment satisfied all the substantive tests of §§ 162 and 274. Dowell was recognized by the District Court as a man possessed with his bank and other businesses, devoting 14-15 hours a day to these enterprises. *Id.* at 72. The Court of Appeals simply noted Dowell as "a dynamo without a doubt. . . ." 522 F.2d at 714.

131 370 F. Supp. at 75.

132 522 F.2d at 708-09.

133 370 F. Supp. at 70.

134 *Id.* at 73-74.

"of every dime . . . inserted into parking meters . . . is impractical and not required" under the law.[135]

The Court of Appeals disagreed. After tracing the legislative history of § 274 and quoting most of the substantiation regulations, it concluded that the District Court "misapprehended the specificity with which the taxpayer must substantiate each expenditure deducted under § 274(d)."[136] As to the dimes inserted in parking meters, the court saw "*no reason* why the taxpayer should not be held to the *same substantiation requirements* regarding these miscellaneous expenditures as for [others]" (emphasis added).[137] It thus said that the taxpayer would have to substantiate even those *local travel expenses* such as parking, taxis, telephone calls, or even a coke.[138]

The District Court was overly liberal with Dowell and needlessly neglectful in not recognizing the distinction between "adequate records" and "other sufficient evidence." But the Court of Appeals was not only harsh, but in some instances dead wrong. For one thing, the substantiation rules of § 274(d) apply only to travel "away from home,"[139] and under no definition of that phrase would daytime trips of Dowell downtown fall under the section. Therefore, the expenditures for parking, taxis, and telephone calls on these short trips would not have to meet "the same substantiation requirements" as other expenditures, contrary to the court's holding.

The court also seems to ignore the fact that Dowell's documentation and testimony overlap. It seems to say that the documentation fails under the "adequate records" test and that the testimony fails under the "other sufficient evidence" test.[140] Together, the written and oral evidence would seem to satisfy the "other sufficient evidence" requirement. Moreover, it is not necessary to obtain documentation for expenditures of less than twenty-five dollars.[141] Most of Dowell's expenditures were for lunches or dinners and, therefor,

---

135 *Id.* at 75.
136 522 F.2d at 714.
137 *Id.* at 716-17.
138 *Id.* at 716.
139 Treas. Reg. § 1.274-5(a)(1) (1962).
140 522 F.2d at 716.
141 Treas. Reg. § 1.274-5(c)(iii)(b) (1962).

would probably fall within this "no documentation necessary" rule.[142]

Another omission of the Fifth Circuit, as well as of other courts interpreting § 274(d), is in not looking to the relief provision in the regulations permitting allowance of a deduction in the absence of complete substantiation where the taxpayer has "substantially complied" with the "adequate records" rules.[143] Much was made of this section when the regulations were announced;[144] little has been done with it since. The Fifth Circuit obviously demands "strict," not "substantial," compliance.

The *Dowell* decision violates the basic spirit of the substantiation rules. The Congressional reports as well as the statements of the Internal Revenue Service say that no "detailed itemizations" are required. Flexibility is the keynote.[145]

Dowell is best seen as exemplifying the courts' persistent inability adequately to enforce § 274. In an effort to crack down on T & E deductions, the Fifth Circuit went too far in *Dowell* — misapplying the statute and rendering a decision unfair to the taxpayer, who only has the rules and the courts' interpretations of them on which to rely. Yet on the opposite end of the spectrum stands *LaForge*, where the Second Circuit tried to be flexible and substantially weakened the new law. Taken together, the cases demonstrate that substantiation under § 274(d) is subject to the same unpredictability and vagaries as the *Cohan* rule.

But without substantiation the door would be open to the same type of abuse that the new law was passed to combat. As long as the basic allowances for deductions are available, no satisfactory answer seems possible. The compromise struck by § 274 — keeping the deduction but tightening the procedure — is obviously failing. Is there a better way? An analysis of the handling of T & E expenses by other countries may give us an idea.

142 370 F. Supp. at 72.

143 Treas. Reg. § 1.274-5(c)(2)(v) (1962).

144 *See, e.g.*, Hannon, *supra* note 61, at 22, 24.

145 *See* Rev. Proc. 63-4, 1963-1 Cum. Bull. 474, 476 (question 15); Treas. Reg. §§ 1.274-5(b)(1), 1.274-5(b)(3) (1962); Announcement 63-8, News Release IR-574, December 28, 1962 (detailed itemizations not required). H.R. Rep., *supra* note 4, at 23; S. Rep., *supra* note 27, at 34-35.

V. The Countries Compared

A. *Great Britain*

As a general rule, § 411 of Great Britain's Income and Corporation Taxes Act of 1970 disallows any deduction for any expenses incurred in providing business entertainment.[146] As a stimulus to foreign trade, there is a very narrow exception allowing deductions for the entertainment of an "overseas customer."[147]

All gifts are treated as entertainment and thus subject to the blanket disallowance of § 411(1).[148] An exception here is made for "an article incorporating a conspicuous advertisement for the donor . . . which is not food, drink, tobacco or a token . . . exchangeable for goods."[149] Even as to these items, however, a limit is placed of two pounds per recipient per year.[150] Thus, a "business gift" in the sense of Internal Revenue Code § 274(b) is never deductible. Only small advertising items, those within our § 274(b)(1)(A) exception, are deductible at all, and then only to two pounds a year. The British code therefore takes one of the U.S. code's exceptions as the allowance and puts a statutory ceiling on it which is less than one-sixth of the limit on business gifts under § 274(b):[151] stiff treatment indeed.

---

146 Income and Corporation Taxes Act 1970 (I.C.T.A.), c. VI, § 411(1). "Business entertainment" is defined as

> entertainment (including hospitality of any kind) provided by a person, or by a member of his staff, in connection with a trade carried on by that person, but does not include anything provided by him for bona fide members of his staff unless its provision for them is incidental to its provision also for others.

I.C.T.A. § 411(5). Thus, the definition of entertainment does not include what is also excepted from the provisions of our § 274 (including the substantiation requirements) by §§ 274(e)(2), (5) relating to food, beverages, and recreational expenses for employees. The British exclusion appears to be strictly construed, however, so that if there are others enjoying the "entertainment" besides the employees, the entire deduction is lost. I.C.T.A. § 411, Commentary to Subs. 5.

147 I.C.T.A. § 411(2). The overseas customer must be "not ordinarily [a] resident of the United Kingdom." *Id.* § 411(6). Furthermore, the entertainment must be "of a kind and on a scale which is reasonable having regard to all the circumstances." *Id.* § 411(2).

148 I.C.T.A. § 411(8).

149 I.C.T.A. § 411(8)(a).

150 I.C.T.A. § 411(8)(b).

151 See note 31 *supra*.

The British tax treatment of traveling expenses demonstrates the fundamentally different concept of net income operating to disallow entertainment expenses entirely. Section 189 provides for deductibility of the necessary expenses of traveling[152] but the perception of what is "necessary" is far more stringent than the "appropriate and helpful" definition of "necessary" under our § 162.[153] Deductions are allowed under the British section only if "the holder of an office or employment is necessarily obliged to incur . . ." the expenses.[154]

The leading opinion discussing this "necessary" test describes these words as "notoriously rigid, narrow, and restricted in their operation."[155] The condition effectively means that it must be impossible to hold the office or job without incurring the expenditure, *i.e.*, that there is nothing optional on the part of the taxpayer.[156]

The severity of the rule has been acknowledged, but nevertheless applied, in several cases.[157] The effect of the rule on traveling expenses is to disallow all commuting costs, all "optional" business trips, and any combined business and pleasure trips. Only a trip devoted wholly to business and vital to that business is deductible,[158] and only expenses *ex necessitate* are allowable.[159]

In Britain, then, it's little travel and no entertainment.

---

152 I.C.T.A. § 189(1). There is an alternative test under this section which permits expenses incurred "wholly, exclusively and necessarily in the performance of . . . [one's] duties." *Id.* This basically allows for travel between two places of work in connection with the same employment, but has little general impact. *See* CCH BRITISH TAX GUIDE, INCOME TAX ¶ 14-52 (1972).

153 *See* Welch v. Helvering, 290 U.S. 111, 113 (1933).

154 I.C.T.A. § 189(1).

155 Lomax v. Newton [1953] 1 W.L.R. 1123, 1125 (Ch.). *Accord*, Sanderson v. Durbridge, [1955] 1 W.L.R. 1087, 1090-91 (Ch.); Brown v. Bullock [1961] 1 W.L.R. 1095, 1100-01 (C.A.).

156 *See* Ricketts v. Colgrahoun, 1926 A.C. 1, 7-8.

157 *See* CCH BRITISH TAX GUIDE, INCOME TAX ¶ 14-44, 45 (1974), and cases cited therein. *But see* Cretney, *Travelling Expenses*, 118 SOL. J. 838 (1974) (discussing two recent cases of the House of Lords permitting deductions); Edwards, *Travelling Expenses*, 115 SOL. J. 217 (1971). The Cretney article provides a good summary of the British treatment of traveling expenses.

158 CCH BRITISH TAX GUIDE, INCOME TAX ¶ 14-51 to -55 (1974).

159 The *ex necessitate* test was announced in Ricketts v. Colgrahoun, 1926 A.C. 1, by Lord Blanesburgh and has been the controlling standard since. CCH BRITISH TAX GUIDE, INCOME TAX ¶ 14-44 (1974).

B. *West Germany*

The very concept of the business entertainment expense in West Germany is also substantially narrower than that in the United States. This fact is best seen by analysis of the word used to denote the concept: *Gesellschaftsfreundbewirtung*, literally "business associate 'wining and dining.'" The basic root of the operative concept, *Bewirtung*, is *Wirt*, the traditional German word for host or inn-keeper. Thus *Bewirtung* does *not* include hunting lodges, country club dues, parties, or many of the other categories of entertainment with which this article has dealt.[160] Indeed, the extent of allowable entertainment in West Germany falls roughly into what is categorized under § 274 as entertainment in a "clear business setting"[161] or as one of the exceptions to the law, notably business meals.

At the present, entertainment expenses are governed by § 4 of the Income Tax Law of 1975.[162] Following the American pattern, there is a general section allowing deduction of "business . . . expenditures that are a result of the business."[163]

Section 4(5) then sets forth certain expenditures that are either not deductible or deductible only to a limited degree.[164] Specifically within these categories of nondeductability are "expenses [which] do not reduce profit" including those of subsection 2:[165]

> Expenditures for the entertainment of persons that are not employees of the taxpayer, insofar as they are unreasonable from the perspective of normal business relations or insofar as the amount of the expense or the business purpose of the expenditure is not proved. To prove either the amount of the expenditure or the business purpose of the expense, the taxpayer should make the following entries on a slip prepared by the tax authorities: day and place of entertainment, persons entertained, reason for the entertainment and the amount of expenditures. If the entertainment took place at a restaurant,

160 Commentary to the 1969 tax law, which was less strict than the new revision, expressly disallowed such activities as golfing, tennis, boating, and flying. BLÜMICH-FALK, EINKOMMENSTEUERGESETZ 434 (10th ed. 1971).

161 *See* Treas. Reg. § 1.274-2(c)(4) (1963).

162 *Einkommensteuergesetz* 1975 (EStG 1975), [1974] *Bundesgesetzblatt I* (BGBl I), 2166 § 4.

163 EStG 1975 §§ 4(1) and 4(4).

164 EStG 1975 § 4(5).

165 EStG 1975 § 4(5)(2).

> the receipt must be included as proof, complete with the signature of the restaurant owner.

This strict provision is a result of an amendment of the law in 1975.

A survey of the literature reveals that there was little specific concern over the abuse of the *Bewirtung* provision prior to 1972. However, the Finance Courts of various German states handed down decisions in this area late in 1971 expressing a new-found concern for the possible abuses of the business entertainment expense deduction. The Niedersachsen Finance Court, for example, decided that the entertainment of business associates could not give rise to any legitimate deduction when it took place in the taxpayer's own home.[166] The reason given by the court was that such an event provided no objective, reliable evidence of the propriety of the deduction or the veracity of the claimant. Similarly, the Hessian Finance Court held that expenditures for food by film producers for their colleagues and employees are not deductible business entertainment expenses because they are not separable from the daily costs of living.[167]

Under German law, as under American law, all personal expenses are nondeductible. But the German analog to § 262 goes much further, and excludes living expenses even when business related.[168] This strong presumption against the deductibility of entertainment expenses generated the decisions just discussed limiting deductions under the 1969 law.

Concern over the reliability and thoroughness of the evidence submitted to support an entertainment deduction is reflected on a broader level by a decision of the Bundesfinanzhof, which is similar to a national tax court. The court decided that the name of the entertained guests must be included on the records setting forth the occasion for the deduction.[169] Just as that of the Finance Court in Niedersachsen, the reason given was that the records must be complete so that they may be checked for reliability.

166 Judgment of October 27, 1971, 1972 FinG 233 (Finanzgericht Niedersachsen), *cited in* [1972] DEUTSCHES STEUERRECHT 472 (DStR).

167 Judgment of March 9, 1972 *cited in* [1972] DSER 573.

168 EStG § 12.

169 [1972] DEUTSCHE STEUER-ZEITUNG (Ausgabe A) 325 (DStZ), *cited in* [1972] DStR 509.

These decisions followed a period of public interest in thoroughgoing tax reform. The coalition of the Social Democratic and Free Democratic Parties that assumed control in 1970 undertook to follow up on some of its proposals, such as for sweeping tax reform, that had been asserted as campaign promises.[170]

Many argued in Germany that an overly rigorous approach could result in driving capital interests out of the country to the more favorable tax climates of the Bahamas, Panama, or even tiny Liechtenstein.[171] Despite this danger, the coalition proposals for reform included the total elimination of the entertainment expense deduction.[172] The plan was aimed at abuses of the deduction estimated at up to five hundred million DM.[173] The complaints raised echo many of those made about "expense account living" in America,[174] though the scope of the problem was much smaller in Germany.

The widespread reaction to the proposals was critical. Despite public awareness of abuses, the total elimination of business expense deductions for entertainment was seen as too blunt an instrument. One commentator saw the proposal as illogical because such expenses, when justified, were certainly expenditures that were made as a direct result of the business.[175]

The German Hotel and Restaurant Association was opposed to outright elimination and suggested that since the problem was one of abuses, an easily policed system of validation would be preferable.[176] They proposed a detailed form to be filled out after each entertainment and retained as proof of the legitimacy of the claimed deduction.[177] The provision of the 1975 law finally enacted is largely an adoption of this suggestion.[178]

Only the official government form may be used to document the

---

170 DER SPIEGEL, July 6, 1970, at 21.

171 *See* DER SPIEGEL, January 4, 1971, at 36; DER SPIEGEL, July 20, 1970, at 32.

172 *Eckwerte des 3 Steuerreformgesetzes* [1974] BLATTER FUR STEUERRECHT, SOZIALVERSICHERUNG UND ARBEITSRECHT 109.

173 DER SPIEGEL, March 18, 1974, at 30, 32.

174 See text accompanying notes 8-15 *supra*.

175 Blotekamp, *Bewirtungskosten — ein Politisches oder Steuerrechtliches Problem*, [1974] DStR 407.

176 *Id.* at 408.

177 DER SPIEGEL, March 18, 1974, at 33.

178 The German equivalent of the income tax regulations illustrates the functioning of the new tax law. The key sections provide:

expense, and thus, the German law avoids the burdensome and often unjust complexities of its American counterpart. There is no need for a court to determine what are "adequate records" or "other sufficient evidence" or "substantial compliance" or any of the plethora of other irksome questions generated by § 274(d).[179] Exceptions to the documentation requirement are minimal.[180]

Gesellschaftsfreundbewirtung represents the furthest reach of possible deductions for entertainment expenses under German law. Only limited meals and activities may be deducted, and these must be substantiated with the official form. Deductions for entertainment facilities are not allowed.[181] Even the expenses of activities in such facilities, such as hunting trips or yacht cruises, are expressly disallowed.[182] The travel expense provisions, however, are more liberal, reflecting the nonpersonal nature of these expenses.[183] Thus, West Germany has gone a long way in estab-

---

> 4.3.1. Expenditures for the entertainment of business associates in the home of the taxpayer are not considered business expenses, but rather are a part of the living expenses . . .
>
> . . . .
>
> 4.3.3. The amount and the business reason for the expenditure can be proved only by filling out a slip that is prepared by the taxing authorities completely. . . . If the slip is not filled out or filled out incompletely, the expenditures cannot be deducted, even if the taxpayer can successfully prove their amount and business purpose in another way.
>
> . . . .
>
> 4.3.7. If the entertainment takes place in a restaurant, the slip must be accompanied by the bill, signed by the owner of the restaurant or one who has his authority. The bill can accompany the slip in the following manner: both the bill and the slip can be marked so that they can be put together at any time. The bill should include the name of the restaurant and the day of the entertainment.

EStG 1975, [1974] Bundessteuerblatt I (BStBl I), at 946 et seq. (official commentary).

179 Of course, considerable discretion still exists in determining the adequacy of the information on the form.

180 These exceptions contemplate only de minimis amounts under Reg. 4.3.6 or organized events with a small, identifiable number of people under Reg. 4.3.5.

181 EStG 1975 § 4(5)(3).

182 EStG 1975 § 4(5)(4).

183 Deductions are allowed for the costs of transportation and related travel expenses for any business trip exceeding 15 kilometers from one's home. *Die Lohnslerren-Durchpuhrungsvenordunung 1975 und die Lohnsteurerliche Behandlung Von Rersersekusten bei Arbeitnehmen*, [1975] DStZ 28. However, all travel expenses must be allocated between business and personal use. If one is on a personal trip and conducts some business, no deduction for travel expenses is allowed, though specific costs related directly to the business activity are deductible. The allocation rule is narrowly construed and strictly enforced. BLÜMICH-FALK, EINKOMMENSTEUERGESETZ 419 (10th ed. 1971). Moreover, specific stautory maximums

lishing a restrictive yet workable system which could be a good model for the United States.

VI. Conclusions and Recommendations

The enactment of § 274 has accomplished the goal of eliminating the more flagrant abuses in the deduction of travel and entertainment expenses. Most would agree with this basic conclusion. What remains unclear is whether the section itself has had this effect or whether its expression of a policy position has made the tax authorities more inclined to test all such deductions more stringently. An analysis of the cases suggests the latter explanation.

The substantive rules of § 274(a) have had little impact. As we have seen, these rules really were not new in the first place. The "directly related" test seems to have been the operative standard for many entertainment deductions for decades.[184] Hardly anyone understands, and even fewer bother to apply, the alternative "associated with" test.[185] Both standards are really legislative gloss over prior judicial standards.

The one area in which the substantive rules of § 274 have been effective is in the disallowance of deductions for entertainment facility expenses for the promotion of direct goodwill.[186] As it should be, a taxpayer no longer can deduct an expense which even directly improves his reputation, tightens his relationship with existing clients, or introduces him to potential clients. The "soft-sell" in an entertainment setting is out.[187]

---

have been established for travel deductions, up to a maximum of 47 DM per day (the equivalent of about 19 dollars). [1975] DStZ, at 29.

Unlike the law of the United States, however, the German law allows a taxpayer a deduction for commuting expenses to and from his place of employment. These expenses do not have to be proved directly, and are allowed for one trip each way per day. *See* C. Bottcher, Grosskommentator zur Einkommensteu-everr (Forkel-Verlag 1975).

184 *See* Rephan, *1962 Act: Is the "Directly Related" Test for Entertainment Expenses Really New?*, 17 J. Taxation 366 (1962), and text accompanying notes 77-86 *supra*.

185 See text accompanying notes 49-51 *supra*.

186 See text accompanying notes 87-97 *supra*.

187 *See* Hippodrome Oldsmobile, Inc. v. United States, 474 F.2d 959 (6th Cir. 1973), discussed in text accompanying notes 94-97 *supra*.

Most of what impact the new law has had, though, stems from the demanding substantiation requirements of § 274(d). These often are enforced with a vengeance.[188] They also open the door to arbitrariness, inconsistency, and unfairness in the application of our tax laws. While giving the courts a basis for disallowances, they result in trapping some honest but unwary taxpayers, while permitting others to abuse the system and deduct clearly personal expenses. This result seems inevitable under a structure which continues to allow deduction of mixed business and personal expenses, such as entertainment.

Abolition of the *Cohan* rule not only means that all claims must be documented but has resulted in an end to the excessively generous allocations based on estimates. A positive result has been seen from this change. Gone are the days when a taxpayer could expect "at least a third" of the expenses of a hunting lodge or social party to be picked up by the government. The cases interpreting § 274(d), however, are just as confused as those under the *Cohan* rule. The main difference today is that without allocation it's all or nothing.

These three main effective elements of § 274 — the disallowance of goodwill entertaining, substantiation, and the end to allocation — have had their greatest effect on the category of entertainment facilities. The tax-supported boats and hunting lodges are gone. Section 274, however, has had much less impact on club dues and general entertainment.

The section fails to confront the primary problem with entertainment and travel expense deductions. These activities inherently involve large elements of personal pleasure and benefit, often predominating over business purposes. In a theoretically perfect tax system, allocation between these elements should occur — limiting deductions to business expenses properly allowable under a concept of net business income. But as the failures of §§ 162 and 274 amply demonstrate, our tax system is far from perfect. Allowances of deductions for travel and entertainment inevitably lead to arbitrary enforcement and confusing case law.

With abuses and arbitrariness throughout the current system, it is remarkable that little attention has been focused on this problem

---

188 See discussion of the *Dowell* case at text accompanying notes 128-45 *supra*.

in the past few years. The bills introduced in the Ninety-fourth Congress primarily have reflected the post-Watergate concern with curbing the side benefits to government officials.[189] One Senate proposal, however, would deny a deduction for first class air travel and for any foreign conventions.[190] A House proposal attempts to set guidelines which would limit deductions for attendance of conventions.[191] While sound proposals, they barely scratch the surface of the T & E problem.

The original proposals of President Kennedy offer the soundest basis for change.[192] No deductions for entertainment facilities should be allowed. Beyond this, the German concept of *Bewirtung* can be used to define the limits of what activities should be allowable.[193] Only those which are certain to be for business reasons can

189 One proposal would limit the amount a government official could be reimbursed by a private foundation for attending a convention to $1000 plus 125% of the government's per diem allowance. H.R. 2984, 94th Cong., 1st Sess. (1975). Another bill would generally put a ceiling on the amount legislators could deduct as business expenses. H. R. 10612, 94th Cong., 1st Sess. (1975). Other than these bills and those related to first class travel and foreign conventions, no bills have been introduced or hearings held this year before the House Ways and Means Committee regarding T & E deductions. Telephone Interview with Janice Mays, Legislative Assistant, House Ways and Means Committee, March 10, 1976.

190 Senator Edward Kennedy (D. Mass.) proposed to the Senate Finance Committee on March 18, 1976 a denial of any deduction for first class air travel. 122 CONG. REC., 3755, 3757 (daily ed. Mar. 18, 1976) (Senator Kennedy's Proposals For Tax Reform).

191 Proposed § 274(h), now before the House Ways and Means Committee, would disallow deductions for attendance at more than two foreign conventions. Tax Reform Act of 1975, H.R. 10612, 94th Cong., 1st Sess. § 605. The bill also attempts to set guidelines for the allocation between business and personal expenses. Allowances for traveling expenses are *limited* to the federal per diem rates. Full transportation costs continue to be deductible if the individual spends more than one-half of the time for business purposes, but these will be limited to coach fares only.

The attempt to quantify time spent on business as compared to personal activities so as to end arbitrary and differing results for taxpayers in similar situations represents progress just as do the per diem and coach fare limitations. However, the resort lobby seems to have won again in that the test is not the reasonable one of a location consistent with the territorial membership scope or special purposes of the organization holding the convention, but instead is domestic versus foreign travel.

While the proposed addition to § 274 moves in the right direction, limiting expenses to per diem rates and eliminating the first class travel altogether, it does so in the wrong way. Further complicated tests are added, and the entire section is only relevant to foreign conventions—a very limited category especially since enactment of § 274(c). Arbitrary lines are drawn rather than any attempt at redefinitions.

192 See text accompanying notes 16-24 *supra*.

193 See text accompanying notes 160-67 *supra*.

escape the present inequities and problems of allowing mixed personal and business expense deductions.

"Entertainment" as it is presently used and understood would be nondeductible in any form.[194] Certain "social activity expenses" in the form of the present exceptions would be allowable.[195] This change in terminology is a keynote to a basic change in the tax system, where no longer will personal expenses be subsidized. The tax system will become simpler and more easily administrable. The economically wasteful practice of entertaining as a form of advertising and promotion will be curbed.

Substantiation will not be the boon-doggle it currently is. Very few deductions will be allowed in the first place, and all of these are well defined. Moreover, once the legislature declares that mixed personal and business entertainment expenses are no longer deductible, the courts will be freed from the dilemma of the past — allowing the expenses in theory, but disallowing them in practice. Theory and practice will coincide.

Basic statutory maximums should be set in each category of allowable "social activity expenses." Thus, a business meal allowance of ten dollars for each guest is granted without the need for documentation. Pursuant to the theory of this structure to disallow personal expenses, the *Sutter* rule must be resurrected.[196] No deductions will be permitted for one's own meals or recreation. Beyond the individual maximums in each category, an absolute overall maximum deduction for all "social activity expenses" must be set to eliminate the possibility of abuse of this system.[197] Travel

194 "Entertainment expenses" as such should be disallowed completely under § 162, thus making it clear that such disallowance goes to the very nature of a definition of income.

195 Most of these relate to social and recreational activities within the employment setting, such as those allowed under present exceptions §§ 274(e)(2), (5). The only activity which could qualify as entertainment as such would be what is presently allowed under the "business meal" exception or the provision relating to a "clear business setting" allowing a hospitality room or a gala opening party. *See* Treas. Reg. §§ 1.274-2(c)(4), (f)(2)(i) (1963).

196 See text accompanying notes 104-05 *supra*.

197 There should be a base rate set by the regulations, which is adjustable for inflation, plus a percentage of gross income. There should not be a fear of favoring big companies here, because this provision would serve only as a fallback to prevent use of the new section as a cover for newly disallowed entertainment or other expenses. The suggested limit is closely analogous to that for charitable

expenses would be subject to allocation and statutory maximum deductions.[198]

More fundamentally, our concept of net income needs redefining. People should not pay for other people's entertainment. Business should not waste its money and its customers' time on entertainment which at most has a tenuous connection with productivity. Under § 274, allocation is out and substantiation is in — and that's about all. In an age in which we are witnessing the derision and decline of bigness — big cars, big cities, big government, and Big Nurse[199] — a tax policy recognizing this reality is needed. As the song says, the party's over.

---

deductions of corporations, which is rarely even approached in fact. *See* INT. REV. CODE OF 1954, § 170(b)(2).

This is also similar to the treatment of entertainment expenses in Japan, where the total amount deductible is based on the paid-in capital of a corporation. Tax Special Measures Law (Law No. 26 of 1957). In Japan, however, this quantitative limit is used in lieu of tight nexus tests for deductibility.

198 "First class travel" should not be deductible. The maximum allowable transportation costs would be the regular coach air fares for such travel; meals and lodging should be limited to the federal government per diem rate without documentation and twice this rate with documentation and proof as to the business necessity of exceeding the basic per diem allowance. An amount equal to a percentage of the meals and lodging per diem rate, perhaps 25 percent, should be set for all incidentals.

199 *See* K. KESEY, ONE FLEW OVER THE CUCKOO'S NEST (1962); Statute, *Beyond the "Cuckoo's Nest": A Proposal for Federal Regulation of Psychosurgery*, 12 HARV. J. LEGIS. 610 (1975).

BOOK REVIEW

THE INTERPRETATION AND APPLICATION OF STATUTES. By *Reed Dickerson*, Boston: Little, Brown & Co., 1975. Pp. xx, 309, indices. $16.50. LEGISLATIVE ANALYSIS: HOW TO USE STATUTES AND REGULATIONS. By *William P. Statsky*, St. Paul: West Publishing Co., 1975. Pp. 216, index. $7.50.

*Reviewed by Dennis J. Owens**

Reed Dickerson has long had the subject of legislative drafting to himself. Once, in response to an inquiry as to what materials were used by the Office of Legislative Counsel of the U.S. House of Representatives, Mr. Edward O. Craft, the retired Director of that office, simply said, "[W]e give each new member of Congress a copy of the Dickerson book."[1] No further identification of the book[2] was given — or needed — then. Such brevity no longer suffices, however, for Dickerson has now produced another important book.

His new work is ambitious. It seeks to set forth "a coherent account of 'statutory interpretation' that clarifies the great bulk of what has been taking place under that banner" (p. xvii). That is, it attempts to give us a theory or philosophy of statutory interpretation. To the extent to which it is successful, the fulfillment of its ambition is rooted in Dickerson's grasp of the dual nature of the term "statutory interpretation." The first aspect of its meaning is epistemological, the second judicial; Dickerson uses the terms "cognition" and "creation" respectively.

In explicating "cognition," Dickerson is at his most scholarly. The meaning of "meaning" may seem rather removed from lawyering or even from legislating. However, the author deftly brings us through the journey from semantics to law by keeping our eyes focused on the ultimate destination: legislation and its meaning. To those specializing in linguistics, this treatment is more a rehash of the law review literature that has developed around the

* Member of the Missouri Bar; J.D. 1975, University of Notre Dame.

1 Letter from Edward O. Craft to Dennis J. Owens, January 12, 1974.

2 R. DICKERSON, LEGISLATIVE DRAFTING (1954).

work of Ogden and Richards[3] than it is a new analysis. But given that this book is written for lawyers, this is not a fault; the overview provides the necessary background for what follows.

Faithful readers of the HARVARD JOURNAL ON LEGISLATION will recognize Chapter 5 of Dickerson's new work as an article from the JOURNAL's first volume.[4] This is probably the best part of the book; it addresses a practical legal problem while simultaneously serving as a mind-stretching exercise in theory. The next chapter[5] extends this treatment into the arena of the jurisprudes, where it unfortunately seems to be somewhat sterile. The difference stems from the fact that the first of these two chapters is eminently useful as a lawyer's tool, while the second is more a recounting of a philosophical dispute. This is not to disparage jurisprudence; it is only that in Chapter 5 the question takes on more of the concreteness lawyers prefer in their approach to issues.

Llewellyn once gave us a list of maxims about statutory construction — and immediately matched it with another list of maxims which contradicted the first.[6] The reviewer's experience is that the one canon which Llewellyn did not set up and then knock down was that courts are to give expression to the legislature's intent and purpose. The commentators make much of the differences between intent and purpose,[7] but not so the courts. Legal scholars disfavor "intent" and favor "purpose."[8] Radin made a tremendous attempt to destroy the concept of "legislative intent" some forty-six years ago,[9] but the courts seem to have ignored the attack and the concept is still a vital one.

What do the terms "intent," "purpose," and "motive" mean when ascribed to a legislature? In keeping with a lawyerly preference for concreteness, my answer must take the form of an

3 C. OGDEN & I. RICHARDS, THE MEANING OF MEANING (10th ed. 1956).

4 Dickerson, *The Diseases of Legislative Language,* 1 HARV. J. LEGIS. 5 (1964).

5 Chapter 6, *Basic Concepts: Legislative Certainty and the Purposive Elements,* pp. 54-66.

6 Llewellyn, *Remarks on the Theory of Appellate Decision and the Rules or Canons About How Statutes Are To Be Construed,* 3 VAND. L. REV. 395, 401 (1950).

7 *See, e.g.,* Corry, *Administrative Law and the Interpretation of Statutes,* 1 U. TORONTO L.J. 286, 290 (1936).

8 Thomas, *Statutory Construction When Legislation Is Viewed as a Legal Institution,* 3 HARV. J. LEGIS. 191, 201, 204 (1966).

9 Radin, *Statutory Interpretation,* 43 HARV. L. REV. 863 (1930).

example.[10] A selective service registrant may not burn his registration card.[11] The statute which prohibits this may be said to have an "intent" to stop the destruction of draft cards and a "purpose" of promoting the efficiency of the selective service system.[12] Congress' "motive" probably was to put the screws on war protesters because they had won so much attention with the dramatic burnings of their cards. Courts say that they do not look to motive; it is well that they do not, since the motive I would assign in the example would have been a suppression of symbolic speech. The trouble with the championing of "purpose" is that it presents problems similar to those surrounding "intent." Radin's attack could be equally well applied; for instance, how do you ascertain "purpose"?[13]

The answer to that question is usually given as "look at the legislative history." The wisdom of this advice is debatable, and as a practical matter, it is often impossible to follow. In Missouri, for instance, there is no legislative history preserved.[14] Committee reports are not published and there is no reason that they should be. They simply do not provide any useful information ("The Committee recommends enactment of the following bills: . . . ." Period.). But, this gap causes no appreciable loss in the quality of appellate construction of the state's statutes. The judicial response to the situation is to heed Frankfurter's threefold imperative to law students: "(1) Read the statute; (2) read the statute; (3) read the statute!"[15] The courts refer to any contrasting language both in laws enacted and in laws repealed simultaneously with the relevant statute's enactment. "The statute is read as a

---

10 Dickerson distinguishes the two as follows:

> . . . the word "intent" coincides with the particular immediate purpose that the statute is intended to directly express and immediately accomplish, whereas the word "purpose" refers primarily to an ulterior purpose that the legislature intends the statute to accomplish or help to accomplish (p. 88).

11 50 U.S.C. § 462(b)(3) (1967).

12 *See* United States v. O'Brien, 391 U.S. 367 (1968); O'Brien v. United States, 376 F.2d 538 (1st Cir.), *cert. denied,* 389 U.S. 814 (1967); United States v. Cooper, 279 F. Supp. 253 (1968).

13 Gibson, *Literary Minds and Judicial Style,* 36 N.Y.U.L. REV. 915, 920 (1961).

14 Woody, *Statutory Interpretation — The Need for Improved Legislative Records in Missouri,* 38 Mo. L. REV. 84 (1973). The same situation prevails in most states.

15 H. FRIENDLY, BENCHMARKS 202 (1967).

whole," meaning that note is taken of the entire chapter or section ("context") of the Revised Statutes, as is appropriate. Many decisions attempt to determine "intent," although they often mean "purpose" as defined by Dickerson.

Justice Jackson once sarcastically referred to the study of legislative history as "psychoanalysis of Congress."[16] He was of the view that the history in the case at hand, consisting only of a House committee hearing transcript, was even "more vague than the statute [the Court was] called upon to interpret."[17] Jackson overstated his case, Dickerson suggests, and maybe so. But he had a point: read the statute.

Possibly the most troublesome things about Dickerson's book are his four "constitutional assumptions" (Chapter 2, pp. 7-12). They are troublesome because they are dubious overstatements. The first assumption is "legislative supremacy." Although legislatures may be the sole source of *statutes,* administrative agencies make law with their regulations, courts with their rules and decisions, and executive departments with their day-to-day application of all of the above. Dickerson's familiarity with the British system, where Parliament is supreme, may have misled him.[18] The use of the term "supremacy" in this assumption evokes the title of Justice Jackson's *The Struggle for Judicial Supremacy* (1941), a book of great political sophistication and legal insight. Statutes only come alive when they are enforced, and the courts are supreme in the field of the enforcement of laws. But, in spite of Dickerson's averments to the contrary, this assumption seems to affect very little which follows in the book. In other words, he is wrong, but so what?

The second assumption is that a statute is made only by proper enactment. I fail to see why this obvious statement need be made. A court could use legislative history to negate the express words of a statute, I suppose, but this would be so plainly improper as to not even call for analytic attack. The third assumption is that

---

16 United States v. Public Util. Comm'n, 345 U.S. 295, 319 (1953) (Jackson, J., concurring).

17 *Id.* at 320.

18 *But cf.* Dickerson, *Legislative Drafting in London and in Washington,* 1959 CAMB. L.J. 49.

statutes are written in ordinary English. Well, of course, they are not; they are written in Legislative English, a dialect of Lawyer's Legal English.[19] Dickerson is making this assumption because he wants to venture into linguistic analysis, and evidently feels the assumption is necessary to set the stage. But how and why can such an assumption be constitutional? Maybe it is in the penumbra of a penumbra; more realistically, it is a common sense assumption, even if it is not entirely true.

The fourth assumption is that legislatures are required to make their statutes reasonably available to citizens. While this is an excellent idea and to be strongly recommended, it is at sad variance with the facts. Nothing in the U.S. Constitution or most state constitutions requires this, although the courts may take a dim view of an "unavailable" statute. In the reviewer's experience, many counties have *never* codified or compiled their ordinances[20] and yet no ordinance has been nullified because of its unavailability. There is a real due process problem in such a situation, but it is a situation which exists in spite of Dickerson's assumption. Again, though, the validity or invalidity of the assumption has little effect on the basic themes of the book.

In short, Dickerson's new book serves as an excellent account of the more rational work in the field of statutory interpretation. It is probably impossible to construct a true jurisprudence of statutory interpretation, but Dickerson is successful in providing us with a systematic, consistent approach to the subject. This treatise's faults are a product of Dickerson's total immersion into the field. Law is more than statutes.

Statsky's new book, *Legislative Analysis: How To Use Statutes and Regulations,* stands in sharp contrast to the theoretical and philosophical work of Dickerson. The essence of this book is reflected in the words "how to" in the title. The preface suggests that the book can be used by a large "number of individuals:

---

19 *See* T. SHAFFER, THE PLANNING AND DRAFTING OF WILLS AND TRUSTS 28-32 (1972).

20 *See, e.g.,* the ordinances of St. Joseph County, Indiana. Ordinances from 1820 to 1899 were recorded only in longhand. The only way to locate the law on a given subject is to provide the county clerk with the date of its passage. There is no way to find out the dates of passage by reference to an index, let alone to a compiled or codified set of ordinances.

paralegals, employees of administrative agencies, staff on legislative committees, legal secretaries, students studying to be lawyers, etc." Notice that lawyers themselves are not included. This exclusion is fully intentional, as evidenced by the lack of Lawyer's English and the extremely unpresumptive starting point: "At the federal level of government, the legislature is Congress" (page 1). To the lawyers amongst the readers of this review, I beseech you: hang in there. It gets better.

By "better," I mean more useful. Lawyers have a tendency to ignore statutes. (As Judge Frank said, we have it all to blame on a "brilliant neurotic.")[21] Statsky did not have ambitions of studying and evaluating each tool of statutory interpretation; he wished only to train persons in their correct use. His most laudable innovation is tying the development of skills in drafting statutes to the development of skills in interpreting them. Statsky teaches at Antioch Law School in Washington, D.C., the hotbed of clinical legal education; this is reflected in the approach taken here. The heart of clinical training is to learn by doing; Statsky's book is replete with realistic exercises.

The book will probably be most valuable to lawyers who never really paid attention to statutes during law school. That is, by most lawyers.

---

21 J. FRANK, COURTS ON TRIAL: MYTH AND REALITY IN AMERICAN JUSTICE (1949). The neurotic, of course, was C. C. Langdell who, as Dean of Harvard Law School, gave us the case method of legal education.

RECENT PUBLICATIONS

PRISONS: HOUSES OF DARKNESS. By *Leonard Orland,* New York: The Free Press, 1975. Pp. xv, 224, appendices, index. $10.00.

This book sharply criticizes the efficacy and fairness of existing sentencing and parole processes, and proposes two alternative, though not mutually exclusive, approaches to restructuring the present systems. Orland's "radical approach" calls for the elimination of the parole system and indeterminate sentences, and the imposition of short (no more than five years) fixed-term sentences. The "reformist approach" would leave the existing structure intact, but would open the sentencing process to judicial scrutiny in a tiered system designed to avoid arbitrariness and to promote evenhandedness.

The book is provocative, well-footnoted, and it provides an excellent history of our present penal system. For those who are versed in the subject of penal reform, this book is must reading; for those who are not, it provides an excellent introduction to the topic.

PUBLIC EMPLOYEE PENSION FUNDS. By *Robert Tilove,* New York: Columbia University Press, 1976. Pp. 370, index. $20.00.

This Twentieth Century Fund Report provides a detailed examination of public employee retirement plans, with a special emphasis on the way these plans mesh (or don't mesh) with Social Security coverage. The book criticizes the "loophole" that allows public employee groups to drop out of the Social Security program after benefits have vested, and calls for mandatory imposition of Social Security taxes on all public employees. Tilove also looks at the way these plans are funded (he says they should be funded on an actuarial basis, rather than pay-as-you-go), and argues for liberalized vesting requirements to increase employee mobility.

The book is filled with interesting statistics and tables, and the general contentions made are illustrated by specific references to New York State, Massachusetts, and Illinois. This comprehensive work will prove invaluable to anyone in a policy-making role with regard to public employee retirement plans.

THE INSANITY DEFENSE: A BLUEPRINT FOR LEGISLATIVE REFORM. By *Grant H. Morris,* Lexington, Ma: Lexington Books, 1975. Pp. xv, 133, appendices, index. $10.00.

This study provides an excellent summary of the debate between those who wish to abolish the insanity defense and those who feel it should be retained and applied in a more enlightened manner than is presently the case. Morris carefully describes and considers the various insanity tests and the doctrine of "diminished responsibility." His main thrust is that the insanity defense as now used is virtually worthless, since successfully pleading the defense generally results in indefinite confinement in a mental institution, where the individual is viewed as a sentence-serving convict anyway. The appendices provide excellent analyses of how the various states stand with respect to key insanity defense issues such as burden of proof and the insanity test to be applied.

LEGISLATIVE LAW AND PROCESS IN A NUTSHELL. By *Jack Davies,* St. Paul: West Publishing Co., 1975. Pp. xxi, 279, index. $6.00 paper.

The book contains an extensive discussion of legislative power in general, and as modified and constrained by judicial power, but the heart of the book is a nuts-and-bolts look at the tactics, techniques, and procedures necessary to draft a bill, get it introduced in the legislature, and push it through committee to a vote on the floor. Anyone involved in the legislative process who doesn't know the difference between a "hairy arm" (p. 71) and a "foot-in-the-door" (p. 59) would undoubtedly profit from reading the whole of Davies' book.

PRESIDENTIAL POWER: THE POLITICS OF LEADERSHIP WITH REFLECTIONS ON JOHNSON AND NIXON. By *Richard E. Neustadt,* New York: John Wiley & Sons, Inc., 1976. Pp. 324, index. $10.95.

This update of Neustadt's 1960 classic is especially timely in this election year. One of Neustadt's themes is that the man makes the office, as well as vice-versa; his thoughts on the differences between the Truman and Nixon White Houses make it clear that a candidate's personality is highly relevant to the way he will be able to function in the Oval Office. Neustadt maintains that the general character of presidential power hasn't changed very much in the fifteen years since *Presidential Power* first appeared, but he does highlight six developments which have caused him to modify his original thoughts. His views on the bloated White House staff, which reached a peak (or valley, depending on your point of view) under Nixon are most interesting; one hopes they will be read by the victor in November.

The new material in the book is unfortunately brief; Neustadt leaves the reader yearning for more. Of course, for those who haven't yet read the original *Presidential Power,* it is all here, as good as ever.