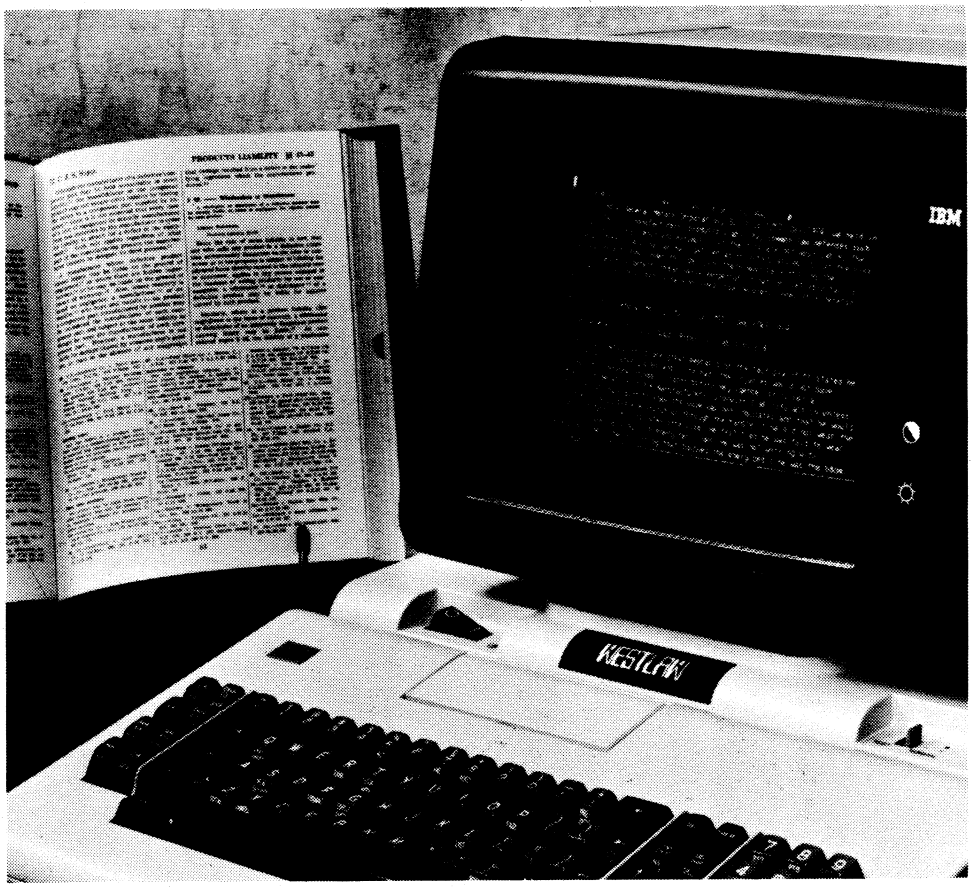
VOLUME 1982 NUMBER 4

UTAH LAW REVIEW

UNIVERSITY OF UTAH COLLEGE OF LAW



WESTLAW

A NEW DIMENSION TO YOUR WEST LIBRARY SYSTEM

When you add WESTLAW to your library, you are giving yourself much more than the speed, flexibility and efficiency of electronic research.

You are adding a new and important law finding tool that is totally interrelated with your basic West library. You are adding a new dimension, a new resource to your library system. WESTLAW works with your books to make your entire library more efficient and more productive.

WESTLAW is designed by West Publishing Co., a company which has served the legal profession for more than 100 years. That is part of the WESTLAW advantage.

West is the company which can offer you electronic research—at the right price—along with a complete and coordinated system of trusted law publications to meet your every research need.

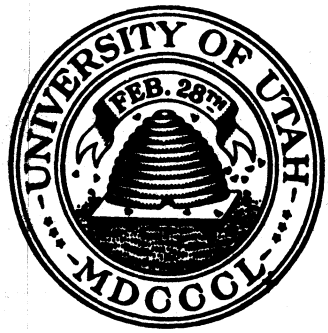
WESTLAW is also easy to add to your library because WESTLAW is compatible with many Word Processing Systems and a variety of video display terminals. You can choose the equipment that's best for you.

Find out how your firm can make WESTLAW an integral part of your total library system. Call us collect today, at 612/228-2536. Or write to WESTLAW, West Publishing Company, P.O. Box 3526, St. Paul, MN 55165

AND ANY FIRM CAN AFFORD IT

UTAH LAW REVIEW

VOLUME 1982 NUMBER 4



EDITORIAL OFFICE: The *Utah Law Review* is published at the University of Utah College of Law, 325 University Street, Salt Lake City, Utah 84112 by the Utah Law Review Society. Second-class postage paid at Salt Lake City, Utah, and at additional mailing offices. All communications should be sent to the editorial offices. Postmaster: Please send form 3579 (change of address) to *Utah Law Review*, University of Utah College of Law, Salt Lake City, Utah 84112.

SUBSCRIPTIONS: The *Utah Law Review* is published quarterly. The subscription rate is $18 for one year, $32 for two years, and $45 for three years. Single issues may be obtained for $5.00 per issue and reprints of articles and notes for $2.50 per copy. If the subscriber wishes his subscription discontinued at its expiration, notice should be sent; otherwise, it is assumed that continuation is desired.

Printed in the United States of America
by Darby Printing Company

UNIVERSITY OF UTAH
COLLEGE OF LAW FACULTY & ADMINISTRATION

AARON, RICHARD I. (1966), Professor of Law. A.B., 1959, Harvard University; J.D., 1962, University of Wisconsin.

ANDERSON, FREDERICK R. (1980), Professor of Law. B.A., 1963, University of North Carolina; B.A. in Jurisprudence, 1965, Oxford University; J.D., 1968, Harvard University.

BENNETT, WALLACE R. (1953), Professor of Law. B.A., 1949, M.S., 1950, J.D., 1953, University of Utah; LL.M., 1957, Harvard University; LL.M. (Trade Regulations), 1963, New York University.

BOYCE, RONALD N. (1966), Professor of Law. B.S.L., 1955, J.D., 1957, University of Utah.

CHANCELLOR, THOMAS H. (1982), Associate Professor of Law. B.A., 1961, North Texas State University; LL.B., 1964, Harvard University.

CORRY, JAQUITA W. (1974), Assistant Dean and Director of Placement, Development and Alumni Relations.

DYER, BOYD K. (1971), Professor of Law. B.A., 1962, Stanford University; LL.B., 1968, Harvard University.

EMERY, ALFRED C. (1947), Professor of Law. B.S., 1940, J.D., 1947, University of Utah.

FIRMAGE, EDWIN B. (1966), Professor of Law. B.S., 1960, M.S., 1962, Brigham Young University; J.D., 1963, LL.M., 1964, S.J.D., 1964, University of Chicago.

FLYNN, JOHN JOSEPH (1963), Professor of Law. B.S., 1958, Boston College; LL.B., 1961, Georgetown University; S.J.D. 1967, University of Michigan.

FORDHAM, JEFFERSON B. (1970), Distinguished Professor of Law, former Dean of the University of Pennsylvania Law School. A.B., 1926, M.A., 1929, J.D., 1929, University of North Carolina; S.J.D., 1930, Yale University; LL.D., 1953, University of North Carolina; L.H.D., 1970, University of Pennsylvania.

FRANCIS, LESLIE PICKERING (1982), Associate Professor of Law and Assistant Professor of Philosophy. B.A., 1967, Wellesley College; Ph. D., 1974, University of Michigan; J.D., 1981, University of Utah.

FRANKEL, LIONEL H. (1966), Professor of Law. B.A., 1953, Ursinus College; J.D., 1956, Yale University; LL.M., 1962, New York University.

HARRIS, LESLIE J. (1978), Associate Professor of Law. B.A., 1973, New Mexico State; J.D., 1976, University of New Mexico.

JENSEN, A. LADRU (1926-1964), Professor Emeritus of Law. A.B., 1917, Brigham Young University; M.A., 1924, J.D., 1925, University of California.

LOCKHART, WILLIAM J. (1964), Professor of Law. B.A., 1955, J.D., 1961, University of Minnesota.

LUND, THOMAS A. (1978), Professor of Law. A.B., 1964, Harvard University; LL.B., 1967, Columbia University; D. Phil., 1978, Oxford University.

MORRIS, JOHN K. (1979), Associate Professor of Law. B.A., 1966, University of California at Los Angeles; J.D., 1969, University of California at Berkeley.

McCORMACK, WAYNE (1978), Professor of Law. B.A., 1966, Stanford University; J.D., 1969, University of Texas.

OBERER, WALTER E. (1975), Dean of the College of Law and Professor of Law. B.A.,

1942, Ohio Wesleyan; J.D., 1948, Harvard University.

OLSON, JEFFERY E. (1977), Associate Dean of the College of Law. B.A., 1975, J.D., 1977, University of Utah.

PALACIOS, VICTORIA J. (1977), Associate Professor of Law and Director of Special Programs. J.D., 1975, University of Nebraska.

SCHMID, ROBERT L. (1954), Professor of Law. J.D., 1954, University of Utah; M.L.L., 1956, University of Washington.

SPURGEON, EDWARD D. (1980), Associate Dean of the College of Law and Professor of Law. A.B., 1961, Princeton University; LL.B., 1964, Stanford University; LL.M., 1968, New York University.

STRACHAN, KRISTINE (1973), Professor of Law. B.F.S., 1965, University of Southern California; J.D., 1968, University of California at Berkeley.

SWENSON, ROBERT W. (1953), James I. Farr Professor of Law. B.S.L., 1940, J.D., 1942, University of Minnesota.

THURMAN, SAMUEL D. (1962), Distinguished Professor of Law, former Dean of the University of Utah College of Law. A.B., 1935, University of Utah; J.D., 1939, Stanford University.

VAN ALSTYNE, ARVO (1966), Professor of Law. B.A., 1943, J.D., 1948, Yale University.

WILKINS, RUTH W. (1966), Professor of Law and Law Librarian. J.D., 1953, University of Utah.

ZILLMAN, DONALD N. (1979), Professor of Law. B.S., 1966, J.D., 1969, University of Wisconsin; LL.M., 1973, University of Virginia.

UTAH LAW REVIEW

VOLUME 1982 NUMBER 4

UTAH LAW REVIEW

VOLUME 1982 NUMBER 4

TABLE OF CONTENTS

ARTICLES

COMMENTS

Commercial Entities' Noncommercial Speech: A Contradiction in Terms

*Richard M. Alderman**

I. INTRODUCTION

The United States Supreme Court has recently extended first amendment protections to purely "commercial" speech.[1] It has done so, however, with qualification. Commercial speech is recognized as having lower first amendment protection than traditionally protected forms of expression.[2] Consequently, a new dichotomy has developed between the first amendment protections afforded noncommercial speech and the "hybrid" protection afforded commercial speech.[3] Although the classification of speech as "commercial" once used to preclude constitutional review, it now dictates a different, though substantial measure of first amendment protection.[4]

---

* Associate Professor of Law, University of Houston Law Center. B.A., 1968, Tulane University; J.D., 1971, Syracuse University; LL.M., 1973, University of Virginia. The author wishes to thank Carole Reed and Steven Kearney for their assistance with the preparation of this article.

1. Virginia State Bd. of Pharmacy v. Virginia Citizens Consumer Council, 425 U.S. 784 (1976). The Court has also held that the corporate nature of the speaker does not deprive the speech of first amendment protection. First Nat'l Bank v. Bellotti, 435 U.S. 765 (1978).

Since the decision in *Virginia State Bd. of Pharmacy*, the Court has applied the first amendment to commercial speech and has prohibited the banning of "For Sale" signs, Linmark Assocs. v. Township of Willingboro, 431 U.S. 85 (1977); the disciplining of lawyers who advertise, *In re* R.M.J., 455 U.S. 191 (1982); Bates v. State Bar of Ariz., 433 U.S. 350 (1977); and the prohibition of promotional advertising by a public utility, Central Hudson Gas & Elec. Corp. v. Public Serv. Comm'n, 447 U.S. 557 (1980).

2. *See* Virginia State Bd. of Pharmacy v. Virginia Citizens Consumer Council, 425 U.S. 748 (1976). In *Virginia State Bd. of Pharmacy*, the Court stated: "In concluding that commercial speech enjoys first amendment protection, we have not held that it is wholly undifferentiable from other forms." *Id.* at 771 n.24. The Court further observed that the common sense differences between commercial and noncommercial speech "suggest . . . a different degree of protection". *Id.* Commercial speech cases to date have reiterated this point. Central Hudson Gas & Elec. Corp. v. Public Serv. Comm'n, 447 U.S. 557, 562-63 (1980); *see In re* R.M.J., 455 U.S. 191, 199-204 (1982); Friedman v. Rogers, 440 U.S. 1, 11 n.9 (1979) (quoting Ohralik v. Ohio State Bar Ass'n, 436 U.S. 447, 456 (1978)).

3. *See infra* notes 34-40 and accompanying text.

4. Commentators have extensively discussed the implications of the Court's decision to extend substantial first amendment protections to commercial speech. *See, e.g.*, Brudney, *Business Corporations and Stockholders' Rights Under the First Amendment*, 91 YALE L.J. 235 (1981) ("Serious doubts exist regarding the validity of constitutional support thus given [corporate political speech]"); Emerson, *First Amendment Doctrine and the Burger Court*,

In granting significant first amendment protection to commercial speech, the Supreme Court utilized the definition of commercial speech developed in earlier cases. Traditionally, commercial speech was narrowly defined for purposes of first amendment analysis in order to avoid the harsh consequences that followed that classificiation.[5] However, it is no longer necessary to so delimit the term "commercial speech." In light of its present treatment by the Court, commercial speech should be defined broadly as "any and all speech of a commercial entity." Only through such an inclusive definition can the "new" commercial speech doctrine be effectively applied. Continuing to use the traditional definitional approach can only lead to the ineffective allocation of judicial resources and a dilution of traditional first amendment protection.

II. THE COMMERCIAL SPEECH DOCTRINE: ADVENT, DEMISE AND RESURRECTION

Historically, the Supreme Court's reaction to commercial speech was to exclude it from the coverage normally associated with freedom of expression. In *Valentine v. Chrestensen*,[6] the Court held that although state and city authorities could not un-

---

68 CALIF. L. REV. 422, 458-61 (1980) ("There is serious question, however, whether the gain [of *Virginia State Bd. of Pharmacy*], if any, is worth the cost"); Jackson & Jeffries, *Commercial Speech: Economic Due Process and the First Amendment*, 65 VA. L. REV. 1, 5 (1979) ("We believe that *Virginia State Bd. of Pharmacy* was decided wrongly"); Prentice, Consolidated Edison *and* Bellotti: *First Amendment Protection of Corporate Political Speech*, 16 TULSA L.J. 599, 601 (1981) ("extension of First Amendment protection to corporate political speech, although not immune from criticism, is a positive development which should be applauded"); Note, *The Corporation and the Constitution: Economic Due Process and Corporate Speech*, 90 YALE L.J. 1833 (1981) (concluding that the decision in First Nat'l Bank v. Bellotti, 435 U.S. 765 (1978), was unwise and should be overturned). *See also* Alexander, *Commercial Speech and First Amendment Theory: A Critical Exchange*, 75 NW. U.L. REV. 307 (1980); Baker, *Commercial Speech: A Problem in the Theory of Freedom*, 62 IOWA L. REV. 1 (1976); Farber, *Commercial Speech and First Amendment Theory*, 74 NW. U.L. REV. 372 (1979); O'Kelley, *The Constitutional Rights of Corporations Revisited: Social and Political Expression and the Corporation After* First National Bank v. Bellotti, 67 GEO. L.J. 1347 (1979); Rotunda, *The Commercial Speech Doctrine in the Supreme Court*, 1976 U. ILL. L.F. 1080; Comment, *First Amendment Protection for Commercial Advertising: The New Constitutional Doctrine*, 44 U. CHI. L. REV. 205 (1976).

5. Under the "commercial speech doctrine," commercial speech was considered outside of the scope of first amendment protections. *See* Valentine v. Chrestensen, 316 U.S. 52 (1942). Courts, therefore, would narrowly define the term "commercial speech" to avoid application of this doctrine. For a discussion of the development and demise of the "commercial speech doctrine," see *infra* notes 6-42 and accompanying text.

6. 316 U.S. 52 (1942). At issue in *Chrestensen* was whether the city could constitutionally prohibit the dissemination of a handbill advertising a submarine and soliciting visitors for a stated admission fee.

duly burden the communication of information, "the Constitution imposes no such restraint on government as respects purely commercial advertising."[7] A long line of cases acknowledged *Chrestensen* as a virtual exclusion of "commercial" speech from first amendment protection.[8] This exclusion is generally referred to as the "commercial speech doctrine."

The constitutional protection afforded commercial speech was gradually increased, however, because a restrictive definition of the term "commercial speech" limited the doctrine.[9] This process of definitional narrowing reached its zenith in *Bigelow v. Virginia*.[10] In *Bigelow*, the Court considered whether an advertisement, placed in a Virginia newspaper by a profit-making New York organization offering to perform abortions, was commercial speech.[11] Finding that it was not, the Court viewed the advertisement to be dissimi-

---

7. *Id.* at 54. According to one commentator, the *Chrestensen* Court "without citing precedent, historical evidence, or policy considerations . . . effectively read commercial speech out of the first amendment." Redish, *The First Amendment in the Marketplace: Commercial Speech and the Values of Free Expression*, 39 GEO. WASH. L. REV. 429, 450 (1971).

The first indication that communications of a commercial nature were not protected by the first amendment appeared in Schneider v. State, 308 U.S. 147, 165 (1939) (dicta) ("We are not to be taken as holding that commercial soliciting and canvassing may not be subjected to such regulation as the ordinance requires"). *See also* Semler v. Oregon State Bd. of Dental Examiners, 294 U.S. 608 (1935) (applying the fourteenth amendment to affirm a state prohibition of advertising by dentists); Packer Corp. v. Utah, 285 U.S. 105 (1932) (applying the fourteenth amendment to affirm a state prohibition of cigarette ads on billboards). For a general discussion of the Court's early treatment of commercial speech, see T. EMERSON, THE SYSTEM OF FREEDOM OF EXPRESSION 413-17 (1970); Rotunda, *supra* note 4, at 1084-96; Note, *supra* note 4, at 1842-59; Comment, *supra* note 4, at 207-13.

8. *See, e.g.*, New York Times Co. v. Sullivan, 376 U.S. 254, 266 (1964) (editorial advertisements distinguished from commercial advertisements); Breard v. Alexandria, 341 U.S. 622, 642 (1951); Martin v. City of Struthers, 319 U.S. 141, 142 n.1 (1943); Morgan v. City of Detroit, 389 F. Supp. 922, 926-28 (E.D. Mich. 1975); Capital Broadcasting Co. v. Mitchell, 333 F. Supp. 582 (D.D.C. 1971), *aff'd mem.*, 405 U.S. 1000 (1972). *Capital Broadcasting* has been described as the "high water mark" of the commercial speech doctrine. J. BARRON & C. DIENES, CONSTITUTIONAL LAW: PRINCIPLES AND POLICIES 834 (1975).

The broad holding in *Chrestensen* was attacked in several dissents. *See, e.g.*, Pittsburgh Press Co. v. Pittsburgh Comm'n on Human Relations, 413 U.S. 376, 398 (1973) (Douglas, J., dissenting); *id.* at 401 & n.6 (Stewart, J., dissenting); Dun & Bradstreet, Inc. v. Grove, 404 U.S. 898, 904-06 (1971) (Douglas, J., dissenting from denial of certiorari).

9. *See generally* New York Times Co. v. Sullivan, 376 U.S. 254 (1964) (paid political advertisement not commercial); Cammarano v. United States, 358 U.S. 498 (1959) (Douglas, J., concurring) (profit motive makes no difference); Murdock v. Pennsylvania, 319 U.S. 105 (1943) (sale of religious books, without profit motive, not commercial speech); Rotunda, *supra* note 4.

10. 421 U.S. 809 (1975).

11. *Id.* at 812. The advertisement in *Bigelow* consisted of information about a New York abortion referral agency and the availability of legal abortions in New York:

lar from the "commercial" advertisement in *Chrestensen*: the advertisement in *Bigelow* "did more than simply propose a commercial transaction."[12] Because the advertisement "contained factual material of clear 'public interest,' " it was outside the ambit of the commercial speech doctrine.[13]

Finally, in *Virginia State Board of Pharmacy v. Virginia Citizens Consumer Council*,[14] the Court reversed its earlier position, holding that even commercial speech, "which does no more than propose a commercial transaction," is within the scope of the first amendment's protection.[15] Noting that earlier decisions had se-

---

UNWANTED PREGNANCY
LET US HELP YOU
Abortions are now legal in New York
There are no residency requirements.
FOR IMMEDIATE PLACEMENT IN ACCREDITED
HOSPITALS AND CLINICS AT LOW COST
Contact
WOMEN'S PAVILLION
515 Madison Avenue
New York, N.Y. 10022
or call any time
(212) 371-6670 or (212) 371-6650
AVAILABLE 7 DAYS A WEEK
STRICTLY CONFIDENTIAL
We will make all arrangements for you and
help you with information and counseling.

Id.

12. *Id.* at 822.

13. *Id. See also* New York Times Co. v. Sullivan, 376 U.S. 254, 266 (1964) (editorial advertisements are not "purely" commercial speech); *cf.* Pittsburgh Press Co. v. Pittsburgh Comm'n on Human Relations, 413 U.S. 376 (1973) (dicta) (certain commercial advertisements have some degree of protection).

Some lower courts read *Bigelow* as preserving a narrow category of unprotected, purely commercial speech. *See, e.g.*, Terminal-Hudson Elecs., Inc. v. Dep't of Consumer Affairs, 407 F. Supp. 1075, 1082 (C.D. Cal.) (Whelan, J., dissenting), *vacated*, 426 U.S. 916 (1976); Population Servs. Int'l v. Wilson, 398 F. Supp. 321, 337 (S.D.N.Y. 1975) (dictum), *aff'd*, 431 U.S. 678, 682 (1977); Urowsky v. Board of Regents, 38 N.Y.2d 364, 369-72, 342 N.E.2d 583, 586-87, 379 N.Y.S.2d 815, 819-21 (1975).

14. 425 U.S. 748 (1976).

15. *Id.* at 762. The Court acknowledged that after *Bigelow*, "some fragment of hope for the continuing validity of a 'commercial speech' exception arguably might have persisted because of the subject matter of the advertisement in *Bigelow*." *Id.* at 760. It is clear that after its decision in *Virginia State Board of Pharmacy*, even that fragment of hope is gone.

As Justice Rehnquist noted in dissent: "The logical consequences of the Court's decision in this case, a decision which elevates commercial intercourse between a seller hawking his wares and a buyer seeking to strike a bargain to the same plane as has been previously reserved for the free marketplace of ideas, are far reaching indeed." *Id.* at 781 (Rehnquist J.,

verely limited the application of the commercial speech doctrine,[16] and motivated at least in part by its inability to define commercial speech, the Court "squarely faced" the question of whether a first amendment exception exists for commercial speech:

> If there is a kind of commercial speech that lacks all First Amendment protection . . . it must be distinguished by its content. Yet the speech whose content deprives it of protection cannot simply be speech on a commercial subject.
>
> . . . .
>
> Our question is whether speech does "no more than propose a commercial transaction," is so removed from any "exposition of ideas," and from " 'truth, science, morality, and arts in general, in its diffusion of liberal sentiments on the administration of Government,' " that it lacks all protection. *Our answer is that it is not.*[17]

Although the Court recognized that commercial speech may be restricted in ways that ideological speech may not,[18] it brought all regulation of commercial speech under constitutional scrutiny.[19]

---

dissenting).

16. *Id.* at 764-65.

17. *Id.* at 761-62 (emphasis added) (citations omitted).

18. The Court stated:

> In concluding that commercial speech enjoys first amendment protection, we have not held that it is wholly undifferentiable from other forms. There are common sense differences between speech that does "no more than propose a commercial transaction," and other varieties. Even if the differences do not justify the conclusion that commercial speech is valueless, and thus subject to complete suppression by the State, they nonetheless suggest that a different degree of protection is necessary to insure that the flow of truthful and legitimate commercial information is unimpaired. The truth of commercial speech, for example, may be more easily verifiable by its disseminator than, let us say, news reporting or political commentary . . . . Also, commercial speech may be more durable than other kinds. Since advertising is the *sine qua non* of commercial profits, there is little likelihood of its being chilled by proper regulation and foregone entirely.
>
> Attributes such as these, the greater objectivity and hardiness of commercial speech, may make it less necessary to tolerate inaccurate statements for fear of silencing the speaker. They may also make it appropriate to require that a commercial message appear in such a form, or include such additional information, warnings, and disclaimers as are necessary to prevent its being deceptive. They may also make inapplicable the prohibition against prior restraints.

Id. at 771 n.24 (citations omitted).

19. Since the decision in *Virginia State Board of Pharmacy*, the Court has applied the first amendment to protect commercial speech. *See supra* note 1. In some cases, however, the Court has upheld the regulation restricting commercial speech based on the substantiality of the states' countervailing interests. *See, e.g.*, Friedman v. Rogers, 440 U.S. 1 (1979) (ban on the use of trade names by optometrists is constitutional); Ohralik v. Ohio State Bar Ass'n, 436 U.S. 447 (1978) (disciplinary rules regulating in-person solicitation by attorneys for pecuniary gain are constitutional). *But see* Metromedia, Inc. v. City of San Diego, 453 U.S. 490 (1981) (ban on billboards valid as to prohibition of commercial advertising, but

The commercial speech doctrine had apparently been laid to rest.

To the casual observer, it appeared that commercial speech could be regulated or prohibited only if the speech, like any other speech, lost its protection by virtue of its content[20] or if the state established a countervailing interest sufficient to justify the regulation.[21] The Court, in *Virginia State Board of Pharmacy*, however, laid the foundation for a new commercial speech doctrine by recognizing that there are "common sense differences" between speech that does no more than propose a commercial transaction and other varieties of speech:

> Even if the differences do not justify the conclusion that commercial speech is valueless, and thus subject to complete suppression by the State, they nonetheless suggest that a different degree of protection is necessary to insure that the flow of truthful and legitimate commercial information is unimpaired.[22]

Shortly after the decision in *Virginia State Board of Pharmacy*, the Court began the process of giving constitutional significance to the common sense distinctions between commercial and noncommercial speech. In *Bates v. State Bar of Arizona*,[23] the Court considered the constitutionality of a prohibition on attorney advertising. Departing from its analytical approach in pre-*Virginia State Board of Pharmacy* decisions, the Court in *Bates* found it

---

ordinance declared unconstitutional on other grounds).

20. For example, obscenity is excluded from the realm of constitutionally protected speech based on its content. *See, e.g.*, Miller v. California, 413 U.S. 15, 23 (1973); Roth v. United States, 354 U.S. 476, 485 (1957); see also New York v. Ferber, 102 S.Ct. 3348 (1982) (child pornography is without first amendment protection).

21. The first amendment appears to speak in absolute terms: "Congress shall make no law . . . abridging the freedom of speech." U.S. CONST. amend. I. The Supreme Court, however, has not interpreted it in that manner. As Justice Harlan stated:

> Throughout its history this Court has consistently recognized at least two ways in which constitutionally protected freedom of speech is narrower than an unlimited license to talk. On the one hand, certain forms of speech, or speech in certain contexts, has been considered outside the scope of constitutional protection . . . . On the other hand, general regulatory statutes, not intended to control the content of speech but incidentally limiting its unfettered exercise, have not been regarded as the type of law the first or fourteenth amendment forbade Congress or the States to pass, when they have been found justified by subordinating valid governmental interests, a prerequisite to constitutionality which has necessarily involved a weighing of the governmental interest involved. . . .

Konigsberg v. State Bar, 366 U.S. 36, 49-51 (1961) (citations and footnote omitted). *See also Virginia State Board of Pharmacy*, 465 U.S. at 770-73 (regulation of false speech or of time, place, and manner permissible).

22. 425 U.S. at 771 n.24.

23. 433 U.S. 350 (1977).

was unnecessary first to classify the speech as commercial or noncommercial. Instead, the Court assumed first amendment protection and analyzed the issue by balancing the individual's free speech rights against the state's justifications for banning the speech.[24] After reviewing each of the state's justifications, the Court concluded: "we are not persuaded that any of the proffered justifications rises to the level of an acceptable reason for the suppression of all advertising by attorneys."[25] In this initial stage of the Court's analysis, the only relevancy of the commercial nature of the speech was its weight in the balancing process that is utilized in all first amendment decisions.

In the later stages of the Court's analysis in *Bates*, however, the commercial nature of the speech took on added significance. In the usual first amendment case involving a restraint on noncommercial speech, a showing that the state regulation unconstitutionally suppressed speech would end the analysis. The Court would not require the person making the challenge to demonstrate that his or her specific speech was suppressed.[26] This doctrine, known as the "overbreadth doctrine," is based on the notion that overbroad statutes might serve to chill protected speech.[27] In *Bates*, however, the Court held that because of the nature of commercial speech, the regulation could not be challenged on the ground that it might be applied unconstitutionally in circumstances other than those before the court.[28] By refusing to apply the overbreadth doc-

---

24. In *Bates*, the Court briefly reviewed its decision in *Virginia State Board of Pharmacy* to reemphasize the protected nature of commercial speech. *Id.* at 363-65. The Court then considered the numerous justifications proffered by the state in support of the advertising ban.

25. *Id.* at 379.

26. *Id.* at 379-80. *See* Bigelow v. Virginia, 421 U.S. 809, 815-16 (1975); Lewis v. City of New Orleans, 415 U.S. 130, 133-34 (1974); Gooding v. Wilson, 405 U.S. 518, 520-21 (1972); Dombrowski v. Pfister, 380 U.S. 479, 486-87 (1965); Thornhill v. Alabama, 310 U.S. 88, 96-98 (1940).

27. For a discussion of the overbreadth doctrine, see generally Note, *The First Amendment Overbreadth Doctrine*, 83 HARV. L. REV. 844 (1970).

28. 433 U.S. at 380-81. In a recent case, the Court discussed this rule:

> The traditional rule is that a person to whom a statute may constitutionally be applied may not challenge that statute on the ground that it may conceivably be applied unconstitutionally to others in situations not before the court . . . . [I]t "would indeed be undesirable for this Court to consider every conceivable situation which might possibly arise in the application of complex and comprehensive legislation." By focusing on the factual situation before us, and similar cases necessary for development of a constitutional rule, we face "flesh-and-blood" legal problems with data "relevant and adequate to an informed judgment." This practice also fulfills a valuable institutional purpose: it allows state courts the opportunity to construe a law to avoid constitutional infirmities.

trine to commercial speech, the Court made the first constitutional distinction between commercial and noncommercial speech since the demise of the "commercial speech doctrine."

The continued importance of the classification of speech as commercial or noncommercial following *Bates* was further demonstrated by *Ohralik v. Ohio State Bar Association*[29] and *In re Primus*.[30] In these decisions the Court considered the constitutionality of state prohibitions of client solicitation by attorneys. In *Ohralik*, the prohibitions were considered in the context of a solicitation for purely pecuniary gain (commercial speech), while in *Primus*, the solicitation by an ACLU attorney was viewed as an associational aspect of expression (noncommercial speech). Using the analytical model set out in *Bates*, the Court reached a different conclusion with respect to the two types of speech. In the case of solicitation that proposed merely a "commercial transaction," the Court in *Ohralik* recognized the need for "prophylactic regulation in furtherance of the State's interest in protecting the lay public," and upheld the ban.[31] In *Primus*, however, the Court noted:

> Where political expression or association is at issue, [the] Court has not tolerated the degree of imprecision that often characterizes government regulation of the conduct of commercial affairs . . . . Although a showing of potential danger may suffice in the former context, [commercial speech] appellant [Primus] may not be disciplined unless her activity in fact involved the type of misconduct at which South Carolina's broad prohibition is said to be directed.[32]

Returning to the terminology it had apparently abandoned in *Virginia State Board of Pharmacy*, the Court distinguished between the scope of constitutional protections afforded speech that simply "proposes a commercial transaction" and protections for noncommercial expression.[33]

New York v. Ferber, 102 S.Ct. 3348, 3360 (1982) (citations and footnotes omitted). *See, e.g.*, Broadrick v. Oklahoma, 413 U.S. 601, 610 (1973); United States v. Raines, 362 U.S. 17, 21 (1960).

29. 436 U.S. 447 (1978).
30. 436 U.S. 412 (1978).
31. 436 U.S. at 468.
32. 436 U.S. at 434.
33. *Id.* at 437. The *Ohralik* court stated:

> At bottom, the case against appellant [Primus] rests on the proposition that a State may regulate in a prophylactic fashion all solicitation activities of lawyers because there may be some potential for overreaching, conflict of interest, or other substantive evils whenever a lawyer gives unsolicited advice and communicates an offer of representation to a layman. Under certain circumstances, that approach is appropriate in the case of speech that simply '*propose[s] a commercial transaction.*'

The commercial/noncommercial speech dichotomy was formally constitutionalized in two 1980 Supreme Court decisions concerning the constitutionality of regulations of New York's Public Service Commission as applied to the commercial and noncommercial speech of public utilities.[34] Although the Court held that both the commercial and noncommercial speech were constitutionally protected, it proposed distinct analytical models to determine the constitutionality of the regulations. In *Consolidated Edison Co. v. Public Service Commission*,[35] the Court enumerated three theories to support the state's ban on noncommercial speech: whether the prohibition was (1) a time, place, or manner regulation; (2) a permissible subject-matter regulation; or (3) a narrowly tailored means of serving a compelling state interest.[36] On the other hand, in *Central Hudson Gas & Electric Corp. v. Public Service Commission*,[37] the Court announced a new four-step analysis by which to evaluate regulations on commercial speech:

> At the outset, we must determine whether the expression is protected by the First Amendment. For commercial speech to come within that provision, it at least must concern lawful activity and not be misleading. Next, we ask whether the asserted governmental interest is substantial. If both inquiries yield positive answers, we must determine whether the regulation directly advances the governmental interest asserted, and whether it is not more extensive than is necessary to serve the interest.[38]

Id. (emphasis added) (citations omitted) (quoting Pittsburgh Press Co. v. Pittsburgh Comm'n on Human Relations, 413 U.S. 376, 385 (1973)). Note that *Pittsburgh* is a *Chrestensen*-type commercial speech doctrine decision.

34. Central Hudson Gas & Elec. Corp. v. Public Serv. Comm'n, 447 U.S. 557 (1980) (commercial speech); Consolidated Edison Co. v. Public Serv. Comm'n, 447 U.S. 530 (1980) (noncommercial speech).

35. 447 U.S. 530 (1980).

36. *Id.* at 535.

37. 447 U.S. 557 (1980). In *Central Hudson*, the Supreme Court held unconstitutional a ban on advertising promoting the use of electricity.

38. *Id.* at 566. The first step in the Court's four-part analysis reflects its earlier refusal to extend the overbreadth doctrine to commercial speech. *See supra* notes 27-28 and accompanying text. Because the regulation of noncommercial speech is subject to close scrutiny, the four-part test announced in *Central Hudson* should be viewed as supplementing the generally applicable standards considered in *Consolidated Edison*. *See supra* note 36 and accompanying text. If a regulation meets the standards applicable in the case of noncommercial speech, it would necessarily be valid as applied to commercial speech. The converse is not true. Commercial speech may be regulated in a manner that does not satisfy any of the three tests applicable in the case of noncommercial speech. Thus, the regulation of commercial speech would be valid if it could be justified under any of the *Consolidated Edison* standards *or* under the four-part test of *Central Hudson*.

Because the commercial/noncommercial speech dichotomy grants lower relative protec-

Although the Court used a different test in each case, it held both regulations of the Public Service Commission unconstitutional.[39] The decisions, however, clearly establish separate analytical models for the evaluation of regulations concerning commercial and noncommercial speech.[40]

The relevancy of a determination that speech is commercial appears to have come full circle. In *Chrestensen*, the determination was of singular importance because "commercial speech" was outside the protection of the first amendment. In subsequent cases, *Chrestensen*'s "commercial speech doctrine" survived, but was severely limited in its application by narrow definitions of the term "commercial speech." In *Virginia State Board of Pharmacy*, it appeared that the constitutional significance of the distinction between commercial and noncommercial speech was diminished because of the Court's holding that commercial speech was protected by the first amendment. But in that decision, the Court also recognized that commercial speech is entitled to a "different degree of protection" than noncommercial speech and thus provided the basis for renewed distinctions between them. Although the traditional commercial speech doctrine has been abolished, a "new" commercial speech doctrine has emerged. This "new commercial speech doctrine" extends first amendment protection to commercial speech, yet subjects it to a different analytical model than is used in noncommercial cases when balanced with competing state

---

tion to commercial speech, it has produced interesting results in a case involving the constitutionality of a billboard regulation applicable to both types of speech. In Metromedia, Inc. v. City of San Diego, 453 U.S. 490, 512-17 (1981), a plurality held that because noncommercial speech is entitled to heightened first amendment protection, a government decision not to proscribe commercial speech in certain instances constitutionally precludes proscription of noncommercial speech in the same circumstances. Therefore, a billboard ordinance, granting greater rights to commercial speech, was unconstitutional. *Id.* at 512-17. Justice Stevens, dissenting in part, found the plurality's holding somewhat "ironic" because it "concludes that the ordinance is an unconstitutional abridgement of speech because it does not abridge enough speech." *Id.* at 540 (Stevens, J., dissenting).

Metromedia's bifurcated approach, requiring courts or legislative bodies to clearly distinguish between commercial and noncommercial speech, lends further support to the need for a clear definition of the term "commercial speech."

39. 447 U.S. at 544; 447 U.S. at 571.

40. It should be noted that the Supreme Court has extended a substantial degree of first amendment protection to corporations whether or not they are engaged in the business of communication. In First Nat'l Bank v. Bellotti, 435 U.S. 765 (1978), the Court stated that the first amendment protects corporate speech, and that the corporate character of the speaker does not deprive the speech of "what otherwise would be its clear entitlement to protection." *Id.* at 778. For a discussion of *Bellotti*, and its ramifications, see generally Brudney, *supra* note 4; O'Kelley, *supra* note 4; Prentice, *supra* note 4.

interests.[41] The emergence of this new doctrine requires a rethinking of the definitional framework in which it is applied.[42]

III. A New Doctrine—A New Definition

Due to the different analytical models that apply in determining the constitutionality of commercial and noncommercial speech, it is essential to define clearly the term "commercial speech." The Supreme Court has never explicitly defined commercial speech, but has described it as "expression related solely to the economic interests of the speaker and its audience" or "speech proposing a commercial transaction."[43] Some commentators have suggested that commercial speech "does *no more* than solicit a commercial transaction or state information relevant thereto."[44]

---

41. *See supra* notes 27-40 and accompanying text.

42. For example, judges have recognized that the old definitions do not apply. *See* Suffolk Outdoor Advertising Co. v. Hulse, 43 N.Y.2d 483, 491-92, 373 N.E.2d 263, 267, 402 N.Y.S.2d 308, 372 (1977) (Fuchsberg, J., dissenting) ("Clearly the earlier cases in this area [commercial speech] can no longer be accepted uncritically"), *appeal dismissed*, 439 U.S. 808 (1978).

It should be noted that the recent Supreme Court activity in the commercial speech field has not been fully appreciated by all of the lower courts that have considered the classification issue. For example, in Scott v. Association for Childbirth at Home Int'l, 85 Ill. App. 3d 311, 407 N.E.2d 71 (1980), *modified*, 88 Ill. 2d 279, 430 N.E.2d 1012 (1981), the Attorney General filed a complaint alleging that the dissemination of information by the Association was an unfair or deceptive trade practice and unlawful under state law. In order to consider the Association's claim that the complaint violated its first amendment rights, the court discussed whether the defendant's speech—"advertising, educating, and training parents in the area of childbirth education"—was protected. Instead of relying on *Virginia State Board of Pharmacy*, the *Scott* court relied solely on Bigelow v. Virginia, 421 U.S. 809 (1975), to find the speech protected. 407 N.E.2d at 75. Further, the *Scott* court implied that the speech was noncommercial. *Id.* at 76. On appeal, the Illinois Supreme Court recognized the speech as commercial but found it was protected by the first amendment. That court stated that false speech may be regulated regardless of its commercial character, and noted that the overbreadth doctrine does not apply to commercial speech. Scott v. Association for Childbirth at Home, 88 Ill. 2d 279, 430 N.E.2d 1012, 1016-17 (1982). That decision incorrectly applied *Bigelow* as the standard for determining first amendment applicability and erroneously classified the speech as commercial. *See generally infra* notes 49-55 and accompanying text.

43. *Central Hudson*, 447 U.S. at 561-62. The Court has consistently limited the term "commercial speech" to purely commercial transactions. *See, e.g., In re* R.M.J., 455 U.S. 191, 204 n.17 (1982); *Bates*, 433 U.S. at 363-64; *Virginia State Board of Pharmacy*, 425 U.S. at 762; *Chrestensen*, 316 U.S. at 54. Other courts have adopted similar definitions. *See, e.g.*, Dunagin v. City of Oxford, 489 F. Supp. 763, 769 (N.D. Miss. 1980) ("speech designed to culminate commercial transactions"); Rutledge v. Liability Ins. Indus., 487 F. Supp. 5, 8 (W.D. La. 1979) (commercial aspect not enough, must propose commercial transaction); Adler, Barish, Daniels, Levin & Creskoff v. Epstein, 482 Pa. 416, 393 A.2d 1175, 1179 (1978), *appeal dismissed*, 442 U.S. 907 (1979) (speech that does no more than propose a commercial transaction).

44. Jackson & Jeffries, *supra* note 4, at 1 (emphasis added). *See also* Ferber, *supra*

Even if justification for such narrow definitions existed under the old commercial speech doctrine, they are no longer appropriate because of the doctrine's demise. Prior to *Virginia State Board of Pharmacy*, the classification of speech as *commercial* precluded constitutional analysis of state regulations affecting it.[45] Under the "new" commercial speech doctrine, a determination that the speech is commercial simply changes the constitutional analysis; it does not preclude it. Therefore, the need for an underinclusive definition of "commercial speech" that limits the amount of speech falling outside first amendment protection no longer exists. A new, rational definition should be adopted that permits the effective use of the commercial/noncommercial distinction in constitutional analysis.

In *Metromedia, Inc. v. City of San Diego*,[46] Justice Brennan noted that "our cases recognize the difficulty in making a determination that speech is either 'commercial' or 'noncommercial.' "[47] This difficulty, and the need for a new definitional framework within which to apply the "new" commercial speech doctrine, can be demonstrated by an example.

Assume that Greedy Oil Company, a major retailer of petroleum products, wants to increase business. The company would also like to see Proposition 10, a new oil tax, defeated. Furthermore, Senator Sally Long, a long-time friend of the oil industry, is running for re-election. Greedy decides to run a series of one-page advertisements, which picture a scene of the city on a beautiful sunny day with cars driving down the freeway. Each advertisement has different copy imprinted over this backdrop. There are four different advertisements:

---

note 4, at 389-90 (commercial speech is expression in which there is a direct functional relationship or nexus between the message and a later commercial transaction). One student has suggested three possible definitions of commercial speech: speech that does no more than propose a commercial transaction, speech of interest to a nondiverse consumer audience, and speech about a brand name product or service. Comment, *supra* note 4, at 228-34. The author of the Comment believes that "the scope and operation of the three proposed definitions would not vary greatly," *id.* at 233, and favors the brand name definition because it is overinclusive and because of its ease of application, *id.* at 234. It is for similar reasons that this article proposes an even more inclusive definition.

45. *See* Valentine v. Chrestensen, 316 U.S. 52 (1942). For a discussion of *Chrestensen*, see *supra* notes 6-8 and accompanying text.

46. 433 U.S. 490 (1981).

47. *Id.* at 539 (Brennan, J., concurring). *See also Virginia State Board of Pharmacy*, 425 U.S. at 764 (there are few commercial messages to which an element of public interest could not be added). Even *Chrestensen*, the decision from which the commercial speech doctrine evolved, the Court recognized that the dissemination of public information in a commercial message could cause difficulty in classifying speech. 316 U.S. at 55.

1. To move around this beautiful city, buy Greedy Oil, $1.39 a quart.
2. Keep your city running smoothly. Vote against Prop. 10. Paid for by Greedy Oil.
3. This is your city. Vote for Sally Long for Senator.
4. This is your city. A vote for Prop. 10 could cost 3 cents more for a quart of oil. Vote against Prop. 10. Help keep Greedy Oil the best value in town, only $1.39 a quart.

Although each of the above examples was motivated by the economic interests of Greedy Oil, and all were viewed as a means to the same end, only example 1 would clearly be classified as commercial speech under the present definition. Examples 2 and 3 would probably be classified as noncommercial speech. It is difficult to predict how example 4, a mixture of commercial and noncommercial speech, would be classified under the present definition.[48] All of these advertisements, however, should be classified as commercial, because they are motivated by the economic decision-making process of a business entity whose existence is solely for

48. The classification of "mixed" speech was first considered in *Chrestensen* where the respondent contended that his speech was not commercial because he appended a message to his advertisement that protested an action by the City Dock Department. The Supreme Court in *Chrestensen* avoided consideration of the first amendment's application to mixed speech, stating: "It is enough for the present purpose that the stipulated facts justify the conclusion that the affixing of the protest against official conduct to the advertising circular was with the intent, and for the purpose, of evading the prohibition of the ordinance." 316 U.S. at 55. In cases subsequent to *Chrestensen*, particularly those decided in the waning years of the commercial speech doctrine, the Court appears to have solved the problem by classifying all mixed speech as noncommercial. *See, e.g.*, Bigelow v. Virginia, 421 U.S. 809 (1975). Justice Rehnquist disagreed with this approach. Citing *Chrestensen*, he stated: "Whatever slight factual content the advertisement may contain and whatever expression of opinion may be laboriously drawn from it does not alter its predominately commercial content." *Id.* at 831-32 (Rehnquist, J., dissenting). Continuing, Justice Rehnquist noted that any advertiser wishing to avoid a prohibition of his advertising need "only append a civic appeal, or a moral platitude, to achieve immunity from the law's command." *Id.* at 832 (quoting *Chrestensen*, 316 U.S. at 55).

The presumption in *Bigelow* is against a finding of commercial speech. Therefore, example 4 would probably be classified as noncommercial. Subsequent cases and commentary also discussed the difficulty of distinguishing commercial and noncommercial speech. *See, e.g., In re* Primus, 436 U.S. 412, 438 n.32 (1978) (line between commercial and noncommercial speech will not always be easy to draw); Virginia State Bd. of Pharmacy v. Virginia Citizens Consumer Council, 425 U.S. 748, 764-65 (1976) (there are few advertisements to which an element of public interest could not be added); Brudney, *supra* note 4, at 284 (court's decision to bring commercial speech under first amendment rested in part in difficulties of classification); Kaufman, *The Medium, the Message and the First Amendment*, 45 N.Y.U. L. REV. 761, 769 (1970) (it is frequently difficult to separate commercial and noncommercial speech). *But cf.* Jackson & Jeffries, *supra* note 4, at 19-25 (line may be hard to draw in some cases "but the distinction between commercial speech and protected speech is relatively easy to maintain").

commercial purposes. This does not mean they are not protected by the first amendment; rather, it means that they are subject to the substantial, though different, first amendment protection afforded commercial speech.

To simplify the process of defining commercial speech, the Supreme Court should adopt a broad, inclusive definition, which would permit a court to concentrate on the application of constitutional principles rather than on definitional vagaries.[49] For purposes of first amendment analysis, commercial speech should be defined as *any* speech by a "commercial entity."[50] "Commercial entity" should mean any business entity[51] whose existence is based on profit,[52] excluding entities whose business is communication or

---

49. A recent comment describes three criteria that a definition of commercal speech should meet:

> In view of the definitional problems posed by the bifurcated approach, an adequate definition of commercial speech must satisfy three criteria: it must be sufficiently clear so that ordinances imposing greater restrictions on commercial speech will not significantly chill noncommercial expression; it must prevent any substantial degree of evasion by commercial advertisers; and it must permit an explanation of why commercial speech is less valuable than noncommercial speech. Although the Supreme Court has formulated three alternative definitions of commercial speech, none of these definitions, nor any reasonable modification of them, satisfy all three criteria.

Comment, *Standard of Review for Regulations of Commercial Speech:* Metromedia, Inc. v. City of San Diego, 66 MINN. L. REV. 903, 919 (1982) (footnotes omitted). The author of the comment notes that none of the current definitions satisfy these criteria. The definition proposed in this article satisfies all three.

50. It must be recognized that commercial entities are, in reality, incapable of expression. Although a corporation may not speak, it may be held legally responsible for the speech of its agents. Because commercial entities such as corporations are legal fictions having no independent lives of their own, they only communicate through their officers or representatives. For example, pronouncements by the board of directors represent the corporation, and each partner may speak for the partnership. This is not to imply that commercial entities should be denied first amendment rights. This article does not challenge the extension of broad-based first amendment rights to corporate speech. Instead, it proposes that commercial entities' speech should be protected as commercial speech. For an excellent discussion of the Supreme Court's extension of first amendment rights to corporations, see O'Kelley, *supra* note 4. *See also* Prentice, *supra* note 4. A rather unique approach to the treatment of corporate political speech may be found in Brudney, *supra* note 4 (corporate noncommercial speech may be less protected than corporate commercial speech).

51. "Business entity" should be broadly interpreted to include any business run by an individual, a partnership or a corporation.

52. Some difficulties exist in classifying the speech of trade unions and trade associations. The speech of trade unions or trade associations represents a united voice on a particular issue relevant to its members. The union or association is not a profit-making organization and protection of its speech is consistent with traditional first amendment theory. The Supreme Court has long recognized this concept. *See, e.g.*, AFL v. Swing, 312 U.S. 321, 325-26 (1941); Thornhill v. Alabama, 310 U.S. 88, 102 (1940). Although the interests of such organizations are primarily economic, the reason for their existence distinguishes them from

entertainment,[53] or that exists primarily for religious, charitable or civic purposes.[54] Thus, any speech of Greedy Oil, including all of

---

commercial entities.

The exclusion of trade unions or trade associations from the definition of commercial entity, however, results in an anomaly. In a labor dispute, for example, one party's speech would be classified as commercial while the other's would be classified as noncommercial. That result is mandated because, by definition, all speech of a commercial entity is commercial. However, the fact that speech is classified as commercial does not mean that it may be freely regulated. In fact, certain commercial speech, such as the speech of an employer in a labor dispute, could be protected to the same extent as noncommercial speech.

53. This definition excludes commercial entities engaged in the communications or entertainment business and those that exist to further social, political, or ideological interests. The first amendment rights of such institutions, unlike other commercial entities, are more closely related to the rights of an individual and should be protected by traditional first amendment notions. The Supreme Court, for example, has long recognized the special and constitutionally recognized role of the communications industry in informing and educating the public, offering criticism, and providing a forum for discussion and debate. *See* Bigelow v. Virginia, 421 U.S. 809, 828 (1975) (prosecution against publisher and editor of newspaper incurs more serious first amendment overtones); Saxbe v. Washington Post Co., 417 U.S. 843, 863-64 (1973) (Powell, J., dissenting) ("[News reporting] is the means by which the people receive that free flow of information and ideas essential to intelligent self-government"); Mills v. Alabama, 384 U.S. 214, 219 (1965) ("Suppression of the right of the press to praise or criticize governmental agents . . . muzzles one of the very agencies the Framers of our Constitution thoughtfully and deliberately selected to improve our society and keep it free"). As the Court has recognized, its decisions involving corporations in the business of communication or entertainment are based not only on "the role of the first amendment in fostering individual self-expression but also in affording the public access to discussion, debate, and the dissemination of information and ideas." First Nat'l Bank v. Bellotti, 435 U.S. 765, 783 (1978); *see* Red Lion Broadcasting Co. v. FCC, 395 U.S. 367, 389-90 (1969) (right of listeners paramount over right of broadcasters); *see also* Stanley v. Georgia, 394 U.S. 557, 564 (1969) (the Constitution protects right to receive information and ideas); Time, Inc. v. Hill, 385 U.S. 374, 389 (1966) ("Those guarantees are not for the benefit of the press so much as for the benefit of all of us").

The exclusion of entities in the business of communication also resolves a difficult concern voiced by Chief Justice Burger. In a concurring opinion in First Nat'l Bank v. Bellotti, 435 U.S. 765 (1978), the Chief Justice emphasized the adverse consequences that a decision to define communication entities as commercial would have on the rights of media corporations. *Id.* at 795-802 (Burger, C.J., concurring). The proposed exclusion recognizes the special nature of the individual's freedom of expression and the significance of the communications industry as an integral means of furthering such expression.

54. The Supreme Court has recognized that activities, including speech, of civic, religious or charitable organizations are modes of expression and association entitled to substantial first amendment protection. *See, e.g.*, NAACP v. Claiborne Hardware Co., 102 S.Ct. 3409, 3422-27 (1982); NAACP v. Button, 371 U.S. 415, 428-29 (1963); NAACP v. Alabama, 357 U.S. 449, 460-61 (1958). As the Court recently acknowledged, "the practice of persons sharing common views banding together to achieve a common end is deeply embedded in the American political process." Citizens Against Rent Control v. Berkeley, 102 S.Ct. 434, 436 (1982). This includes the solicitation of clients by attorneys associated with such organizations. *See In re* Primus, 436 U.S. 412 (1978). The exclusion of the speech of such organizations does not imply that it cannot be regulated or prohibited. It simply means that regulations or prohibitions must withstand "the exacting scrutiny applicable to limitations on core first amendment rights." Buckley v. Valeo, 424 U.S. 1, 44-45 (1976). The state must show a

the above advertisements, would be evaluated as commercial speech. The speech of an officer speaking for the corporation, an attorney acting in his or her capacity as an attorney, or a partner acting for the partnership, would also be defined as commercial speech. Of course, a member of the board of directors, an attorney, or a partner speaking as an individual is not included within the scope of this definition. The speech of such individuals, as well as the speech of any commercial entity excluded from the proposed definition of commercial entity, should not be presumed to be either commercial or noncommercial. Instead, speech not classified as commercial under the proposed definition should be classified on a case-by-case basis, in accordance with the developing judicial definitions. For example, the speech of a newspaper could be classified as either commercial or noncommercial, depending on the nature of the speech. An editorial run by the paper clearly should be classified as noncommercial. On the other hand, an advertisement soliciting subscriptions should be classified as commercial.[55]

IV. A New Definition: Justification

Under present definitions, the speech of commercial entities may be classified as either commercial or noncommercial/political.[56] Although the notion of commercial entities' noncommercial speech has superficial appeal, the distinction is unnecessary to preserve the first amendment rights of commercial entities and is inconsistent with the realities of the marketplace and our political system. The definition of commercial speech proposed here would eliminate commercial entities' noncommercial speech. It is proffered for three reasons. First, all speech of commercial entities is by its very nature "commercial." Second, judicial resources should be spent on constitutional analysis rather than definitional quibbling. Finally, the present practice of narrowly defining commercial speech may dilute traditional first amendment protections. The proposed overinclusive definition recognizes the "common sense" distinctions between individuals and commercial entities, maintains an independent and separate framework within which to analyze the individual's first amendment freedoms, and decreases the

---

compelling state interest to justify regulations. *Primus*, 436 U.S. at 432.

55. This type of advertisement would be classified as commercial under any of the definitions currently employed by the Court. *See supra* notes 43-44.

56. *Compare* Central Hudson Gas & Elec. Corp. v. Public Serv. Comm'n, 447 U.S. 557 (1980) (commercial entity's commercial speech) *with* Consolidated Edison Co. v. Public Serv. Comm'n, 447 U.S 530 (1980) (commercial entity's noncommercial speech).

likelihood of first amendment dilution.[57]

A. *Commercial Realities*

A commercial entity exists for the purpose of making a profit. It is not a direct participant in our democratic system, it does not vote and it cannot hold office. Moreover, its speech has no real relationship to either the freedom of individual self-expression or freedom of expression.[58] All speech of a commercial entity is presumably related to the reason for its existence (profit) and should, therefore, be classified as commercial.[59]

This is not to say that the speech of commercial entities is not of value to society because it often is. Instead, such speech should be classified and protected consistent with its societal role. Consider, for example, an advertising campaign by Mobil Oil explaining its position on various energy issues. These advertisements clearly inform the public and "critics would be hard put to deny that Mobil's editorial insistence has brought new facts to the public debate on energy, and in the process has influenced editorial thought and political actions."[60] Yet Mobil Oil would be equally hard put to deny that the advertisements were placed to better the economic position of Mobil Oil, at least indirectly.[61] While this

---

57. An additional benefit of the proposed definition is that it enables states to determine the degree of scrutiny to which their regulations will be subject. For example, commenting on the decisions in *Primus* and *Ohralik*, Justice Rehnquist stated, "I do not believe that any State will be able to determine with confidence the area in which it may regulate prophylactically and the area in which it may regulate only upon a specific showing of harm." *Primus*, 436 U.S. at 443 (Rehnquist, J., dissenting).

58. Although corporations are a legal fiction physically incapable of speech, it is not improper to write in terms of the "corporation's speech." It is, however, a serious error to conclude that corporations have constitutional free speech rights concurrent with individuals. Corporate speech, or for that matter the speech of any commercial entity, must be recognized as constitutionally distinguishable from the speech of individuals. The definition proposed here serves to establish such a distinction. It should be noted that in very limited circumstances a corporation may be given a right to vote. *See, e.g.*, Sayler Land Co. v. Tulare Water Dist., 410 U.S. 719 (1973) (voting rights in water district given to corporation).

59. In extending first amendment protection to commercial speech, the Court expressly recognized the profit motive behind such speech. *See Virginia State Board of Pharmacy*, 425 U.S. at 772 n.24 ("commercial speech may be more durable than other kinds. Since advertising is the *sine qua non* of commercial profits"); *see also* People v. Mobil Oil Corp., 48 N.Y.2d 192, 203, 397 N.E.2d 724, 730, 422 N.Y.S.2d 33, 40 (1979) (Gabrielli, J., dissenting) ("For one thing, commercial speech is firmly founded upon the profit motive."); Baker, *supra* note 4, at 9-14 (in our economic order, market forces dictate the content of commercial enterprises' speech by forcing the enterprises to orient their communications toward maximizing profits).

60. Banks, *The Rise of the Newsocracy*, ATL. MONTHLY Jan. 1981, at 54, 58.

61. This type of advertising is generally referred to as corporate image advertising,

type of speech may contribute to the marketplace of ideas, its purpose of increasing the market of the speaker should deny it equality with traditional political speech.[62] In extending broader first amendment protections to commercial speech, the Supreme Court

---

"advertising that describes the corporation itself, its activities or its policies, but does not explicitly describe any products or services sold by the corporation." FEDERAL TRADE COMM'N, STATEMENT OF ENFORCEMENT POLICY BY THE FEDERAL TRADE COMMISSION REGARDING CORPORATE IMAGE ADVERTISING 1-2 (Dec. 4, 1974). *See also* Bird, Goldman & Lawrence, *Corporate Image Advertising: A Discussion of the Factors that Distinguish Those Corporate Advertising Practices Protected Under the First Amendment From Those Subject to Control by the Federal Trade Commission*, 51 J. URB. L. 405 (1974); Note, *The Regulation of Corporate Image Advertising*, 59 MINN. L. REV. 189 (1974).

What is the practical justification for corporate image advertising? As is the case with traditional advertising, the motives of the corporation and the expected effect of the speech are essentially for the economic betterment of the corporation. For example, Falstaff Brewing Corporation recently ran a full-page advertisement that contained an open letter to President Reagan and encouraged consumers to "buy American." Although on its face the advertisement appeared to be ideological, the vice-president of Falstaff acknowledged it was a marketing device. *See* Wall St. J., Feb. 25, 1982, at 4, col. 2. Under the present definitional trend, corporate image advertising would probably be classified as noncommercial. *See* Consolidated Edison Co. v. Public Serv. Comm'n, 447 U.S. 530 (1980). Under the definition proposed in this article, that advertisement clearly would be commercial.

62. Consider whether "issue" or "institutional" advertising should be classified as noncommercial, thereby placing it on a constitutional par with the individual's traditionally protected speech. More and more commercial entities are using this type of advertising, which stresses issues rather than products. In considering issue advertising, the courts have evaluated it from a content standpoint and have concluded that such speech is noncommercial. The nature of the speaker or the reason that the advertisement was placed, though often recognized, is not given significance in the courts' analyses.

For example, in Rutledge v. Liability Ins. Ind., 487 F. Supp. 5 (W.D. La. 1979), the court had to classify insurance industry advertisements as commercial or noncommercial. These advertisements stressed loss prevention and legislative reform as the best ways to cope with current conditions in the field of torts. The purpose of the ads clearly was to reduce jury awards in tort cases: "It is clear that these ads do have a commercial aspect; a reduction in jury awards would operate to the financial advantage of liability insurance carriers." *Id.* at 8. Noting, however, that the "ads make no attempt to sell insurance or to recommend any particular type of insurance coverage; they propose no commercial transaction," the court concluded the advertisements were noncommercial speech. *Id.* It is interesting to observe that the court relied on pre-*Virginia State Board of Pharmacy* decisions to find significance in the fact that the ads communicate information about the industry and the industry's opinion.

For other examples of commercial entities' noncommercial issue-oriented speech, see Consolidated Edison v. Public Serv. Comm'n, 447 U.S. 530, 533 (1980); First Nat'l Bank v. Bellotti, 435 U.S. 765, 784-86 (1978); Quinn v. Aetna Life & Casualty Co., 616 F.2d 38 (2d Cir. 1980) (per curiam).

In one notable decision, a court has looked beyond the language itself to conclude that speech appearing to be ideological or noncommercial on its face may, in fact, be commercial. In National Comm'n on Egg Nutrition v. FTC, 570 F.2d 157 (7th Cir. 1977), *cert. denied*, 439 U.S. 821 (1978), the court classified speech as commercial even though it related to a controversial issue and did not propose a direct sale. The court apparently was swayed in its classification by the fact that the speech was a joint effort on the part of egg producers to increase sales.

has recognized that commercial speech is protected because of its value as a means of informing, educating and enlightening others.[63] Unlike traditionally protected speech, commercial speech is not protected as a means of self-expression.[64] Similarly, the so-called noncommercial speech of corporate entities serves a role in society quite distinct from that of an individual's speech. For example, in *First National Bank v. Bellotti*,[65] while holding that corporate political speech is entitled to a measure of first amendment protection, the Court implicitly conceded that corporate speech involves neither self-expression nor self-fulfillment. In *Bellotti*, the Court emphasized the "expression" itself rather than the speaker's right of expression.[66] Relying on the same analysis employed in the commercial speech cases,[67] the Court sought to protect the listener's interest in the free interchange of information, ideas and

---

63. *See, e.g.*, Central Hudson Gas & Elec. Corp. v. Public Serv. Comm'n, 447 U.S. 557, 561-63 (1980); Virginia State Bd. of Pharmacy v. Virginia Citizens Consumer Council, 425 U.S. 748, 762-65 (1976).

Note that the Supreme Court has declared that the central purpose of the first amendment is to protect political expression. *See, e.g.*, Buckley v. Valeo, 424 U.S. 1, 14 (1976); Monitor Patriot Co. v. Roy, 401 U.S. 265, 272 (1971); Mills v. Alabama, 384 U.S. 214, 218 (1966); New York Times Co. v. Sullivan, 376 U.S. 254, 270 (1964) (quoting Whitney v. California, 274 U.S. 357, 375-76 (1927) (Brandeis, J., concurring)). Professor Meiklejohn defines political expression as "speech which bears, directly or indirectly, upon issues with which voters have to deal." A. MEIKLEJOHN, POLITICAL FREEDOM 79 (1960). According to Meiklejohn, commercial speech would be entitled only to "due process" protection as private speech. *Id.* at 79-83.

64. Professor Emerson observed:

> The first amendment is said to serve four basic speech interests in a democratic system. . . .
>
> First, freedom of expression is essential as a means of assuring individual self-fulfillment. . . .
>
> Second, freedom of expression is an essential process for advancing knowledge and discovering truth. . . .
>
> Third, freedom of expression is essential for providing participation in decision-making by all members of society. . . .
>
> Finally, freedom of expression is a method of achieving a more stable community [by substituting reason for force in politics].

T. EMERSON, *supra* note 7, at 6-7. Commercial speech as defined here essentially serves to fulfill only the second of these interests. Commercial entities cannot achieve individual self-fulfillment, they do not participate in our democratic system and they only indirectly participate in the process of open discussion. It is not necessary to constitutionally recognize commercial entities' noncommercial speech to advance knowledge and the discovery of truth. A broad definition of commercial speech adequately furthers these speech interests, consistent with the function of the first amendment.

65. 435 U.S. 765 (1978).

66. *Id.* at 776, 783.

67. *See, e.g.*, Virginia State Bd. of Pharmacy v. Virginia Citizens Consumer Council, 425 U.S. 748, 750-57 (1976).

opinions.[68]

The theoretical basis on which the Court has extended first amendment protection to commercial speech recognizes two facts: commercial speech, unlike traditionally protected speech, serves a more limited function in our society;[69] and by its very nature, commercial speech will not be inhibited by virtue of lessened first amendment rights.[70] In essence, commercial speech is entitled to substantially the same protections as noncommercial speech, subject only to minor differences based primarily on the speech's content and the speaker's motivation. For example, in *Bates v. State Bar of Arizona*,[71] the Court refused to extend the overbreadth doctrine to commercial speech.[72] After explaining the justification for the application of the overbreadth doctrine in traditional speech cases, the Court stated:

> But the justification for the application of overbreadth analysis applies weakly, if at all, in the ordinary commercial context. As was acknowledged in *Virginia Pharmacy Board v. Virginia Consumer Council*, there are "common sense differences" between commercial speech and other varieties. Since advertising is linked to commercial well-being, it seems unlikely that such speech is particularly susceptible to being crushed by overbroad regulation. Moreover, concerns for uncertainty in determining the scope of protection are reduced;

---

68. 435 U.S. at 777.

69. In a representative democracy, the first amendment is primarily an instrument to enlighten public decisionmaking. *See generally*, A. MEIKLEJOHN, FREE SPEECH AND ITS RELATION TO SELF-GOVERNMENT 24-27 (1948). Traditionally protected speech, related to the exposition of ideas, truth, science, morality and the arts in general, directly furthers this goal. Commercial speech, on the other hand, is protected because of its value in preserving a predominantly free enterprise economy through a free flow of information. Virginia State Bd. of Pharmacy v. Virginia Citizens Consumer Council, 425 U.S. 748, 765 (1976). This view was expressed in Justice Stewart's concurring opinion:

> Ideological expression, be it oral, literary, pictorial, or theatrical, is integrally related to the exposition of thought—thought that may shape our concepts of the whole universe of man . . . . Commercial price and product advertising differs markedly from ideological expression because it is confined to the promotion of specific goods or services. The First Amendment protects the advertisement because of the "information of potential interest and value" conveyed, rather than because of any direct contribution to the interchange of ideas.

Id. at 779-80 (Stewart, J., concurring) (footnote and citation omitted).

70. In Friedman v. Rogers, 440 U.S. 1 (1979), the Court noted that "[c]ommercial speech, because of its importance to business profits, and because it is carefully calculated, is also less likely than other forms of speech to be inhibited by proper regulation." *Id.* at 10. Although the Court's discussion is in reference to "pure commercial speech," it is equally applicable to commercial speech as defined in this article.

71. 433 U.S. 350 (1977).

72. *Id.* at 381. *See supra* notes 33-36.

> the advertiser seeks to disseminate information about a product or service that he provides, and presumably he can determine more readily than others whether his speech is truthful and protected.[73]

Although in *Bates* the Court was discussing "purely commercial speech," its reasoning is equally applicable to commercial speech as defined in this article. In each case, the speech is well planned, motivated by financial considerations, and not subject to being "crushed" by overbroad regulation. While the content of "pure" commercial speech and commercial speech as defined in this article differs, the motives of the speakers and the context are the same: a commercial entity spending vast amounts of money to place its name before the public in a favorable and profitable light.[74]

B. Ease of Classification

In *In re Primus*,[75] Justice Powell, writing for the majority, correctly noted that although the line between commercial and noncommercial speech may be hard to draw, that is no reason for

---

73. 433 U.S. at 380-81 (citations omitted).

74. The wealth and power of large business corporations have a "potent effect" on local and national politics. Brudney, *supra* note 4, at 237. This is an additional justification for establishing classifications of speech which permit a constitutional distinction between the speech of commercial entities and individuals. This fact has been recognized by the courts and has been given constitutional significance. For example, in FEC v. Weinstein, 462 F. Supp. 243 (S.D.N.Y. 1978), the court upheld the constitutionality of a prohibition on corporate political contributions, stating that:

> A corporation, which is a creature of state law, may expend its resources to speak on particular issues. *See Bellotti.* The right to self-expression, however, does not extend to corporate financial contributions. Such contributions by themselves say little, but given the realities and expense of modern communications, they may permit the indirect purchase of votes. To permit even small political contributions by corporations would alter the structure and presentation of political issues. Instead of encouraging individual free speech, the allowance of corporate contributions would obscure it.
>
> This danger is heightened by the aggregation of economic power possessed by the modern corporation, too evident to require demonstration.

Id. at 249 (footnote omitted). Comparing corporate contributions with those of an individual, the court noted that "the making of a political contribution by an individual is a means of self-expression which is directly linked to that individual's right to vote." *Id.* The definition of commercial speech proposed in this article recognizes this distinction. *See also* Baker, *Realizing Self-Realization: Corporate Political Expenditures and Redish's The Value of Free Speech*, 130 U. PA. L. REV. 646 (1982) ("Corporate resources are sufficient to dominate the financing of electoral as well as initiative and referendum campaigns"); Wright, *Money and the Pollution of Politics: Is the First Amendment an Obstacle to Political Equality?*, 82 COLUM. L. REV. 609 (1982) ("In two vitally important and, in my judgment, tragically misguided first amendment decisions, *Buckley v. Valeo* and *First National Bank of Boston v. Bellotti*, the Court has given protection to the polluting effect of money in election campaigns.").

75. 436 U.S. 412 (1978).

avoiding the undertaking.[76] Difficulty of classification is an insufficient reason for avoiding constitutional scrutiny,[77] unless a more objective, equally appropriate standard is available. Rather than abandoning the distinction between commercial and noncommercial speech because of the difficulties of its application, the inclusive definition of commercial speech proposed here should be adopted to eliminate many of the difficulties arising out of the existing definitional ambiguities.

Under the definition presently employed by the Supreme Court, speech is commercial if it is "expression related solely to the economic interests of the speaker and its audience," or if it simply "proposes a commercial transaction."[78] Although the Court has recognized that those definitions may be difficult to apply,[79] it has not yet been faced with an ambiguous set of facts giving rise to complex definitional analysis.[80] In light of the restrictive nature of

---

76. *Id.* at 438 n.32.

77. The difficulty of classifying speech for purposes of first amendment analysis is not confined to the commercial/noncommercial speech distinction. For example, in Abood v. Detroit Bd. of Educ., 431 U.S. 209 (1977), the Court made a constitutional distinction between funds spent for "collective bargaining, contract administration and grievance adjustment purposes," and funds spent for "political and ideological purposes unrelated to collective bargaining." *Id.* at 232-37. *See also* Railway Employees' Dept. v. Hanson, 351 U.S. 225 (1956) (holding that requiring non-union employees to contribute financially to unions did not violate the first amendment). In making this distinction, the Court recognized that some speech falls clearly into neither category, and that "[t]here will . . . be difficult problems in drawing lines between collective bargaining activities, for which contributions may be compelled, and ideological activities unrelated to collective bargaining, for which such compulsion is prohibited." 431 U.S. at 236. *See generally*, Gaebler, *Union Political Activity or Collective Bargaining? First Amendment Limitations on the Uses of Union Shop Funds*, 14 U.C.D. L. REV. 591 (1981).

78. Central Hudson Gas & Elec. Corp. v. Public Serv. Comm'n, 447 U.S. 557, 561-62 (1980). *See supra* notes 43-44.

79. *See, e.g.*, cases cited *supra* note 47.

80. Lower courts, however, have had some difficulty classifying speech. *E.g.*, consider Quinn v. Aetna Life & Casualty Co., 616 F.2d 38 (2d Cir. 1980) (per curiam). At issue in *Quinn* were advertisements placed by the defendant entitled: "Too Bad Judges Can't Read This to a Jury" and "And Now, the Big Winners in Today's Lawsuits." The essence of both advertisements was that the current system for handling tort claims resulted in excessive jury awards. They included language such as:

> "Every payer of liability insurance premiums is a loser."
>
> "The jury is cautioned . . . to bear in mind that money doesn't grow on trees. It must be paid through insurance premiums from uninvolved parties such as yourselves."
>
> "We can ask juries to take into account a victim's own responsibility for his losses."
>
> "Insurers, lawyers, judges—each of us shares the blame for this mess. But it is you, the public, who can best begin to clean it up."

Quinn v. Aetna Life & Casualty Co., 96 Misc. 2d 545, 409 N.Y.S.2d 473, 475 (Sup. Ct. 1978). The trial court, apparently relying on the fact that the purpose of the advertisements was the economic benefit of the advertiser, held the ads to be "commercial speech." *Id.* at 478.

the definition of commercial speech, however, it is likely that lawyers will attempt to clothe otherwise commercial speech in the guise of noncommercial expression.[81]

The definition proposed here would not remove all of the problems associated with classifying speech as commercial or noncommercial (for example, noncommercial entities may still engage in commercial speech), but many cases will be easily classified simply by identifying the speaker as a commercial entity. This ease of classification would conserve judicial resources that might otherwise be spent distinguishing various degrees of commercial and noncommercial speech. Results would be based on constitutional analysis rather than definitional classification.[82]

The federal district court, upon removal, disagreed, calling the trial court classification a "fundamental misconception." Quinn v. Aetna Life & Casualty Co., 482 F. Supp. 22, 29 (E.D.N.Y. 1979), *aff'd*, 616 F.2d 38 (2d Cir. 1980). The advertisement, according to the district court, was "fully protected political expression." *Id.* For a discussion of *Quinn*, see Comment, *Corporate Advocacy Advertising: When Business' Right to Speak Threatens the Administration of Justice*, 1979 DET. C.L. REV. 623, 647-55, 666-70.

81. The possibility of manipulating the commercial/noncommercial distinction has been recognized by the present Court. Justice Rehnquist, for example, noted that "clever practitioners" could manipulate the present definitions of commercial and noncommercial speech:

> If Albert Ohralik, like Edna Primus, viewed litigation "not [as] a technique of resolving private differences" but as "a form of political expression" and "political association" for all that appears he would be restored to his right to practice. And we may be sure that the next lawyer in Ohralik's shoes who is disciplined for similar conduct will come here cloaked in the prescribed mantle of "political association" to assure that insurance companies do not take unfair advantage of policyholders.

In re Primus, 436 U.S. 412, 442 (1978) (Rehnquist, J., dissenting) (quoting the majority opinion, *id.* at 428, which in turn is quoting NAACP v. Button, 371 U.S. 415, 429, 431 (1962)) (citations omitted).

The Court long ago recognized that speakers may try to circumvent the appropriate constitutional standard by altering the context, or substance, of their speech. In Valentine v. Chrestensen, 316 U.S. 52 (1942), for example, the speaker changed a handbill containing pure advertising into a two-sided handbill with advertising on one side and a protest on the other. The Court avoided consideration of the nature of "mixed speech" by finding that:

> It is enough for the present purpose that the stipulated facts justify the conclusion that the affixing of the protest against official conduct to the advertising circular was with the intent, and for the purpose, of evading the prohibition of the ordinance. If that evasion were successful, every merchant who desires to broadcast advertising leaflets in the streets need only append a civic appeal, or a moral platitude, to achieve immunity from the law's command.

Id. at 55.

82. The problem of distinguishing between commercial and noncommercial speech may also confront city or state officials faced with the task of proposing the enforcing regulations on speech. For example, assume that a state desires to ban commercial billboards. In Metromedia, Inc. v. City of San Diego, 453 U.S. 490 (1981), the Court indicated that a prohibition of billboards, insofar as it regulated commercial speech, was constitutional. *Id.* at 512. Justice Brennan, however, noted the difficulties that could arise in enforcing an ordi-

Consider, for example, whether the decision in *Ohralik v. Ohio State Bar Association*[83] should rest even partially on a determination that the speech is commercial or noncommercial. As Justice Marshall noted: "What is objectionable about Ohralik's behavior here is not so much that he solicited business for himself, but rather the circumstances in which he performed that solicitation and the means by which he accomplished it."[84] It should not be necessary to determine whether Ohralik's speech "involved political expression or an exercise of associational freedom" to conclude that the state has the power to prevent attorneys from engaging in similar conduct.[85] By classifying all of an attorney's speech as commercial, and evaluating it under the four-part analysis set out in *Central Hudson*,[86] the courts could directly address the constitu-

---

nance prohibiting commercial billboards:

> It is one thing for a court to classify in specific cases whether commercial or noncommercial speech is involved, but quite another—and for me dispositively so—for a city to do so regularly for the purpose of deciding what messages may be communicated by way of billboards . . . . I would be unhappy to see city officials dealing with the following series of billboards and deciding which ones to permit: the first billboard contains the message "Visit Joe's Ice Cream Shoppe;" the second, "Joe's Ice Cream Shoppe uses only the highest quality dairy products;" the third, "Because Joe thinks that dairy products are good for you, please shop at Joe's Shoppe;" and the fourth, "Joe says to support dairy price supports: they mean lower prices for you at his Shoppe." Or how about some San Diego Padres baseball fans—with no connection to the team—who together rent a billboard and communicate the message "Support the San Diego Padres, a great baseball team." May the city decide that a United Automobile Workers' billboard with the message "Be a patriot—do not buy Japanese-manufactured cars" is "commercial" and therefore forbid it? What if the same sign is placed by Chrysler?
>
> I do not read our recent line of commercial cases as authorizing this sort of regular and immediate line-drawing by governmental entities. If anything, our cases recognize the difficulty in making a determination that speech is either "commercial" or "noncommercial."

Id. at 538 (citation and footnote omitted).

Adoption of the proposed definition would permit concise classification of speech, in many cases, simply through identification of the speaker. The bifurcated approach taken by the Court in *Metromedia* demonstrates the need for definitional precision.

83. 436 U.S. 447 (1978).

84. *Id.* at 470 (Marshall, J., concurring).

85. *Id.* at 458. In its opinion, the Court implied that if Ohralik's conduct had involved political expression or an exercise of associational freedom its analysis may have been different. *Id.* The nature of Ohralik's speech should not have influenced the outcome of the decision. The resultant abuse should have been sufficient to justify the state's action. Similarly, it should not have been necessary to determine whether the speech was commercial or noncommercial, or whether it had informational value, to establish constitutional protection. *See supra* notes 9-13 and accompanying text. The speech should simply have been classified as commercial. The regulations should then have been evaluated under the applicable standard, and the same outcome, with added simplicity of analysis, would have resulted.

86. *See supra* notes 34-38 and accompanying text.

tional issue of balancing the state's interest against the commercial speaker's first amendment rights.[87] Adoption of the proposed definition would establish an independent analytical basis for the speech of commercial entities and that of individuals. This approach would provide a logically consistent and principled rationale for the Court's prior cases and would focus judicial attention on the appropriate constitutional issues in future speech cases.[88]

C. *Dilution of Traditional First Amendment Protections*

Under the old commercial speech doctrine, speech was protected only if it was noncommercial.[89] Therefore, to challenge successfully a limitation on speech, it was first necessary to establish a protected speech interest. As a result, the Supreme Court occasionally went to extremes to "decommercialize" speech presented in a commercial context.[90] Once the speech was found to be protected, the same constitutional analysis was employed regardless of the speech's context.[91] For purposes of the first amendment, protected speech was protected speech. Although the result reached in the balancing process would differ, the process and standards were essentially the same, whether a ban on advertising or a ban on politi-

---

87. Applying the proposed definition, while sometimes changing the classification, would not alter the results of any of the noncommercial speech cases to date. *See, e.g.*, Consolidated Edison Co. v. Public Serv. Comm'n, 447 U.S. 530, 540-44 (1980) (speech would be commercial; asserted justifications without sufficient state interest and not drawn narrowly enough); First Nat'l Bank v. Bellotti, 435 U.S. 765, 786-95 (1978) (speech would be commercial; asserted state interests not substantially related to prohibition); Bigelow v. Virginia, 421 U.S. 809, 826-29 (1975) (speech would be commercial; asserted justifications not a substantial state interest); New York Times Co. v. Sullivan, 376 U.S. 254 (1964) (entity whose business is communication, speech still classified noncommericial); Quinn v. Aetna Life & Casualty Co., 482 F. Supp. 22 (E.D.N.Y. 1979), *aff'd*, 616 F.2d 38 (2d Cir. 1980) (speech would be commercial; asserted justification not a substantial state interest).

88. The proposed definition should answer Justice Rehnquist's complaint that the cases to date have not made a "principled distinction" between commercial and noncommercial speech. *See In re* Primus, 436 U.S. 412, 442 (1978) (Rehnquist, J., dissenting).

89. *See* Valentine v. Chrestensen, 316 U.S. 52 (1942). This doctrine was severely limited in its subsequent application. *See supra* notes 6-21 and accompanying text.

90. *E.g.*, Bigelow v. Virginia, 421 U.S. 809 (1975). In *Bigelow*, the Court took seven pages of analysis to determine that an advertisement for out-of-state abortions was protected. *Id.* at 818-25. The Court found that the advertisement did more than simply propose a commercial transaction. "It contained factual material of clear 'public interest.'" *Id.* at 822.

91. For example, in *Bigelow*, the Court appeared willing to apply the first amendment overbreadth doctrine to commercial speech. *Id.* at 815-18. In a subsequent case, after recognizing commercial speech as being protected by the first amendment, the Court held the overbreadth doctrine inapplicable to commercial speech. Bates v. State Bar, 433 U.S. 350, 379-81 (1977). *See supra* notes 23-28 and accompanying text.

cal activity was under consideration.[92]

The extension of a substantial degree of first amendment protection to commercial speech may have been effectuated, at least in part, to eliminate problems that could arise from a unified constitutional analysis of all speech. Due to the factual distinctions between speech arising in a commercial context and traditionally protected speech,[93] a dilution of the protections afforded the individual's political or ideological speech could result under a unified analysis. The Supreme Court itself has noted this danger:

> To require a parity of constitutional protection for commercial and noncommercial speech alike could invite dilution, simply by a leveling process, of the force of the Amendment's guarantee with respect to the latter kind of speech. Rather than subject the First Amendment to such a devitalization, we instead have afforded commercial speech a limited measure of protection, commensurate with its subordinate position in the scale of First Amendment values, while allowing modes of regulation that might be impermissible in the realm of noncommercial expression.[94]

---

92. There is language in pre-*Virginia State Board of Pharmacy* decisions suggesting that commercial speech was protected "differently" than noncommercial speech. *See, e.g.*, Pittsburgh Press Co. v. Pittsburgh Comm'n on Human Relations, 413 U.S. 376, 389 (1973) (dictum) (advertisements may receive some degree of first amendment protection); Bigelow v. Virginia, 421 U.S. 809, 821 (1975). What the Court was apparently referring to, however, was not a different analytical standard for commercial speech but rather the relative weight given the state's justifications when balanced against the speaker's first amendment rights. As the Court noted in *Bigelow*:

> Advertising, like all public expression, may be subject to reasonable regulation that serves a legitimate public interest. To the extent that commercial activity is subject to regulation, the relationship of speech to that activity may be one factor, among others, to be considered in weighing the First Amendment interest against the governmental interest alleged. Advertising is not thereby stripped of all First Amendment protection. The relationship of speech to the marketplace of products or of services does not make it valueless in the marketplace of ideas.

Id. at 826 (citations and footnotes omitted). It was not until the Court's decision in *Consolidated Edison* that a separate and distinct analytical model for evaluating commercial speech was established.

93. For example, consider insurance company advertisements that criticize the present system of adjudicating tort cases, or public utility advertisements that advocate the use of nuclear power. For purposes of constitutional analysis, the Court's present definitional approach equates these advertisements with an individual's political speech. *See, e.g.*, Consolidated Edison Co. v. Public Serv. Comm'n, 447 U.S. 530, 533-35 (1980) (public utility's billing insert advocating increased use of nuclear power is noncommercial speech); Quinn v. Aetna Life & Casualty Co., 616 F.2d 38, 40 (2d Cir. 1980) (per curiam) (insurance company's advertisements relating large jury awards to high insurance premiums is noncommercial speech); Rutledge v. Liability Ins. Indus., 487 F. Supp. 5, 8 (W.D. La. 1979) (advertisements similar to those in *Quinn* "not to be characterized as 'commercial speech' "). *See also supra* note 62.

94. Ohralik v. Ohio State Bar Ass'n, 436 U.S. 447, 456 (1978).

By specifically recognizing commercial speech as constitutionally protected, the Court partially resolved the leveling problem by dealing with commercial speech on an independent constitutional basis. Separate and distinct constitutional analysis may now be employed for commercial and noncommercial speech.[95] Persistence in narrowly defining commercial speech by relying on pre-*Virginia State Board of Pharmacy* definitions,[96] however, may thwart the benefits the Court hoped to gain through the "new" commercial speech doctrine. A parity of commercial entities' and individuals' speech can result in a leveling of the individual's protections, just as the *Ohralik* Court feared would result from a parity of constitutional protection for commercial and noncommercial speech.[97] By recognizing broad noncommercial speech rights of commercial entities, the Court invites dilution of first amendment protections. The "common sense distinctions" between the speech of commercial entities and that of individuals dictate the separate constitutional analyses which would result from the definition proposed here.[98]

---

95. Additionally, the Court's refusal to extend the first amendment overbreadth doctrine to commercial speech permits the Court to deal with commercial speech problems in the specific context in which they appear. *See supra* notes 27-28 and accompanying text.

96. The Court's definition of commercial speech has not reflected the change in the commercial speech doctrine. Many decisions still rely on Bigelow v. Virginia, 421 U.S. 809 (1975), and Pittsburgh Press Co. v. Pittsburgh Comm'n on Human Relations, 413 U.S. 376 (1973). *See, e.g., In re* Primus, 436 U.S. 412, 437 (1978) (relying on *Pittsburgh Press* for definition of commercial speech); Rutledge v. Liability Ins. Indus., 487 F. Supp. 5, 8 (W.D. La. 1979) (relying on *Bigelow* for definition of commercial speech); Scott v. Association for Childbirth at Home Int'l, 85 Ill. App. 3d 311, 407 N.E.2d 71, 75 (1980) (relying on *Bigelow* to establish first amendment protection). *But see* Princess Sea Indus. v. Nevada, 635 P.2d 281 (Nev. 1981), *cert. denied*, 102 S.Ct. 1972 (1982) (proper use of *Bigelow*). In light of subsequent judicial developments, *Bigelow* and *Pittsburgh Press* are of doubtful precedential value in the commercial speech context.

97. *See supra* note 94.

98. A parity of classification of the political speech of commercial entities and individuals does not dictate a parity of results. A court would still be free to distinguish between the nature of the speech and the context in which it appears to reach differing results. For example, a state's asserted justification for limiting corporate issue advertising, *see supra* note 62, while sufficient to support a ban of such advertisements, may not be as "compelling" when applied to a ban prohibiting an individual's political statement. Because of the applicability of the overbreadth doctrine, however, commercial entities would be able to successfully argue against the application of a regulation to their speech by demonstrating that the regulation would violate an individual's corresponding first amendment rights. The classification of all of a commercial entity's speech as commercial precludes analogy to the individual's first amendment rights. *See supra* notes 34-36 and accompanying text.

There is some question as to the tendency of the proposed definition to delimit "editorial advertising" such as that contained in New York Times Co. v. Sullivan, 376 U.S. 254 (1964). It is true that most editorial advertising, or issue advertising, would be classified as commercial under the proposed definition. For example, if the advertisement considered in *New York Times* were placed by a commercial entity it would, under the proposed defini-

By broadly defining commercial speech to include all speech of commercial entities, first amendment rights can be adequately protected without affecting the rights of individuals. On the other hand, if the present definitional approach continues, and commercial speech is narrowly defined to include only expression concerning "purely" commercial transactions, most speech of commercial entities will be classified as noncommercial.[99] Eventually, constitutional analogy to the concurrent rights of individuals is inevitable and the possibility of dilution of the individual's rights is quite real.

There is another possible outcome that could result from the broad recognition of commercial entities' noncommercial speech interests. Rather than diluting individuals' first amendment rights through a leveling process, commercial entities could be given first amendment rights commensurate with those of the individual. This possibility has not gone unnoticed by the Court. In *First National Bank v. Bellotti*,[100] Justice White, dissenting, considered the

---

tion, be classified as commercial as regards that entity (but not as regards the New York Times). This does not necessarily mean that it would not be granted significant constitutional protection. In fact, under the analytical test applied to commercial speech, such advertising would be entitled to the same protections afforded noncommercial speech under *New York Times*. It is doubtful that any state could establish a "substantial state interest" to override what the Court in *New York Times* viewed as a constitutional guarantee of "a federal rule that prohibits a public official from recovering damages for a defamatory falsehood relating to his official conduct unless he proves that the statement was made with 'actual malice'—that is, with knowledge that it was false or with reckless disregard of whether it was false or not." *Id.* at 279-80. For a discussion of the test applicable to commercial speech, see *supra* notes 34-38 and accompanying text. In this case, the classification of the speech as commercial would serve simply to distinguish commercial entities from individuals with no resultant lessening of constitutional protections.

99. As is demonstrated by Bigelow v. Virginia, 421 U.S. 809 (1975), most commercial messages, including direct advertising of a product or service, can be found to serve a "noncommerical" informative function. It would be easy for any advertiser to include a noncommercial element in any advertisement. *See* People v. Remeny, 40 N.Y.2d 527, 531, 355 N.E.2d 375, 378, 387 N.Y.S.2d 415, 417 (1976) (Fuchsberg, J., concurring) ("Indeed, as a practical matter, most communications may have a thread of both the commercial and the noncommercial running through them").

100. 435 U.S. 765 (1978). In *Bellotti*, the Court held unconstitutional a state criminal statute that prohibited expenditures by banks and business corporations for the purpose of affecting a vote on state referendum proposals not "materially affecting" any of the property, business or assets of the corporation. Writing for the majority, Justice Powell stated that the issue was whether the Massachusetts law "abridges expression that the first amendment was meant to protect." *Id.* at 776. The majority held that the law violated the first amendment. Although the Court classified the speech in question as political, and applied the traditional compelling state interest test, the result could have been the same if the speech had been classified as commercial. In the case of commercial speech, the test would have been the four-part test proposed in *Central Hudson*. *See supra* note 38. Under that test, the justifications proffered by the state in support of the ban would have to rise to the

implications of the Court's decision to recognize broad corporate first amendment rights:

> More importantly, the analytical framework employed by the Court clearly raises great doubt about the Corrupt Practices Act. The question in the present case, as viewed by the Court, "is whether the corporate identity of the speaker deprives this proposed speech of what otherwise would be its clear entitlement to protection," . . . . But the Court has previously held in *Buckley v. Valeo* that the interest in preventing corruption is insufficient to justify restrictions upon individual expenditures relative to candidates for political office. If the corporate identity of the speaker makes no difference, all the Court has done is to reserve the formal interment of the Corrupt Practices Act and similar state statutes for another day.[101]

Although the majority opinion in *Bellotti* suggests that Justice White's concerns are ill-founded,[102] his comments point out the ba-

---

level of a "substantial state interest." The state's justifications were: (1) its "interest in sustaining the active role of the individual citizen in the electoral process;" and (2) its "interest in protecting the rights of shareholders whose views differ from those expressed by management on behalf of the corporation." 435 U.S. at 787. The Court dismissed the first proffered justification because there was "no showing that the relative voice of corporations [was] overwhelming or even significant in influencing referenda in Massachusetts, or that there [was] any threat to the confidence of the citizenry in government." *Id.* at 789-90 (footnote omitted). As to the second asserted justification, the Court stated that the "purpose of the statute is belied, however, by the provisions of the statute, which are both underinclusive and overinclusive." *Id.* at 793. As viewed by the Court, the justifications would clearly have been insufficient to have established the substantial state interest necessary to support the statute if the speech was classified as commercial. It should be further noted that in *Bellotti* the Court relied, in part, on commercial speech cases to support its conclusions. *See id.* at 791 n.31 ("The First Amendment rejects the 'highly paternalistic' approach of statutes like § 8 which restrict what the people may hear"). The Court relied on Virginia State Bd. of Pharmacy v. Virginia Citizens Consumer Council, 425 U.S. 748, 770 (1976), and Linmark Assocs. v. Township of Willingboro, 431 U.S. 85, 97 (1977), both commercial speech cases.

101. 435 U.S. at 820-21 (White, J., dissenting). The relevant section of the Federal Corrupt Practices Act has been re-enacted in the Federal Election Campaign Act of 1971, Pub. L. No. 225, 86 Stat. 3 (current version at 2 U.S.C. § 441b (Supp. IV 1980)). *See also* Cox, *Foreword: Freedom of Expresson in the Burger Court*, 94 HARV. L. REV. 1, 70 (1980) (*Bellotti* will "increas[e] the relative influence of organizations with large financial resources, . . . shrinking the attention paid to truly individual voices [resulting in] a net loss of human freedom").

102. 435 U.S. at 788 n.26. Subsequent to *Bellotti*, the lower courts considering the constitutionality of the Federal Election Campaign Act (sometimes referred to as the Federal Corrupt Practices Act) have agreed with the majority. *See* FEC v. National Right to Work Comm., 501 F. Supp. 422 (D.D.C. 1980), *rev'd on other grounds*, 665 F.2d 371 (D.C. Cir. 1981), *cert. granted*, 102 S.Ct. 1766 (1982) (mem.); FEC v. Weinstein, 462 F. Supp. 243 (S.D.N.Y. 1978). *Cf.* FEC v. Lance, 635 F.2d 1132 (5th Cir. 1981) (en banc), *appeal dismissed*, 453 U.S. 917 (1981) (§ 441b(a), regulation forbidding contributions or expenditures by a national bank is constitutional). *See generally* Symposium, *Political Action Committees and Campaign Finance*, 22 ARIZ. L. REV. 351 (1980).

sic problem inherent in a recognition of commercial entities' noncommercial speech: the first amendment rights of individuals do not naturally extend to commercial entities. The problem foreseen by Justice White demonstrates the difficulties that lie ahead if the Court continues to ignore the true nature of the speech under consideration and maintains the narrow definition of commercial speech.

If the present trend of increased recognition of commercial entities' noncommercial speech continues, it seems inevitable that either the individual's first amendment rights will be reduced, or that commercial entities will be given first amendment rights coexistent and coextensive with individuals. Neither alternative is desirable. The definition proposed by this article, which permits the Court to establish parallel, but not equal, paths for the constitutional rights of commercial entities and individuals, should eliminate, or at least reduce, this problem.

V. Conclusion

The proposition that commercial speech enjoys a substantial degree of first amendment protection can no longer seriously be questioned. Yet the courts, in considering what constitutes commercial speech, continue to define the term in the context of the old commercial speech doctrine, which excluded commercial speech from constitutional protection. As a result, profit-motivated speech of commercial entities is equated with ideological speech of individuals for purposes of first amendment analysis.

The present practice of narrowly defining commercial speech will cause a dilution of traditional first amendment values and result in the misallocation of judicial resources. A new definition of commercial speech that simplifies the analytical process and clearly distinguishes between the profit-motivated speech of commercial entities and the speech traditionally protected by the first amendment should be adopted. This separation preserves traditional first amendment analysis while maintaining substantial first amendment protections for commercial speech.

Simply stated, commercial entities exist for commercial purposes and their speech is commercial by its very nature. To classify such speech as noncommercial, based on its content or its informational value, ignores the realities of the marketplace. Constitutional analysis, to be effective, must not ignore such realities. Classifying commercial speech according to its actual role in society

affords adequate constitutional protection to commercial speech and gives vitality to the new commercial speech doctrine.

Umbrella Pricing and Antitrust Standing: An Economic Analysis

*Roger D. Blair**
*Virginia G. Maurer***

One of the purposes of the antitrust laws is to protect plaintiffs who have paid excessive prices for goods because of price-fixing conspiracies. The confused state of present case law, however, leaves uncertain the question of whether plaintiffs who have paid excessive prices for goods purchased from a competitor of price-fixing conspirators will be accorded standing to sue on the same basis as those who purchase directly from the conspiring price-fixers. This article explores the state of existing antitrust standing doctrines and provides an economic rationale for extending standing to plaintiffs who are injured by purchasing from a competitor of price-fixing conspirators. The article concludes that granting standing under an umbrella pricing theory to all who have been injured by a price-fixing conspiracy is consistent with the policy goals of deterring antitrust violations and compensating injured victims.

I. Background

Many price-fixing conspiracies do not include every firm in the industry. If an industry is comprised of a core of dominant firms plus a fringe of small firms, a price-fixing conspiracy may involve cooperation only among the dominant firms. When this happens, the small, competitive fringe firms will likely be ignored by the conspirators.[1]

There are a number of reasons why the conspirators will not want to include the fringe firms in their price-fixing conspiracy.

---

* Acting Chairman and Professor of Economics, University of Florida. B.A., 1964, M.A., 1966, Ph.D., 1968, Michigan State University.

** Assistant Professor of Business Law, University of Florida. B.A., 1968, Northwestern University; M.A., 1969, J.D., 1975, Stanford University.

The authors wish to acknowledge the research support of the Public Policy Research Center at the University of Florida.

1. Although the small firms' opinions will not be solicited, the conspirators will not ignore the existence and the output of the fringe firms. *See infra* notes 70-73 and accompanying text.

Because price-fixing is an illegal activity, conspirators will want to make it a clandestine activity as well.[2] The smaller the number of active conspirators, the easier it will be to conceal their illegal activity. Moreover, keeping the conspiracy small also keeps down the costs associated with group decisionmaking.[3] It therefore seems likely that a price-fixing conspiracy will not include all the firms in an industry.[4]

If the dominant firms fix prices, purchasers from the competitive fringe firms will still pay a price that exceeds what the market price would be in the absence of collusion.[5] That result is mandated by the competitive fringe firms' role as price takers. In other words, fringe firms will not act as though their output decisions have a perceptible influence on price. Accordingly, they charge the current market price and simply adjust their output level to maximize profits. Thus, the fringe firms set their prices under the "umbrella" of the dominant firms.

The ability of the purchasers from the nonconspiring, competitive fringe firms to recover treble damages from the conspirators under section 4 of the Clayton Act[6] is uncertain because of the Su-

---

2. Section 1 of the Sherman Act, 15 U.S.C. § 1 (1976), provides in relevant part that "[e]very contract, combination in the form of trust or otherwise, or conspiracy, in restraint of trade or commerce among the several States, or with foreign nations is declared to be illegal." *Id.*

3. Coordination costs, negotiation costs, etc., tend to rise as membership in the conspiracy increases. For a nonmathematical discussion of group behavior, see J.M. BUCHANAN & G. TURLOCK, THE CALCULUS OF CONSENT, 97-116 (1962); M. OLSON, THE LOGIC OF COLLECTIVE ACTION (1965).

4. The economist's usual marginal analysis applies to determining the optimal number of conspirators. An additional conspirator will be added, as long as the increase in the cartel's profits exceeds the increased costs.

5. For a formal economic demonstration of this result, see *infra* notes 66-78 and accompanying text. For now, it is enough to observe that competition among buyers for the lowest price will not permit a multi-price equilibrium. The low prices will be pushed up and the high prices will be pushed down.

6. 15 U.S.C. § 15 (Supp. IV 1980). Section 4 provides a treble damage remedy for antitrust violations: "Any person who shall be injured in his business or property by reason of anything forbidden in the antitrust laws . . . shall recover threefold the damages by him sustained and the cost of suit, including a reasonable attorney's fee." *Id.* Treble damage claims that arise out of price-fixing activities of the defendant must allege a violation of section 1 of the Sherman Act. *See supra* note 2.

The federal courts have concluded that "Congress did not intend the antitrust laws to provide a remedy in damages for all injuries that might conceivably be traced to an antitrust violation," Hawaii v. Standard Oil Co., 405 U.S. 251, 263 n.14 (1972) (citations omitted); thus, doctrines have been created to limit access to the treble damage actions.

These limiting doctrines are especially important in treble damage antitrust suits when a civil suit follows a successful government action. Section 5(a) of the Clayton Act, 15 U.S.C. § 16(a) (Supp. IV 1980), permits private plaintiffs in a treble damage action to use final

preme Court decision in *Illinois Brick Co. v. Illinois*.[7] In *Illinois Brick* the Court ruled that indirect purchasers may not sue on a theory that a price-fixing overcharge has been passed on to them.[8] The Court stated that to rule otherwise would have subjected the defendant to "a serious risk of multiple liability,"[9] and turned section 4 cases into problems of "massive evidence and complicated theories."[10] Although *Illinois Brick* did not deal with the standing of purchasers from nonconspiring *competitors* of the antitrust violator, that case was relied on heavily by the Third Circuit to deny standing to such purchasers. In *Mid-West Paper Products Co. v. Continental Group, Inc.*,[11] the Third Circuit held that purchasers from nonconspiring competitors could not sue to recover damages.[12] There is some indication in recent cases, however, that *Mid-*

---

judgments from criminal antitrust actions as prima facie proof of a violation of the antitrust law.

Since World War II, the private treble damages suit has become the major force for enforcement of the antitrust laws. The marked growth in private antitrust suits, and in the proportion of private suits to government actions, may be attributed to post-war court decisions that have tended to increase the availability of the action. *See* Blair, *Antitrust Penalties: Deterrence and Compensation*, 1980 UTAH L. REV. 57, 59-62.

7. 431 U.S. 720 (1977).

8. *Id.* at 728-29. *Illinois Brick* dealt with the offensive use of the "pass-on" doctrine in antitrust cases. The "pass-on" doctrine deals with the theory that a purchaser from an antitrust violator passes on any overcharge to its own customers. The State of Illinois sued the Illinois Brick Company and other manufacturers of concrete blocks for fixing prices. Illinois, however, did not purchase concrete blocks. It purchased buildings from general contractors who purchased masonry work from masonry contractors who purchased the concrete blocks from Illinois Brick. The Supreme Court ruled that only direct purchasers of the concrete blocks had standing to sue, precluding indirect purchasers from using the pass-on theory offensively.

The Supreme Court had earlier disallowed the defensive use of the pass-on doctrine as well. *See* Hanover Shoe, Inc. v. United Shoe Mach. Corp., 392 U.S. 481 (1967) (defendant could not reduce damages by claiming that the plaintiff had passed on most of the overcharges to its customers).

Illinois Brick led to legislative proposals to overrule it. For an analysis, see AMERICAN ENTERPRISE INSTITUTE, EXPANDING THE RIGHT TO SUE FOR ANTITRUST VIOLATIONS: PROPOSALS TO OVERRIDE THE *Illinois Brick* DECISION (1978).

There is a wealth of law review literature on the pass-on doctrine and *Illinois Brick*. *See, e.g.*, Harris & Sullivan, *Passing On the Monopoly Overcharge: A Comprehensive Policy Analysis*, 128 U. PA. L. REV. 269 (1979); Landes & Posner, *Should Indirect Purchasers Have Standing to Sue Under the Antitrust Laws? An Economic Analysis of the Rule of* Illinois Brick, 46 U. CHI. L. REV. 602 (1979).

9. 431 U.S. at 730.

10. *Id.* at 745.

11. 596 F.2d 573 (3d Cir. 1979).

12. *See infra* notes 31-45 and accompanying text. The court's holding departed from its earlier decisions and from those of other lower federal courts. *See* Illinois v. Ampress Brick Co., 67 F.R.D. 461, 468 (N.D. Ill. 1975), *rev'd on other grounds*, 536 F.2d 1163 (7th Cir. 1976), *aff'd sub nom.* Illinois Brick Co. v. Illinois, 413 U.S. 720 (1977); Wall Prods. Co.

West Paper may not be followed in other circuits,[13] and that purchasers from nonconspiring competitors may be accorded the protection of the antitrust laws as well as the right to recover for injuries caused by antitrust violations. From the perspective of both economic analysis and antitrust policy, this seems appropriate and desirable.

Antitrust cases and commentaries have distinguished the related concepts of antitrust injury and antitrust standing.[14] The concept of antitrust *injury* limits the type of injury for which compensation can be sought. By contrast, the concept of antitrust *standing* limits the type of plaintiff who can sue to redress an antitrust injury.[15] The concepts are related in that one who has not sustained the type of injury that the antitrust laws are intended to prevent cannot sue to recover damages.

In formulating the doctrine of antitrust injury, the courts have defined the type of injury that is actionable under section 4 of the Clayton Act.[16] This limiting doctrine implicitly assumes that there

---

v. National Gypsum Co., 357 F. Supp. 832, 840 (N.D. Cal. 1973); Washington v. American Pipe & Constr. Co., 280 F. Supp. 802, 807 (W.D. Wash. 1968).

13. *See* Pollock v. Citrus Assocs., Inc., 512 F. Supp. 711 (S.D.N.Y. 1981); *In re* Coordinated Pretrial Proceedings in Petrolum Products Antitrust Litigation, Civ. Nos. 81-5117, 81-5930 (9th Cir. Nov. 9, 1982) (interlocutory order). *But see In re* Coordinated Pretrial Proceedings in Petroleum Products Antitrust Litigation, 497 F. Supp. 218 (C.D. Ca. 1980). The Third Circuit's decision in *Mid-West Paper* has also been soundly criticized. *See* Comment, *Standing of Purchasers from Nonconspirators to Challenge Price-Fixing Conspiracy: Mid-West Paper Prod. Co. v. Continental Group, Inc.*, 93 HARV. L. REV. 598 (1980).

14. *See* Handler, *Changing Trends in Antitrust Doctrines: An Approach to Antitrust Injury*, 77 COLUM. L. REV. 979, 989-97 (1977).

15. P. AREEDA, ANTITRUST ANALYSIS 81-82 (3d ed. 1981). For a discussion of the relationships between antitrust injury and antitrust standing, see Handler, *supra* note 14, at 996-97.

16. The doctrine of antitrust injury permits recovery for injuries that reflect the anticompetitive effects either of the antitrust violation or of anticompetitive acts made possible by that violation. In Brunswick Corp. v. Pueblo Bowl-O-Mat, Inc., 429 U.S. 477 (1977), the Court described the plaintiff's burden of proof of antitrust injury: "Plaintiffs must prove antitrust injury, which is to say injury of the type the antitrust laws were intended to prevent and that flows from that which makes defendants' acts unlawful." *Id.* at 489. *Brunswick* involved alleged violations of the antimerger provisions of the Sherman Act. *See* 15 U.S.C. § 18 (1976). The lower courts' reception of the antitrust injury doctrine has been mixed, however, when defendants have sought to extend the concept in nonmerger cases. *See* Page, *Antitrust Damages and Economic Efficiency: An Approach to Antitrust Injury*, 47 U. CHI. L. REV. 467, 471 n.22 (1980).

Blue Shield v. McCready, 50 U.S.L.W. 4723 (U.S. June 22, 1982) (No. 81-225), is the Supreme Court's most recent decision on the limits of and the relationship between antitrust injury and standing under section 4. In *McCready*, a consumer who was denied reimbursement under her employer-provided health insurance plan for her visits to a clinical psychologist brought a class action suit against the health plan and the Neuropsychiatric Society of Virginia, alleging that the two organizations conspired to exclude and boycott

are some injuries caused in fact by the violation of the antitrust law that are not actionable. Although the lower federal courts have wrestled with theories of causation borrowed from various areas of the law,[17] recent Supreme Court decisions have developed a more policy-oriented approach for analyzing section 4's injury requirement.[18] That approach limits recovery under section 4 to the kind of injury that the antitrust laws are intended to prevent. Those injuries would include the particular losses that are a part of a general loss of consumer welfare resulting from a violation of the antitrust laws.[19]

The related limiting doctrine of antitrust standing is even more restrictive than antitrust injury. It denies court access to some persons who have actually suffered an antitrust injury. Recovery by those persons is barred either because of their relationship to the defendant, or because of their relationship to the injury.[20] Here too, courts have attempted to resolve analytical problems by drawing on theories developed in other areas of the

---

psychologists from receiving compensation under the prepaid health plan, in violation of section 1 of the Sherman Act. The district court for the Eastern District of Virginia granted the defendant's motion to dismiss on the ground that the injury alleged was "too indirect and remote to be considered 'antitrust injury'" to grant standing under section 4. The Fourth Circuit reversed and granted standing under a proximate cause theory. On appeal, the defendant petitioners urged that McCready's alleged injury was not related to a reduction in competition among health care plans, and thus was beyond the range of injuries that could be redressed under section 4. The Supreme Court upheld the circuit court and promulgated a broad rule of standing to include the plaintiff's injuries that were "inextricably intertwined with the injury the conspirators sought to inflict on psychologists and the psychotherapy market." 50 U.S.L.W. at 4727. The *McCready* test of standing appears to limit the use of the antitrust injury doctrine as a barrier to section 4 actions.

17. Some courts have analyzed the section 4 injury requirement as a problem of proximate causation or direct injury. *See, e.g.*, Comet Mechanical Contractors, Inc. v. E.A. Cowen Constr. Co., 609 F.2d 404, 406 (10th Cir. 1980); Reibert v. Atlantic Richfield Co., 471 F.2d 727 (9th Cir.), *cert. denied*, 411 U.S. 938 (1973). *See generally* Berger & Bernstein, *An Analytical Framework for Antitrust Standing*, 86 YALE L.J. 809 (1977) (proposing an analytical framework for balancing the fear of unreasonable damage awards with the compensatory and deterrent purposes of section 4). A similar analysis views antitrust injury as a concept that limits section 4 recovery to damages that "actually flow from the aspect [of the antitrust laws] that causes market inefficiency." Page, *supra* note 16, at 471.

18. *See* Illinois Brick Co. v. Illinois, 431 U.S. 720 (1977); Brunswick Corp. v. Pueblo Bowl-O-Mat, Inc., 429 U.S. 477 (1977). *See also* Areeda, *Antitrust Violations Without Damage Recoveries*, 89 HARV. L. REV. 1127 (1976) (Although the defendant may have violated the antitrust laws, some plaintiffs will be undeserving of damages).

19. *See* Berger & Bernstein, *supra* note 17; Page, *supra* note 16. Factors of economic efficiency shape the concept of antitrust injury and therefore should be used to shape the damage award.

20. *See* Berger & Bernstein, *supra* note 17; Shapiro, *Proof of Damages—A Causation Perspective*, 44 ANTITRUST L.J. 88, 93-96 (1975).

law,[21] resulting in confusion and inconsistent treatment of plaintiffs.[22]

The confusion caused three basic theories of standing to evolve. Those theories are: the direct injury theory, the target area theory and the zone of interests theory. The direct injury rule allows recovery by plaintiffs whose injuries are caused directly or "proximately" by an antitrust violation.[23] The target area test grants standing to plaintiffs who are "targets" of antitrust viola-

---

21. *See, e.g.*, Mid-West Paper Prods. Co. v. Continental Group, Inc., 596 F.2d 573 (3d Cir. 1979). *See also supra* notes 11-13. The concept of proximate cause has also been used explicitly in antitrust cases. *See, e.g.*, Comet Mechanical Contractors, Inc. v. Cowen Constr., Inc., 609 F.2d 404, 405 (10th Cir. 1980); Farnell v. Albuquerque Publ. Co., 589 F.2d 497, 500 (10th Cir. 1978); Reibert v. Atlantic Richfield Co., 471 F.2d 727 (10th Cir.), *cert. denied*, 411 U.S. 938 (1973). For a discussion of the propriety of using the proximate cause doctrine to resolve threshold standing questions, see Tyler, *Private Antitrust Litigation: The Problem of Standing*, 49 U. COLO. L. REV. 269, 270-71 (1978).

22. *Compare* Calderone Enters. Corp. v. United Artists Theatre Circuit, Inc., 454 F.2d 1292 (2d Cir. 1971) (owner of theater did not have standing to challenge motion picture distributor's distribution of first run films), *cert. denied*, 406 U.S. 930 (1972) *with* Congress Bldg. Corp. v. Loew's, Inc., 246 F.2d 587 (7th Cir. 1957) (owner of theater had standing to challenge motion picture distributor's distribution of first run films). The task of defining antitrust standing has challenged the federal courts for years. *See* Berger & Bernstein, *supra* note 17, at 818 nn.37-38; Handler, *The Shift from Substantive to Procedural Innovations in Antitrust Suits*, 71 COLUM. L. REV. 1. (1971). In Ostrofe v. H.S. Crocker Co, 670 F.2d 1378, 1383-88 (9th Cir. 1982), the Ninth Circuit adopted a balancing approach to standing issues that attempts to implement Congressional policy and implicitly rejects standing theories drawn from other areas of the law. The plaintiff allegedly was forced to resign from his job and was denied employment elsewhere in the industry because he refused to participate in a price-fixing conspiracy. The Ninth Circuit granted him standing to sue for treble damages on the theory that it would increase the deterrent effect of the law. *Id. See* Sherman, *Antitrust Standing: From* Loeb *to* Malamud, 51 N.Y.U. L. REV. 374 (1976); Tyler, *supra* note 21.

23. *See* Reibert v. Atlantic Richfield Co., 471 F.2d 727, 731 (10th Cir.), *cert. denied*, 411 U.S. 938 (1973); Sanitary Milk Producers v. Berjans Farm Dairy, Inc., 368 F.2d 679 (8th Cir. 1966); Volasco Prods. Co. v. Lloyd A. Fry Roofing Co., 308 F.2d 383, 395 (6th Cir. 1962), *cert. denied*, 372 U.S. 907 (1963); Snow Crest Beverages, Inc. v. Recipe Foods, Inc., 147 F. Supp. 907 (D. Mass. 1956); Loeb v. Eastman Kodak Co., 183 F. 704 (3d Cir. 1910).

In Montreal Trading Ltd. v. Amax, Inc., 661 F.2d 864 (10th Cir. 1981), *cert. denied*, 102 S.Ct. 1634 (1982), the Tenth Circuit denied standing to a plaintiff who claimed that its inability to purchase potash resulted from a concerted refusal of potash producers to deal with it, and a conspiracy among those producers to create a shortage of potash. The court ruled that the plaintiff was not "directly injured," noting that the conspirators "gained no 'fruits' from nonsales," and that a grant of standing could result in ruinous recoveries; furthermore, the injury was "inherently speculative." *Id.* at 867.

The evolution of the direct injury rule is traced in Berger & Bernstein, *supra* note 12, at 813-19. For articles tracing the evolution of antitrust standing theories, see Note, *Standing to Sue in Private Antitrust Litigation: Circuits in Conflict*, 10 IND. L. REV. 532 (1977); Note, *Antitrust and the Consumer Problems of Standing and Damage Recovery*, 25 S.D. L. REV. 392, 395-96 (1980).

tions, or who are within the "target area" of wrongful activity.[24] The target area concept evolved as a test for directness.[25] An offshoot of the target area test is the "foreseeable target area" test which grants standing to all those who could reasonably have been foreseen as falling within the target area of the alleged antitrust violation.[26] Finally, the "zone of interests" test requires the plaintiff to allege that he has been injured in fact and is within the "zone of interests to be protected or regulated by the statute or constitutional guarantee in question."[27] The test is borrowed from the standard used to measure a private party's standing to challenge an administrative agency ruling,[28] but it has been adapted to antitrust cases.[29]

---

24. *See, e.g.*, Long Island Lighting Co. v. Standard Oil Co., 521 F.2d 1269, 1274 (2d Cir. 1975), *cert. denied*, 423 U.S. 1072 (1976); Billy Baxter, Inc. v. Coca-Cola Co., 431 F.2d 183, 187 (2d Cir. 1970), *cert. denied*, 401 U.S. 923 (1971).

25. *See* Berger & Bernstein, *supra* note 17, at 830. *See generally* Lytle & Purdue, *Antitrust Target Area Under Section 4 of the Clayton Act: Determination of Standing in Light of the Alleged Antitrust Violation*, 25 AM. U. L. REV. 795 (1976) (arguing that a properly applied target area concept should remain the major section 4 standing test). For a discussion of the infirmities and imprecision of this test, see Berger & Bernstein, *supra* note 17, at 131-35; Sherman, *supra* note 18.

26. *See* Mulvey v. Samuel Goldwyn Productions, 433 F.2d 1073, 1076 (9th Cir. 1970), *cert. denied*, 402 U.S. 923 (1971); Hoopes v. Union Oil Co., 374 F.2d 480, 485 (9th Cir. 1967).

Like the target area test, see *supra* note 20, the foreseeable target area test has produced inconsistent results within and among the circuits. *See* Berger & Bernstein, *supra* note 17, at 831. In any event, "foreseeability is arguably irrelevant, for all antitrust injuries are intentionally rather than negligently inflicted." *Id.* at 835. Furthermore, those who may be in the target area, but who elect not to purchase due to excessive prices apparently have no standing to sue for treble damages. *See* Montreal Trading Ltd. v. Amax, Inc., 661 F.2d 864, 867-68 (10th Cir. 1981), *cert. denied*, 102 S.Ct. 1634 (1982). In *Amax*, the Tenth Circuit appears to have denied standing to a nonpurchaser because the injury was inherently speculative. However, the social welfare loss due to a price-fixing conspiracy should not be measured by the overcharge on the units sold at the conspiratorial price. Instead, it should be measured by the difference between the value to potential consumers of the units not sold and the social costs of producing those units. *See* R. POSNER, ECONOMIC ANALYSIS OF LAW 201-05 (2d ed. 1977).

27. Association of Data Processing Serv. Orgs. v. Camp, 397 U.S. 150 (1970).

28. *Id.* at 153.

29. *See* Malamud v. Sinclair Oil Co., 521 F.2d 1142, 1145 (6th Cir. 1975). *See also* Chrysler Corp. v. Fedders Corp., 643 F.2d 1229, 1234-35 (6th Cir. 1981) (holding that plaintiff must plead an antitrust injury "of the type section 4 was intended to remedy," and that the injury must be within the "zone of interests"); California State Council v. Associated General Contractors, 648 F.2d 527, 537-38 (9th Cir. 1980) (approving the zone of interest test in dictum); Oakland County Hearing Aid Service v. Sonotone Corp., 1977-2 Trade Cas. (CCH) ¶61,567 (E.D. Mich. 1977) (applying *Malamud* to determine standing). For a sharp criticism of the use of the zone of interest test in antitrust standing analysis, see Sherman, *supra* note 22, at 392-405. *See generally* Tyler, *supra* note 21, at 274-77 (analyzing the zone of interest test); Comment, *Standing to Sue Under Section 4 of the Clayton Act: Direct Injury, Target Area, or Twilight Zone*, 47 MISS. L.J. 502, 519-23 (1976) (commenting favor-

Some courts have rejected those conceptual tests, and instead have invoked a policy analysis approach to standing.[30] Commentators have recently argued that the policy considerations affecting standing should be drawn from the substantive goals of the antitrust statutes and tempered by the considerations of efficiency, practicality and fairness that inhere in those goals.[31]

II. The Confused State of Existing Case Law

The Supreme Court has not directly addressed the question of whether a purchaser from a competitor of colluding price-fixers may sue for treble damages under section 4 of the Clayton Act.[32] The lower federal courts that have considered this issue have reached inconsistent results.[33] A brief review of the decided cases will demonstrate those inconsistencies.

In *Washington v. American Pipe and Construction Co.*,[34] a

---

ably on the zone of interest test); Comment, *Standing to Sue in Antitrust: The Application of Data Processing to Private Treble Damage Actions*, 11 Tulsa L.J. 542, 556-62 (1976) (discussing the *Malamud* zone of interest test).

30. *See, e.g.*, Ostrofe v. H.S. Crocker Co., 670 F.2d 1378, 1382-83 (9th Cir. 1982); Cromar Co. v. Nuclear Materials & Equip. Corp., 543 F.2d 501, 508-09 (3d Cir. 1979); Braverman v. Bassett Furniture Indus., 552 F.2d 90, 99 (3d Cir. 1977).

31. *See* Berger & Bernstein, *supra* note 17. The authors identify the goals of section 4 as deterrence of would-be wrongdoers and compensation of injured victims, *see id.* at 848-49, and identify efficiency in the adjudicative process and avoidance of duplicative and ruinous damages as tempering considerations, *see id.* at 850-52. Furthermore, Berger and Bernstein suggest that standing decisions should be made by balancing the goals of section 4 against the countervailing considerations by the use of a four-pronged test for standing. *Id.* at 860-65. Under that analysis, courts should accord standing to purchasers from the defendant's nonconspiring competitors unless the policy against "overkill" is implicated. *Id.* at 879. In Ostrofe v. H.S. Crocker Co., 670 F.2d 1378 (9th Cir. 1982), the court adopted a "balancing" test similar to that outlined by Berger and Bernstein: "Since no limiting factor counvervails the compelling reasons for favoring standing the balance of competing policy considerations strongly favors allowing employees who suffer loss inflicted in retaliation for opposition to anticompetitive schemes to maintain suit." *Id.* at 1386.

32. The Supreme Court has consistently denied certiorari in antitrust standing cases despite the confusion in the lower courts. *See* Handler, *supra* note 14, at 994 n.89. *See also* P. Areeda, *supra* note 15, at 81 (Supreme Court has neither approved nor disapproved standing requirements created by lower courts). Nevertheless, the Court has given its apparent approval to some of the judicial limitations on standing. *See* Hawaii v. Standard Oil Co., 405 U.S. 251, 262 n.14 (1972): "The lower courts have been virtually unanimous in concluding that Congress did not intend the antitrust laws to provide a remedy in damages for all injuries that might conceivably be traced to an antitrust violation."

33. *See supra* notes 11-13 and accompanying text.

34. 280 F. Supp. 802 (M.D.L. 9th Cir.), *cert. denied*, 393 U.S. 842 (1968). In *American Pipe*, the defendants had pleaded nolo contendere to a federal criminal price-fixing charge. The plaintiffs, who were purchasers of the kind of pipe sold by the defendant, subsequently brought private suits seeking treble damages under section 4 of the Clayton Act. The defendant moved for summary judgment against those plaintiffs who had not purchased directly

federal district court used a "foreseeable target area" test of standing[35] to permit a plaintiff who had purchased pipe from a nondefendant nonconspirator to sue a company accused of fixing prices in the market for that product. The court viewed standing as involving two separate issues. First, whether the injury to the plaintiff was a proximate result of the defendant's violation of the antitrust laws; and second, whether the injury was so remote as to be beyond the reach of section 4.[36]

The court found the first issue, proximate cause, to be a question of fact for the jury. The court was unwilling to rule as a matter of law that the umbrella price theory did or did not supply the necessary causal relationship between the plaintiff's injury and the defendant's violation of the law. Instead, the court stated that this determination rested "upon inferences from facts within the exclusive province of the jury."[37] In dealing with the second issue, remoteness, the court employed the "foreseeable target area"[38] test for standing: "[I]njury is direct and proximate when it occurs within that . . . area [of economic activity] which it could reasonably be foreseen would be affected by the antitrust violation."[39] Because the conspiracy had attempted to raise the general level of pipe prices, all sales affected by the elimination of competition would lie within the area of economic activity that the defendant should reasonably have foreseen would be affected.[40] That affected area would include sales made by noncolluders.[41]

---

from it. The court split those plaintiffs into four classes and granted summary judgment against only one class consisting of plaintiffs who could not prove any purchases, and also those plaintiffs who could only identify purchases from other defendants and who had already settled their claims against those defendants. Summary judgment was denied if the plaintiffs had purchased from other conspirators, from nonconspirators or from an unknown source. *Id.* at 803-04.

35. *Id.* at 807 n.16. Several Ninth Circuit cases have employed the foreseeable target area test. *See, e.g.*, Hoopes v. Union Oil Co., 374 F.2d 480, 485 (9th Cir. 1967); Twentieth Century Fox Film Corp. v. Goldwyn, 328 F.2d 190, 212 (9th Cir.), *cert. denied*, 379 U.S. 885 (1964); Karseal Corp. v. Richfield Oil Corp., 221 F.2d 358, 363 (9th Cir. 1955); Conference of Studio Unions v. Loew's, Inc., 193 F.2d 51, 54-55 (9th Cir. 1951), *cert. denied*, 342 U.S. 919 (1952).

36. 280 F. Supp. at 806.

37. *Id.* at 806-07 (quoting Story Parchment Co. v. Paterson Parchment Paper Co., 282 U.S. 555, 566 (1931)).

38. *Id.* at 807. *See supra* note 18.

39. 280 F. Supp. at 807 (quoting Hoopes v. Union Oil Co., 374 F.2d 480, 485 (9th Cir. 1967)).

40. *Id.* at 807.

41. *Id.* To recover damages, the plaintiffs would need to prove that they had paid more for pipe purchased from nonconspirators than they would have paid in the absence of a conspiracy; furthermore, they would also have to prove that the defendant's participation

By reasoning that the affected area could include sales made by noncolluders, the court implied that the umbrella pricing theory was sufficiently predictable—and sufficiently probative of causation—as to place injuries explicable by that theory within the range of injuries that are foreseeable and, therefore, direct. A difficulty with the court's analysis is that it treats the standing issue as an aspect of the causation issue and does not distinguish the question of what kinds of injuries are caused by antitrust violations from the question of what kinds of injuries may be remedied under section 4.

In *Mid-West Paper Products Co. v. Continental Group,*[42] the Third Circuit purported to separate the question of actual economic causation from the policy question of standing. The plaintiff purchased ice cream packaging bags from noncolluding competitors of the defendants,[43] who were accused of conspiring to fix prices in the consumer bag market. The court assumed, arguendo, that the plaintiffs fell "within that level of the economy . . . threatened by the price-fixing conspiracy."[44] Nevertheless, the court held that even if the defendant's actions had actually harmed the plaintiffs, the plaintiffs did not have standing to sue because they had not purchased directly from the defendant.[45] The court found that only direct purchasers from the defendant were among those "whose protection is the fundamental purpose of the antitrust laws."[46]

The *Mid-West Paper* court's reliance on a privity test seemed

---

in the conspiracy caused those overcharges. *Id.*

42. 596 F.2d 573 (3d Cir. 1979).

43. Plaintiffs, some of whom had purchased bags indirectly and some of whom had purchased bags directly from competitors of the defendants, sought both injunctive relief and treble damages. All five defendants had been indicted by a federal grand jury and charged with violations of section 1 of the Sherman Act. *See* United States v. Continental Group, Inc., 456 F. Supp. 704 (E.D. Pa. 1978). Two of the defendants entered pleas of nolo contendere, two were convicted after jury trials, and one was found not guilty. *Id.* at 708. The district court, relying on Illinois Brick Co. v. Illinois, 431 U.S. 720 (1977), granted defendants' motions for summary judgment against the civil plaintiffs and dismissed their suits. Mid-West Paper Prods. Co. v. Continental Group, Inc., 596 F.2d 573, 576 (3d Cir. 1979). On appeal, the Third Circuit affirmed the lower court's dismissal of the treble damages claims of both the indirect purchasers and the purchasers from the defendants' competitors, but reversed the dismissal of the claims for injunctive relief. *Id.* at 586-90. The court ruled that injunctive relief required a lower threshold standing requirement than actions for treble damages. *See id.* at 593-94.

44. 596 F.2d at 583.

45. *See id.*

46. *Id.* Some commentators have also identified the narrow view of the "direct injury" test as essentially a rule of privity. Berger & Bernstein, *supra* note 17, at 819 & nn.41 & 43.

to avoid the questions of proximate causation and foreseeability that inhere in other theories of standing. The court justified the privity rule, however, on three bases. First, the court considered the line of causation between the antitrust violation and the nonprivy plaintiffs' injury too tenuous, concluding: "[I]t cannot readily be said with any degree of economic certitude to what extent, if indeed at all, purchasers from a competitor of the price-fixers have been injured by the illegal overcharge."[47] The court believed that possible intervening causes might break the causal connection between the antitrust violation and the injury[48] Second, the court believed that allowing the plaintiff to attempt the necessary proof of causation would require the type of complex economic proceeding that the Supreme Court discouraged in *Illinois Brick*.[49] Finally, the court rejected recovery by nonprivy plaintiffs because of the potential for ruinous damages—the defendant's liability could equal a multiple of all of the overcharges in the industry. That multiple could conceivably include overcharges that had not accrued to the defendant's benefit. The court found that result unfair and potentially injurious to competition as it could cause the failure of the defendant's firm.[50]

A vigorous dissent in *Mid-West Paper* rejected the privity test as too restrictive.[51] The dissent favored the proximate cause analysis as more effectively serving the multiple antitrust goals of compensation, competition and vigorous enforcement of the law. Denying standing, argued the dissent, could leave a real injury uncompensated. Moreover, it could discourage private enforcement of the law because direct purchasers might be wary of suing suppliers on whom they depended for a reliable, albeit expensive, supply

---

47: 596 F.2d at 584.

48. *See id.* The court cited elasticity of demand plus the competitors' inefficiency as possible alternative causes. To account for other possible causes, the plaintiff would have the burden of proving that the defendant's antitrust violation caused the competitor's decision to charge the higher price. *Id.* at 584 n.45. The plaintiff had to show "more than the creation of an umbrella; it must also establish the more difficult proposition that its supplier would have sold to it at a lower price had the conspiracy not existed." *Id.*

49. 431 U.S. 720 (1977). *See supra* note 10 and accompanying text. The *Mid-West Paper* court concluded:

> Illinois Brick represents in effect the proposition that when defendants have fixed prices above the competitive market price, where the benefit derived by them is readily ascertainable, the objectives of the treble damage action are fulfulled when the defendants are required to pay the direct purchasers three times the overcharge.

596 F.2d at 585 (footnotes omitted).

50. 596 F.2d at 586.

51. *Id.* at 595 (Higginbotham, J., dissenting).

of goods.[52] The dissent also accused the majority of misinterpreting *Illinois Brick*.[53] The dissent believed that *Illinois Brick* only dictated avoidance of duplicative recoveries and recoveries that required apportionment of damages among *vertically* related parties.[54] Thus, application of the reasonable foreseeability test to sales made by noncolluding competitors would not contravene *Illinois Brick*.[55]

Both the majority and the dissenting opinions in *Mid-West Paper* assumed that the burden of proving a causal link between antitrust violation and injury should lie with the plaintiff.[56] Furthermore, both majority and dissent agreed that this proof could involve questions of intervening or superceding causes.[57] The two opinions disagreed, however, on the extent to which the mandate of *Illinois Brick* precluded the plaintiff from factually proving the actual effects of the defendant's overcharges.[58] Both opinions assumed that the factual inquiry into the plaintiff's injury would focus on the price the plaintiff would have paid its supplier but for the price-fixing conspiracy, rather than on the price the plaintiff would have had to pay to obtain the same goods in a market untainted by the price-fixing scheme.[59] That assumption reflects a misunderstanding of the relevant economic analysis.[60]

In *Pollock v. Citrus Associates, Inc.*,[61] the district court for the Southern District of New York ruled that in some instances it may be immaterial whether the plaintiff has purchased directly from the defendant.[62] *Pollock* involved a unique fact pattern, however, that did not tempt the court to inquire into the choices or economic constraints of the noncolluding competitors. The plaintiff in *Pollock* sued on behalf of a class of persons who sold November 1977 orange juice futures contracts on the Citrus Associ-

52. *Id.* at 596.
53. *Id.* at 596 n.6.
54. *Id.* at 596-99.
55. *Id.* at 597. The dissent's analysis would allow the plaintiff to prove that its injuries were caused by the defendant's violation of the law. Judge Higginbotham concluded: "It is foreseeable, if not inevitable, that when those with a substantial share of the market fix prices, their competitors will also raise prices under the anticompetitive umbrella established by the price-fixers." *Id.*
56. *See id.* at 584, 598.
57. *See id.* at 585, 598-99.
58. *See id.* at 584 n.45, 599.
59. *See id.* at 584, 599.
60. *See infra* notes 88-89 and accompanying text.
61. 512 F. Supp. 711 (S.D.N.Y. 1981).
62. *Id.* at 719 (emphasis added).

ates of the New York Cotton Exchange and did not liquidate their positions.[63] On the defendant's motion to dismiss for the plaintiff's lack of standing, the district court applied the Second Circuit's "target area" test of "whether there is a legally significant causal relationship between the alleged violation and the alleged injury."[64] The court granted standing to the plaintiffs and went on to distinguish both *Illinois Brick* and the interpretation of *Illinois Brick* developed in *Mid-West Paper.*

The court emphasized two factors in its analysis. First, denying standing to the plaintiffs would effectively preclude all liability for the defendants because there would be no identifiable class of direct purchasers who could sue. Second, not only would the economic proof of causation be relatively simple in this case, but the identity of the plaintiffs' sellers would also be immaterial. The court concluded that: "Regardless of whether the plaintiffs ultimately purchased offsetting contracts from the defendants or from other traders with a long position, the price throughout the market allegedly rose as a result of the defendant's activities."[65]

The court further distinguished its theory of causation from that in *Mid-West Paper* stating:

> In the market in which supply is restricted, prices move up *naturally* pursuant to basic laws of supply and demand. In a market in which an oligopoly or price-fixing arrangement allows a relatively small seller to raise its price to the level protected by the price umbrella, the small seller is not "compelled" to raise his price to the same extent as a seller in a supply restricted market. Thus, the severe difficulties attendant with proving damages in an "umbrella" pricing situation, which troubled the court in *Mid-West Paper Products*, are not present in the case at bar.[66]

The court failed to explain, however, why a seller might be willing

---

63. *Id.* at 711. The plaintiffs claimed that the defendants restrained trade by achieving dominance and control of the long side of the available November contracts, thus driving the price of those contracts to artificially high levels. *Id.* at 715. When an individual buys a futures contract, he pays now for the guarantee of being able to buy a particular commodity on a specific future date at a specific price. The seller is gambling that the spot price on the due date will be lower than the contractual price. If the spot price rises above the contractual price, the seller loses.

64. *Id.* at 718 (quoting Reading v. Kennecott Copper Corp., 631 F.2d 10, 13 (2d Cir. 1980)). The rationale of the target area rule is to accommodate a sufficient number of private plaintiffs to deter antitrust law violations while preventing an "overkill" use of the treble damages weapon. *Id.* (quoting Calderone Enters. v. United Artists Theatre Circuit, 454 F.2d 1292 (2d Cir. 1971)).

65. *Id.* at 719 (footnote omitted).

66. *Id.* at 719 n.9 (emphasis added).

to sell at the apparent market price, instead of a lower price, in one kind of market structure and not in another. Sellers in either market are free to ignore the upward movement of the price level and take less. To inquire into why sellers might sell below market price in *Mid-West Paper* and not in *Pollock*, is to assume that sellers are free to violate basic laws of supply and demand. That inquiry is too hollow to support the court's distinction. It is possible that the court was merely straining to find a factual distinction between *Pollock* and *Mid-West Paper*. It is also possible that the court failed to realize that an oligopolistic market also is a supply-restricted market. Thus, both the noncolluding competitors and the buyers were faced with very similar sets of economic choices in *Mid-West Paper* and *Pollock*.[67]

An equally important aspect of *Pollock* is the court's attempt to distinguish the case from the Second Circuit's decision in *Reading Industries v. Kennecott Copper Corp.*[68] In *Reading*, a buyer of scrap copper alleged that it paid artificially high prices for scrap copper because the defendants had colluded to fix the price of refined copper at an artificially low price. The Second Circuit denied standing under a "target area" theory[69] rejecting as "conjectural" the plaintiff's argument of "attenuated economic causality that would mire the courts in intricate efforts to recreate the possible permutations in the causes and effects of a price change."[70]

The *Pollock* court correctly distinguished *Reading* as necessitating an economic analysis far more complex than was required in *Pollock*. The two cases are superficially similar as both involved defendants who were accused of fixing prices which caused the plaintiffs to overpay, although the defendant sellers had not sold any of their product to the plaintiffs directly. The distinction between the cases, however, lies in the difficulty of identifying the product market in which sales were made.[71] In *Reading*, a court would have to determine the effect of price-fixing in the sale of one form of the product (refined copper) on the price paid by a buyer

---

67. The nature of this similarity is developed in Sections III and IV of this article.

68. 631 F.2d 10 (2d Cir. 1980).

69. The Second Circuit adhered to a narrower target area test for standing. *See supra* note 24.

70. 631 F.2d at 14.

71. Product market definition is a vexing problem in antitrust generally. For a general discussion, see 2 P. AREEDA & D. TURNER, ANTITRUST LAW 346-87 (1978); L. SULLIVAN, HANDBOOK OF THE LAW OF ANTITRUST 41-74 (1977). In essence, a definition is needed that includes all reasonably close substitutes but is not excessively broad. There are enormous practical difficulties in finding a theoretically appealing product market definition.

for what is arguably a different product (scrap copper) sold in a different market. Although an economist might predict and approximate both effect and damages, the analysis necessarily would be more complicated and more tenuous than it would be in a case involving a single market.[72] In *Pollock*, the court correctly perceived that the obvious identity of the buyers' and sellers' markets made the proof of damages—and of economic causation—no more complicated than if the buyer had purchased directly from the seller.

Finally, in the 1981 case of *In re Bristol Bay, Alaska, Salmon Fishery Antitrust Litigation*,[73] the federal district court for the Western District of Washington permitted sellers of raw salmon to sue wholesale buyers who conspired to fix salmon prices, even though the plaintiffs had sold the salmon to nonconspiring competitors of the defendant. The *Salmon Fishery* court separated the issue of proving causation from the standing issue. As to causation, the court recounted the plaintiff's burden of proving a "but for" causal relationship between the defendant's acts and the plaintiff's injury. A decision to grant standing, said the court, "in no way lessens the plaintiff's burden of establishing that it was the defendant's acts and not other market forces which occasioned the price structure that prevailed when the plaintiffs sold their fish."[74]

As to the standing issue, the court applied broadly the "target area" standing test. It found the "target area" to be the market for salmon in Bristol Bay, Alaska. The standing questions, then, were two-fold. First, the court asked whether the plaintiffs were parties whose injuries were the result of defendant's disruption of "free market forces" within the Bristol Bay salmon market.[75] Then the court addressed the effect of *Illinois Brick* on the standing of plaintiffs who deal directly with nonconspirators. The court considered both the factual distinctions between *Illinois Brick* and

---

72. If a single market is involved, the price charged by each of the sellers is influenced by the same economic factors. Although it is not trivial to sort out the quantitative significance of each factor, the effects can be approximated by statistical techniques known as regression analysis. If more than one market is involved, however, the economic factors influencing the prices in question may vary from market to market which necessarily makes the statistical analysis more difficult. For an introductory treatment of regression analysis, see J. KMENTA, ELEMENTS OF ECONOMETRICS 201-16 (1971).

73. 530 F. Supp. 36 (1981).

74. *Id.* at 37 n.3. Presumably, the plaintiffs would be free to use competent economic evidence to establish at trial the existence and effects of a price umbrella. On a motion to dismiss, however, the court found it unnecessary to evaluate the credibility of that proof.

75. *Id.* at 40.

Salmon Fishery and also the policy considerations underlying the Supreme Court's decision in *Illinois Brick*. In *Illinois Brick*, but not in *Salmon Fisheries*, the damage calculation would be complicated by independent pricing decisions made on different levels of distribution. The measure of damages in *Salmon Fishery*, said the court, would be "the difference between the price which would have obtained in the presence of competition and that which actually prevailed under the pressure of the conspiracy."[76] Moreover, the facts of *Salmon Fishery*, unlike those of *Illinois Brick*, did not pose the possibility of dual recovery by direct and indirect dealers for the same injury; only direct sellers would be granted standing to sue and they could recover only to the extent they could prove damages. The court said also that the policies of *Illinois Brick*—"creating a vehicle for recovery by injured plaintiffs, limiting recovery to those for whom it is most appropriate, and providing that damages are not unreasonably large"[77]—were enhanced and followed by granting standing in *Salmon Fishery*.[78]

The approaches to standing in *Illinois Brick, Mid-West Paper, Pollock* and *Reading* reflect the confused development of standing doctrine in section 4 cases. The opinions reveal uncertainty about the theory of economic causation and the proper relationship between economic causation and antitrust policy. Notwithstanding the convoluted and confused state of the antitrust standing doctrine in present case law, it appears rational, both from the standpoint of economic theory and antitrust policy, to grant standing to a plaintiff who has purchased from a noncolluding competitor of a price-fixing defendant. The limit to this theory may lie in the problem of ascertaining whether the product was

---

76. *Id.* at 39.

77. *Id.*

78. The reasoning and holding of *Salmon Fishery* survives the Ninth Circuit's recent decision in *In re Coordinated Pretrial Proceedings in Petroleum Products Antitrust Litigation*, Civ. Nos. 81-5117, 81-5930 (9th Cir. Nov. 9, 1982), where the court denied standing to indirect purchasers from non-conspirators. The Ninth Circuit noted the multi-level distribution problem that distinguished *Petroleum Products* from *Salmon Fishery*:

> We need not decide, however, whether, in a situation involving a single level of distribution, a single class of direct purchasers from non-conspiring competitors of the defendants can assert claims for damages against price-fixing defendants under an umbrella theory. In the case before us, the umbrella claimants purchased gasoline from independent marketers who, in turn, purchased their gasoline from independent refiners. These independent refiners manufactured a percentage of the independent marketers' supply and brokered the remainder of the marketers' supply from major refiners, *i.e.*, the defendants.

Id. slip op. at 6.

purchased in substantially the same market in which the price-fixing occurred.[79] If the product is purchased in substantially the same market, however, *Illinois Brick*'s proscription of complex economic analysis should not preclude standing under an umbrella price theory.

III. The Umbrella Pricing Model: An Economic Analysis

This analysis assumes the existence of an industry that produces a homogeneous product[80] and has a core of large firms plus a competitive fringe.[81] It is further assumed that the dominant firms collude on price and permit the competitive fringe to do as it pleases.[82] This model then examines the pricing behavior of the competitive fringe and the impact on those who buy from the competitive fringe.[83] The analysis clearly demonstrates that all buyers of the industry's products are injured to the same extent and thus should have the same standing to sue under section 4 of the Clayton Act.

A. The Partial Conspiracy

As a benchmark, it is necessary to first establish the preconspiracy price and output. In Figure 1, the consumer demand for

---

79. Market definition problems are among the most troublesome in antitrust. *See supra* note 70.

80. All aspects of the product provided by various sellers must be identical for the industry output to be homogeneous. That assumption is made to abstract from the inevitable complications introduced by heterogeneity. When the product is homogeneous, buyers' decisions are only affected by price. For a brief discussion of product homogeneity, see G. Stigler, The Organization of Industry 60-62 (1968).

81. This industry structure is not uncommon as entry often occurs at small scales. The small firms are described as a "competitive fringe" because they will behave competitively, acting as price-takers. Small firms will simply charge the price that they find in the market and adjust their output in order to maximize profit.

82. Given the assumption that the numerous fringe firms are not included in the conspiracy, the conspirators can treat them in a predatory way and attempt to drive them out of the industry. That result is not very appealing, however, because predatory behavior may attract the attention of the antitrust authorities. Moreover, predatory behavior requires selling at very low prices, which in turn demands a sacrifice of present profits for the uncertain prospect of future profits.

83. Our model will be recognized as a variant of the so-called dominant firm price leadership model. For a brief discussion of the dominant firm model, see Landes, *An Introduction to the Economics of Antitrust*, in R. Posner & F. Easterbrook, Antitrust 1055, 1066-69 app. (2d ed. 1981). This model is found in nearly every standard intermediate microeconomics text. Our collusive variant of the model was presented in mathematical form by T.R. Saving. *See* T.R. Saving, *Concentration Ratios and the Degree of Monopoly*, 11 Int'l Econ. Rev. 139 (1970).

the product is represented by the negatively sloped line labelled D which shows the quantity that consumers are willing to buy at various prices. The curve denoted S_F shows the amount that the competitive fringe will supply at various prices. Similarly, S_C shows the supply response of the dominant firms to various prices. The industry supply curve is obtained by horizontally adding these supply curves together:

$$S = S_C + S_F$$

For prices between R and T, only the dominant firms will supply any output. Consequently, the industry supply curve over that price range coincides with S_C. For all prices above R, however, the two supply curves are added horizontally resulting in the kinked industry supply curve labelled TAS.

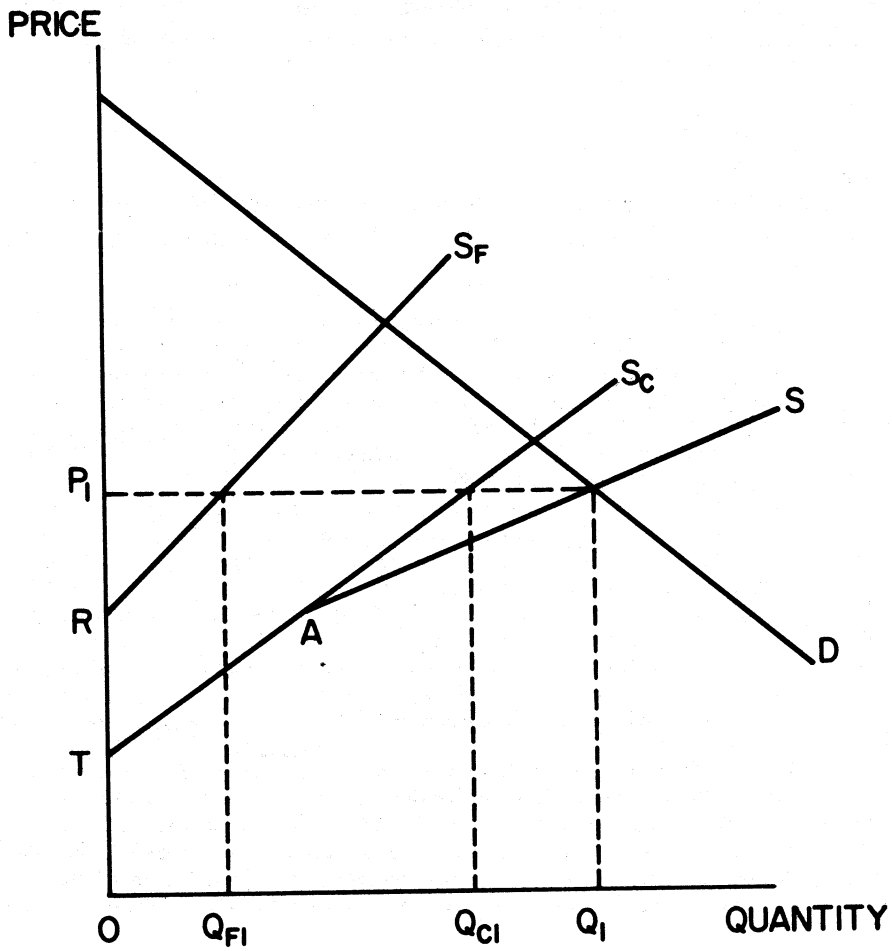


Figure 1

Assuming there is no collusion, the market will clear at a competitive price of P_1 and an output of Q_1 where demand equals industry supply.[84] We can see that the dominant firms produce Q_{C1} while the competitive fringe produces Q_{F1}. The sum of these two is necessarily Q_1:

$$Q_{C1} + Q_{F1} = Q_1$$

This graphical result follows directly from the construction of the industry supply curve.[85]

Suppose that the dominant firms decide to collude in an effort to increase their profits and that they decide to treat the fringe of competitive firms passively. Thus, their task is a difficult one: they must determine the optimal price to charge without having any control over the competitive fringe. One way of accomplishing that end is to determine the profit maximizing price while explicitly recognizing that the competitive fringe will respond to the collusive price as it would to a market-determined price. The way that such a price is determined is depicted in Figure 2.

We have reproduced the industry demand (D) and the three supply curves (S_F, S_C and S) of Figure 1. The dominant firms can estimate the quantities that the competitive fringe will produce at each alternative price that is established. The conspirators' demand is a residual demand that is determined by subtracting the competitive fringe supply (S_F) from the industry demand (D). At a price of P_3, for example, the competitive fringe will produce where S_F intersects D, thereby leaving nothing for the dominant firms to produce. At the other extreme, if the conspirators select a price equal to P_2 or less, the competitive fringe will produce nothing at all, thereby leaving all of the demand (D) for the conspirators. For all prices between P_3 and P_2, the competitive fringe supply must be subtracted from the industry demand to determine the quantity that the conspirators can sell. Thus, the net (or residual) demand available to the conspiring dominant firms is the industry demand less the competitive fringe supply:

---

84. The competitive behavior of all firms in the industry will lead them to supply Q_1 units of output at a price of P_1. Consumers will just be willing to take Q_1 units at a price of P_1. Thus, the intersection of supply and demand leads to an industry equilibrium in the sense that there is neither excess demand nor excess supply.

85. Recall that the industry supply curve (S) was obtained by adding the supply responses of the dominant firms (S_C) and those of the competitive fringe (S_F). Consequently, at any price the total quantity supplied will equal the sum of the quantities supplied by the dominant firms and by the fringe.

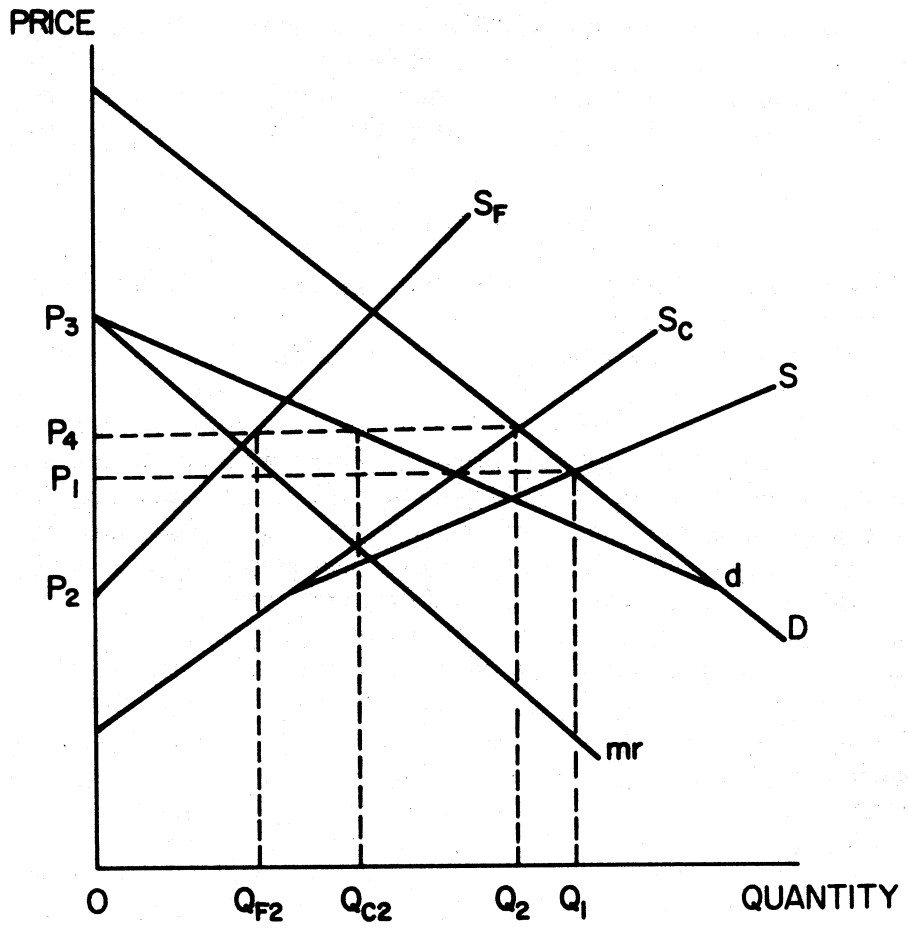


Figure 2

Net demand = D - S_F

This net demand is labelled P_3dD in Figure 2. The marginal revenue curve[86] associated with this net demand curve has been constructed and labelled mr.

If the conspirators are to maximize their profits, they must select an output where marginal revenue (mr) equals marginal cost (S_C).[87] When the dominant firms select their output in Figure 2

86. Marginal revenue is the change in total sales revenue when the quantity sold is changed by a small amount. For an excellent, concise discussion, see E. GELLHORN, ANTITRUST LAW AND ECONOMICS 89-91 (1976).

87. Marginal cost is the change in total cost when the quantity produced is changed by a small amount. Profit is defined as the difference between total revenue and total cost.

where mr intersects S_C, they will charge a price of P_4 and produce Q_{C2}. Comparing Figures 1 and 2, we see that the conspirators have restricted their output from Q_{C1} in Figure 1 to Q_{C2} in Figure 2. In contrast, the competitive fringe will respond to the higher price of P_4 by expanding its output to Q_{F2} from Q_{F1}. On balance, industry output will decline from Q_1 to Q_2. As a consequence, the market will clear at the price of P_4. Not surprisingly, the dominant firms will increase their profits as a result of their conspiratorial efforts. Interestingly, the competitive fringe is a beneficiary of the conspiracy as its profits also increase.

B. Impact on Competitive Fringe

The impact of the price conspiracy on a typical firm in the competitive fringe can be seen in Figure 3. Again it is first necessary to establish the competitive benchmark. Competitive firms behave as though their output decisions have no discernible impact on the market-determined price.[88] Consequently, they are characterized as price-takers and perceive demand as being horizontal at the market-determined price.[89] Prior to the dominant firms' price-fixing, the fringe firm's quest for profit led it to produce where marginal cost (MC) was equal to the competitive price (P_1). At that point, the firm produced q_1 and enjoyed profits equal to the difference between price and per unit, or average, cost (AC) times the quantity sold:

$$\text{profit} = (P_1 - AC)q_1$$

Given the positively sloped fringe supply curve (S_F) in Figure 1, we know that there is a continuum of fringe firms with respect to profitability. The marginal firm[90] earns no profit, that is its total

---

When marginal cost equals marginal revenue, the firm has produced the profit maximizing output. For output beyond that point, marginal cost exceeds marginal revenue, which means that the firm has increased total cost by more than the increase in total revenue. Consequently, profits are lowered due to the expanded output. If the firm produces less than that amount, marginal revenue will exceed marginal cost. This means that an expansion in output will increase total revenue more than it will increase total cost. Since that will increase total profit, profit is not maximized at lower outputs. *See* E. GELLHORN, *supra* note 86, at 52-54. Note that the industry's marginal cost curve is equivalent to the industry's supply curve. *Id.* at 55.

88. This behavior is the essence of a competitive market and the reason that competitive firms are described as price-takers.

89. When a firm is under the impression that its actions have no impact on price, it behaves as if it can sell all that it wants to sell at the market-determined price. Thus, that firm's demand curve is a horizontal line at a height equal to the market price.

90. The marginal firm just breaks even and earns a competitive return on its invest-

revenue is just sufficient to cover the opportunity costs[91] of the resources that it employs. All of the infra-marginal[92] firms earn positive profits to some degree.

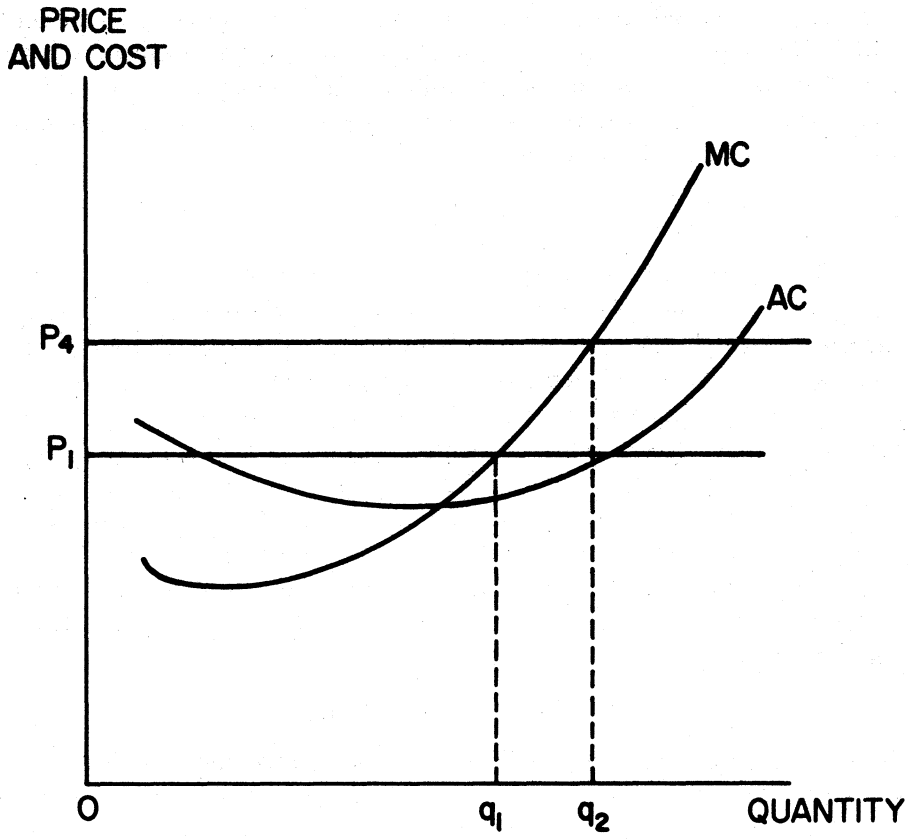


Figure 3

ment. All of the other firms are more efficient than the marginal firm, which means that they have lower production costs. As a result, each of them earns some profit beyond that which is necessary to keep them in the industry.

91. Opportunity cost is the cost imputed to owned resources that are employed by a firm. The value that is imputed is equal to the return that could be earned in the next best occupation.

92. The infra-marginal firms are those firms whose entry-inducing price is below the market price. When that occurs, the marginal firm is induced to enter.

The effect of the price fixing conspiracy is to raise prices from the competitive level of P_1 to the collusive level of P_4. When the conspiracy elevates price to P_4, the fringe firm will expand its output to the point where its marginal cost (MC) intersects the horizontal price line at P_4. Consequently, its output expands from q_1 to q_2. As all fringe firms do the same sort of thing, the collective effect is to raise their output from Q_{F1} to Q_{F2} in Figure 2. As shown by Figure 3, the gap between price and per unit cost (AC) has widened. Consequently, the typical fringe firm will enjoy greater profits under the umbrella price established by the conspiracy as per unit profit is greater and output is also greater.

C. *Impact on Customers of Fringe Firms*

It should be clear from the preceding analysis that all customers will pay the higher price. Figure 2 shows that the preconspiracy price and quantity were P_1 and Q_1, respectively. In contrast, Figure 2 identifies the collusive price and the resulting output as P_4 and Q_2, respectively. Each customer who buys at the collusive price of P_4 is injured by an amount equal to $P_4 - P_1$ per unit purchased, irrespective of the seller's identity. The fact that one customer bought from a conspiring firm while another bought from a nonconspiring fringe firm is immaterial. Every customer is paying precisely the same overcharge as a direct result of the conspiracy. The customer of the fringe firm unequivocally has been injured due to an antitrust violation.

IV. Implications of the Model for Causation

Courts that have addressed the antitrust standing of purchasers, who purchased from firms under the price umbrella established by illegal price-fixing, have reasoned that the chain of causation between price-fixing and the plaintiff's injury has been broken by the independent pricing decisions of the noncolluding seller.[93] It has been urged that a nonconspirator might not raise its price, but instead might set prices below the artificial umbrella raised by the defendants.[94] That reasoning, however, misunderstands the nature

93. *See supra* note 48.

94. Mid-West Paper Prods. Co. v. Continental Group, Inc., 596 F.2d 573, 583-84 (3d Cir. 1979); *In re* Coordinated Pretrial Proceedings in Petroleum Prods. Antitrust Litigation, 497 F. Supp. 218, 227 (C.D. Ca. 1980) ("The nonconspirator makes independent decisions concerning price and output. In order for a plaintiff to recover, the effect of the umbrella price on these pricing and output decisions would have to be shown.").

of competitive behavior. A firm in the competitive fringe does not engage in strategic behavior.[95] Instead, it accepts the prevailing price as an economic parameter beyond its influence. When the price that prevails increases for any reason, the fringe firm simply adjusts its output to the point where further increases are not profitable, that is, where marginal cost equals the new price.

Assume that a competitive fringe firm refused to increase its price following the creation of a higher umbrella price through the dominant firms' collusion. The firm in Figure 3 could not expand its output beyond q_1 because it would lose money on each additional unit sold. There would be tremendous pressure on the firm to expand output, however, because one segment of the industry (the conspirators) would now be charging a higher price, and everyone would want to buy from the competitive fringe at the lower original price. Consequently, many potential customers would have to be turned away because, as Figure 1 illustrates, the competitive fringe collectively could not supply more than Q_{F1} at a price of P_1.[96] In the ordinary course of events, those disappointed customers would bid up the prices of the competitive fringe firms. This process would expand the fringe supply until Q_{F2} was being supplied at the conspiratorial price of P_4. By raising their prices and expanding their output the fringe firms are acting competitively. This should not be construed as antisocial or illegal activity although it is directly caused by the dominant firms' illegal price-fixing.

Suppose that the competitive fringe simply refused for economically irrational reasons[97] to raise its price. Its supply would remain at Q_{F1}.[98] That decision would make the conspirators ecstatic because it would restrict industry output and permit the conspirators to raise prices above P_4 and earn even higher profits. This result can be seen in Figure 4, which contains the demand and supply curves from Figure 1.

If the competitive fringe refuses to adjust its price, and

---

95. Competitive fringe firms do not engage in strategic behavior because they are small enough to doubt that such behavior would have any favorable impact on price.

96. The quantity Q_{F1} represents the maximum amount that the fringe producers will provide at a price of P_1. A greater output would result in some of the fringe firms losing money. Consequently, we should not expect production beyond that point.

97. When an economic agent behaves in a way that is inconsistent with the pursuit of its own self-interest, that behavior can be characterized as economically irrational. That behavior does not necessarily have anything to do with psychological rationality.

98. The quantity supplied remains at Q_{F1} because we have assumed that the firms have not increased their prices. *See supra* note 86.

thereby cannot expand output beyond Q_1, the conspirators will face a residual demand curve equal to the market demand minus Q_{F1}. This net demand is labelled d and the associated marginal revenue is denoted as mr. Now the conspirators will produce an output of Q_{C4}, which is less than Q_{C2}, and will charge a price of P_5, which is higher than P_4. Consequently, the conspirators' profits are now higher than they were when the competitive fringe responded to a higher price. Of course, total output is now smaller, too. In this case, those who buy from the competitive fringe would suffer no injury, but those who purchase from the conspirators would have a larger injury equal to $P_5 - P_1$ per unit purchased.

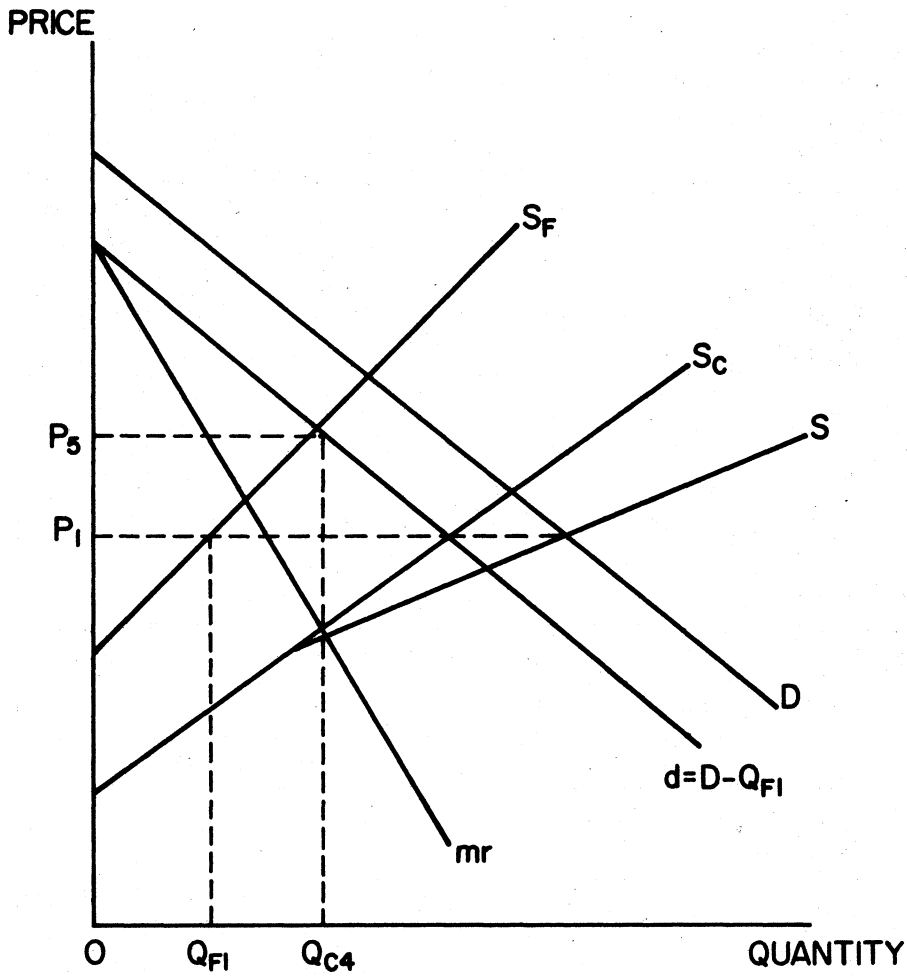


Figure 4

This analysis of umbrella price theory demonstrates that purchasers from the noncolluding competitors of the price-fixers suffer the same actual economic harm that is suffered by purchasers from the price-fixers. In both cases, the extent of the injury varies directly with the success of the price-fixing scheme and the willingness of all sellers to take the fixed price as given. Moreover, it is unreasonable to characterize the noncolluding seller's competitive behavior as a contributing cause of the plaintiff's injury. Rather, the noncolluding seller's competitive behavior is a predictable result of the defendant's illegal behavior, and is actually desirable from the standpoint of reducing the overall economic harm to the public from the illegal activity.

Implicitly following this umbrella price theory, the court in the *Pollock* case[99] found that the identity of the plaintiff buyer's sellers not only could not be ascertained but did not matter.[100] The umbrella pricing model demonstrates that the identity of the seller is no more important in a market in which individual sales are identifiable than it is in one in which they are not. In both situations, the question of economic injury is whether the buyer paid more for the product in a market in which the price-fixing scheme operated than it would have paid in that market without a price-fixing scheme.[101]

It should be noted that this model is not particularly useful where the buyer from the noncolluder and the buyer from the colluder have purchased different products,[102] or if they have purchased the same product in essentially different markets.[103] Similarly, the umbrella pricing model does not account for a causal

---

99. Pollock v. Citrus Assocs., Inc., 512 F. Supp. 711 (S.D.N.Y. 1981). *See supra* notes 58-68 and accompanying text. *Pollock* was decided under the Second Circuit's relatively narrow target area test. Presumably, it would have been decided the same way under the broader test of "foreseeable target area," *see supra* note 22, or the "zone of interest" test, *see supra* notes 23-25 and accompanying text.

100. 512 F. Supp. at 719. *See supra* notes 61-72 and accompanying text.

101. The plaintiff's burden of proof is identical in either instance. He must estimate the market price but for the conspiracy and establish that he has paid a higher price. This estimated overcharge is then multiplied by the quantity purchased to obtain an estimate of the economic injury suffered.

102. The dominant firm model assumes that the competitive fringe is producing the same product as the dominant firm(s). When the products are different, the buyer may still have suffered a substantial economic injury; however, the economic analysis of this situation is far more complicated and beyond the scope of this article.

103. The umbrella pricing model implicitly assumes that all sales have taken place in the same geographic market. When that is not the case, the economic analysis is more complicated. Nonetheless, it may still be possible to provide some convincing evidence of economic injury.

relationship between price-fixing and the price of a good that has a high degree of substitutability with the good that is subject to a price-fixing scheme.[104] Commentators have noted that price-fixing schemes succeed best in fungible goods markets[105] and, therefore, the applicability of the umbrella pricing theory usually will not be compromised by problems of market and product definitions.[106] That logic, however, suggests only that the problem will not arise often; it does not guide courts in ruling on standing questions where a plaintiff claims that the price paid for product X was affected by price-fixing of product Y. The model used in this article, and the arguments that flow from it, do not apply where proof of causation would require an analysis of the extent of product substitutability.[107] It is also inapplicable if proof of causation requires an analysis of the effect of price-fixing on sales in disparate geographic or product markets.[108] This proviso is consistent with *Illinois Brick*'s proscription of complex economic analysis in threshold standing determinations.

---

104. This is not to say that no causal relationship exists. The umbrella pricing model does not establish this causality; however, a different analysis is necessary.

105. Where goods are standardized and thus interchangeable or fungible, the market distinguishes among sellers mainly on the basis of price. Typically, in a fungible goods market there is a prevailing price that can be exceeded only at the risk of loss of market share. *See* R. POSNER, ANTITRUST LAW 59-60 (1976). For an interesting discussion, see Stigler, *A Theory of Oligopoly*, 72 J. POL. ECON. 44, 45-46 (1964). *See also* Hay & Kelley, *An Empirical Survey of Price-Fixing Conspiracies*, 17 J. LAW & ECON. 13, 15 (1974) (providing empirical evidence). In essence, fungibility simplifies the collusive price structure because differences in the products which cause equilibrium differences in prices need not be taken into account; those differences do not exist.

106. Comment, *Standing of Purchasers from Nonconspirators to Challenge Price-Fixing Conspiracy:* Mid-West Paper Prod. Co. v. Continental Group, Inc., 93 HARV. L. REV. 598, 606 (1979-80).

107. Questions of product substitutability involve complex issues of cross-elasticities. *See, e.g.*, United States v. E. I. duPont de Nemours & Co., 351 U.S. 377 (1956). If two products are substitutes and the producers of one collude on price, the demand for a substitute increases. The original purchasers of the substitute now will pay a higher price, thereby suffering an economic injury. The umbrella pricing model does not deal with this problem.

108. Buyers of the same product at different locations face a different complication. If geographic markets are disjointed in an economic sense, a conspiracy in one area will not taint the free market forces in the other area. But two apparently separate markets may be linked together by the buyers' search for the lowest price. A conspiracy in one market can cause a demand shift in the second market. That shift will injure the original customers in the second market. The umbrella pricing model does not deal with this problem. For an interesting attempt to deal with the vexing problem of geographic market definition, see Elzinga & Hogarty, *The Problem of Geographic Market Delineation in Antimerger Suits*, 18 ANTITRUST BULL. 45 (1973). For a suggestion that Elzinga and Hogarty have not been completely successful and for an alternative view, see Horowitz, *Market Definition in Antitrust Analyses: A Regression-Based Approach*, 48 So. ECON. J. 1 (1981).

V. RUINOUS DAMAGES

Several courts that have addressed antitrust standing issues have been impressed with the possibility that imposing damages on a defendant that reflect the total cost to the market of the defendant's behavior might destroy the defendant's business.[109] The availability of treble damages exacerbates the possibly ruinous effect of the judgment on the firm. Some courts also have been concerned with the equity of imposing damages for overcharges that did not accrue to the defendant.[110]

At least where the damages are a multiple of the ill-gotten gains of the defendant, however, it has been widely noted that treble damages rarely are ruinous. Instead, they tend to offset the economic benefits of the collusive activity that would not otherwise

---

109. *See* Jeffrey v. Southwestern Bell, 518 F.2d 1129, 1131 (5th Cir. 1975); Calderone Enters. v. United Artists Theatre Circuit, Inc., 454 F.2d 1292, 1295 (2d Cir. 1971), *cert. denied*, 406 U.S. 930 (1972). *See also* Berger & Bernstein, *supra* note 17, at 852. Berger and Bernstein argue that the destruction of the defendant firm removes a competitor in the market and permits further concentration of the market, and suggest "procedural adjustments" to avoid ruinous damage awards. The dissolution of the defendant firm does not necessarily result in injury to the market through increased concentration. To the extent it does permit greater concentration of the market, it may well be concentration by the fringe firms that were the innocent beneficiaries of the price-fixing scheme in the first instance. *Id.* at 852, 864.

110. *E.g.*, Mid-West Paper Prods. Co. v. Continental Group, Inc., 596 F.2d 573 (3d Cir. 1979). The *Mid-West Paper* court concluded:

> Allowing recovery for injuries whose causal link to defendants' activities is as tenuous as it is here could subject antitrust violators to potentially ruinous liabilities, well in excess of their illegally-earned profits, because . . . *price-fixers would be held accountable for higher prices that arguably ensued in the entire industry.*

Id. at 586 (footnotes omitted) (emphasis added). The dissent argued instead that price-fixers on whom the burden of paying damages to purchasers from competitors would fall would be "[c]ompanies who are best able to withstand such losses. . . . Thus the operation of the market would tend to prevent recoveries in suits such as this from being of a ruinous or anticompetitive dimension." *Id.* at 598 (Higginbotham, J., dissenting). Responding to that argument the majority claimed that it is quite possible for a firm with a low but substantial market share to be an industry price leader, advertently or inadvertently, but that "does not mean that it can withstand the burdens of a treble damage recovery that is based upon profits obtained by the rest of the industry." *Id.* at 586-87 n.51. The majority opinion's response confuses theories of price leadership and price-fixing; nevertheless, it is possible for the defendant to be the only conspirator sued and thus face the problems of contribution. The prospect of recovery by plaintiffs who have not purchased from any of the conspirators simply increases the risk and cost of the illegal activity.

See also In re Coordinated Pretrial Proceedings in Petroleum Prods. Antitrust Litigation, 497 F. Supp. 218, 227 (C.D. Ca. 1980) ("[W]here the plaintiff is allowed to recover from a defendant for excessive prices charged by a nonconspirator, the defendant is not disgorging illegally earned profits—those have gone to the competitor. There is a possibility for ruinous recovery in allowing treble damages to be awarded in such circumstances.").

be offset by a single damage recovery.[111] Even if the treble damage recovery exhausts the defendant's resources, or pushes the defendant over the brink of insolvency, it does not follow necessarily that competition will be injured by the demise of a single competitor. The firm's assets may be returned to productive and competitive activity under different ownership.[112] Additionally, the large recovery from one firm in an industry may deter collusive behavior by other firms and thereby enhance competition in that market.[113] It is difficult to assert with confidence that antitrust policy—or other social policy—is served by the continued existence in commerce of any particular firm, especially one guilty of collusive anticompetitive behavior.[114]

The likelihood of ruinous treble damage awards is apparently increased where recovery by nonprivy purchasers is allowed under the umbrella pricing theory, as those damages cannot be paid out

---

111. For a lucid discussion of the many factors that lower the costs to a colluder of obtaining these benefits, see Wheeler, *Antitrust Treble Damage Actions: Do They Work?*, 61 CALIF. L. REV. 1319, 1321-34 (1973). *See also* Tyler, *supra* note 21, at 287 (even if detected and successfully prosecuted, antitrust violations may prove profitable). *See generally* Blair, *supra* note 6 (the deterrence and compensation objectives of the antitrust laws are not adequately served by the treble damage remedy). Additionally, the treble damage remedy offsets to some extent the plaintiff's difficulty in bringing suit against a group of powerful colluders. *See* Brunswick Corp. v. Pueblo Bowl-O-Mat, Inc., 429 U.S. 477, 486 n.10 (1977).

112. *See* Tyler, *supra* note 21, at 289.

113. *See* Page, *supra* note 16, at 500. For a contrasting view, see K. ELZINGA & W. BREIT, THE ANTITRUST PENALTIES (1976) (in which the authors hypothesize that large treble damage recoveries provide a "perverse incentive" for customers to continue purchasing from monopolists).

The "perverse incentive" effect argued by Elzinga and Breit, could only become significant, however, if the probability of a judgment against the price-fixer were high and if the costs were low. In the real world, those transaction costs are very significant and the probability of a successful judgment is difficult to estimate. Thus, it is questionable whether the perverse incentive effect of the antitrust laws is ever reflected in a purchaser's behavior.

If is not clear what effect that price-fixing has on the price at which the buyer and seller transact, except that the price may be influenced by the buyer's expected return on an antitrust suit. This is the expected effect in a transaction between a colluding seller and a buyer who is operating under a "perverse incentive." Giving standing to the "umbrella" purchaser, however, should undo whatever "perverse incentive" there may be for dealing with the price-fixer rather than a noncolluder. Moreover, the operation of the "perverse incentive effect" in the transaction between the umbrella seller and purchaser should increase the cost of price-fixing in a way that could be ameliorated only by abandoning the price-fixing scheme and eliminating its effects from the marketplace.

114. The potentially ruinous effect of a treble damage recovery will be recognized by the plaintiff, who will have an incentive to settle for an amount or on terms that are within the means of the firm as an ongoing business. It has been noted, too, that the defendant who has engaged in anticompetitive behavior, resulting in widespread injurious effects, should not be protected by standing doctrine from liability "because of the outrageous scope of his own misconduct." Berger & Bernstein, *supra* note 17, at 868.

of the defendant's overcharges. The seeming inequity of imposing damages on price-fixers that are disproportionate to their illegal gains, however, is more apparent than real. There is no indication in section 4 and nothing that inheres in the treble[115] or punitive damages[116] theory which indicates section 4 damages should be proportionally related to the ill-gotten gains rather than to the injury sustained by the victim. The purpose of section 4 is not to achieve symmetry in the distribution of dislocated gains and losses, or to achieve restitution of ill-gotten gains, but to enforce the antitrust laws by compensating victims of antitrust violations[117] and by deterring potential violators.[118]

---

115. The language of section 4 refers to a remedy measured by the extent of the plaintiff's injuries: "[A]ny person . . . injured . . . shall recover threefold the damages by him sustained. . . ." 15 U.S.C. § 15 (Supp. IV 1980). The courts have separated the question of proving the fact of injury from the question of proving the amount of damages. In a price-fixing case, the plaintiff proves the amount of damages by showing the extent to which the price it paid exceeded the price that would have prevailed in the market in the absence of the price-fixing conspiracy. Usually, the damage amount must be estimated from pre- or post-conspiracy price data and other related information, but it is measured by the excessive price paid by the plaintiff. *See* P. AREEDA, *supra* note 15, at 78-79 and authorities cited therein. Damage estimation in antitrust cases generally raises a multitude of problems. *See generally* Hoyt, Dahl & Gibson, *Comprehensive Models for Assessing Lost Profits to Antitrust Plaintiffs*, 60 MINN. L. REV. 1233 (1976); Lanzillotti, *Problems of Proof of Damages in Antitrust Suits*, 16 ANTITRUST BULL. 329 (1971); Parker, *Measuring Damages in Federal Treble Damage Actions*, 17 ANTITRUST BULL. 497 (1972); Comment, *Proof Requirements in Antitrust Suits: The Obstacles to Treble Damage Recovery*, 18 U. CHI. L. REV. 130 (1950). *But see* Montreal Trading Ltd. v. Amax, Inc., 661 F.2d 864 (10th Cir. 1981), *cert. denied*, 102 S.Ct. 1634 (1982).

In many instances, the defendant's gains would not accurately depict the full extent of the social injury caused by their actions. In some price-fixing conspiracies, market shares are allocated through nonprice competition. When that is the case, some of the overcharge will be dissipated by expenditures on nonprice variables. Consequently, the gains to the defendants will understate seriously the extent of the damages.

116. Antitrust treble damages are recognized as punitive in nature. *See* P. AREEDA, *supra* note 15, at 75-76. Although the criteria for jury assessment of punitive damages are vague and inconsistent, and may include factors such as the wealth of the defendant or the egregiousness of the defendant's gain, historically punitive damages have been justified on the basis of punishment and deterrence rather than restitution.

117. Deterrence and compensation may be competing goals of antitrust law; however, granting standing to a plaintiff under a "price umbrella" theory would serve both goals. Where deterrence and compensation conflict, deterrence is considered the superior goal. *See* K. ELZINGA & W. BREIT, *supra* note 113, at 66. The authors argue that the goals of the antitrust laws require that deterrence be preferred because the achievement of deterrence could obviate the need for compensation, but not vice versa. The Supreme Court, however, has indicated the importance of compensation as a primary goal: "[T]he treble-damages provision, which makes awards available only to injured parties, and measures the awards by a multiple of the injury actually proved, is designed primarily as a remedy." Brunswick Corp. v. Pueblo Bowl-O-Mat, Inc., 429 U.S. 477, 485 (1977) (footnote omitted).

118. *See* American Soc. of Mechanical Eng'rs, Inc. v. Hydrolevel Corp., 102 S.Ct. 1935, 1947 (1982); Hawaii v. Standard Oil Co., 405 U.S. 251, 262 (1972); Bigelow v. RKO Radio

The economic analysis of umbrella pricing in Section III demonstrated that purchasers from noncolluding competitors of price fixers suffer exactly the same injury as direct purchasers from the conspirators.[119] Allowing recovery by these nonprivy purchasers would therefore satisfy the antitrust law's goal of compensation because the plaintiff would be compensated for a real, ascertainable loss that is part of the social interest that the law was intended to protect. The goal of deterring price-fixing could be served also, although the precise deterrent effect merits closer scrutiny. It is useful to separate analytically two dimensions of section 4 deterrence: the trebling of damages[120] and the basis for determining damages. Treble damages are designed to increase damages to a level that more than compensates the victim, and that imposes a penalty to deter the wrongful action.[121] Collusive activities such as

---

Pictures, Inc., 327 U.S. 251 (1976); Perma Life Mufflers v. International Parts Corp., 392 U.S. 134, 139 (1968). For a discussion of the importance of deterrence in the legislative history of the Sherman Act, see Page, *supra* note 16, at 473-74. *But see* Parker, *Treble Damage Actions—A Financial Deterrent to Antitrust Violations*, 16 ANTITRUST BULL. 483, 486-92.

Deterrence is also the primary goal of the criminal sanctions for antitrust violations. Violations of Sherman Act sections 1 and 2 may be punishable by imprisonment up to three years and fines up to $100,000 for an individual and up to $1 million for a corporation, or by both fine and imprisonment. *See* 15 U.S.C. §§ 1, 2 (1976). In addition, section 14 of the Clayton Act provides criminal penalties of up to one year imprisonment or a $5,000 fine, or both, for corporate directors, officers or agents involved in a corporation's violation of the criminal provisions of the Sherman Act. *See* 15 U.S.C. § 24 (1976). The Clayton Act penalties do not limit Sherman Act liabilities. United States v. Wise, 370 U.S. 405, 411-15 (1962). In general, the criminal sanctions are considered to be of less deterrent effect than the private action suit because of the infrequency with which they are invoked. *See* P. AREEDA, *supra* note 15, at 54-55; Baker & Reeves, *The Paper Label Sentences: Critique*, 86 YALE L.J. 619 (1979); Blair, *supra* note 6, at 57-59; Flynn, *Criminal Sanctions Under State and Federal Antitrust Laws*, 45 TEX. L. REV. 1301 (1967).

119. *See supra* section III C.

120. Treble damages as a remedy for restraint of trade dates to the English statute on monopolies. 1623, 31 Jac. 1, ch. 3, § IV. *See* ELZINGA & BREIT, *supra* note 113, at 63-96. For a legislative history of treble damages, see generally Bicks, *The Department of Justice and Private Treble Damage Actions*, 4 ANTITRUST BULL. 5 (1959); Loevinger, *Private Action—The Strongest Pillar of Antitrust*, 3 ANTITRUST BULL. 167 (1958); Parker, *supra* note 118.

121. *See* ELZINGA & BREIT, *supra* note 113 at 66:

> [I]f the appropriate purpose of the private treble damage suit is deterrence, the litigation of private suits should then be encouraged in the courts and the awarding of damages becomes paramount. Damages should be awarded to someone (so that a deterrent effect is manifested), even if that party is not the one specifically injured by the antitrust violation. However, if compensation is the goal, then those not injured by anticompetitive activity have no business in court (even if their suit attacks a bona fide antitrust violation), and a greater role for the government as the agent of enforcement is logical.

See Wheeler, *supra* note 101.

price-fixing involve deliberate decisionmaking, planning and coordination; they are rational responses to market opportunity and are, therefore, susceptible to rational deterrence. Moreover, these illegal activities are engaged in by employees several rungs down the organizational ladder. Thus, the firm must expend resources to monitor and educate those in the firm who are likely to decide whether or not to collude.[122] The potential for treble damages should motivate the firm to avoid collusive activity by increasing the costs of the activity. The rational firm may compare the potential benefits of collusive behavior with the potential damages, discounted by the probability of detection and successful prosecution.[123] In effect, the trebling of damages offsets to some extent the fact that many price-fixing schemes go undetected.

The question of what kinds of injury might be the basis for treble damages in collusive price-fixing suits involves a different aspect of deterrence. It is obvious that the prospect of recovery by purchasers from noncolluding competitors should have a greater deterrent effect than recovery limited to direct purchasers, assuming a constant probability of detection. The rational actor model[124] suggests that damages ought to be based on some measure greater than the defendant's gain in order to deter behavior which has a lower than one-in-three probability of detection. On the other hand, the prospect of unlimited damages may cause a risk averse firm to avoid activity that, in fact, does not violate antitrust law and may even be desirable.[125] The potential for unlimited damages

---

122. The deterrent effect may be diminished in the large corporate setting where the firm itself rather than the professional manager is likely to bear the entire brunt of the penalty. *See* Blair, *supra* note 6, at 63-65; Liman, *The Paper Legal Sentences: Critique*, 86 YALE L.J. 630, 631 (1977); Wheeler, *supra* note 111, at 37.

123. This assumption about the behavior of the firm assumes that the firm is a single decisionmaker that assesses information rationally and acts in accord with value-maximizing decisions. *See* Blair, *supra* note 6, at 62-63; Wheeler, *supra* note 111. *See generally* R. POSNER, ANTITRUST LAW: AN ECONOMIC PERSPECTIVE 221-24 (1976) (a prospective violator rationally weighs the cost of punishment against the probability that he will be caught); Becker, *Crime and Punishment, An Economic Approach*, 76 J. POL. ECON. 169, 176-85, 191-93 (1968).

124. *See supra* note 123. *See generally* C. STONE, WHERE THE LAW ENDS: THE SOCIAL CONTROL OF CORPORATE BEHAVIOR 48 (1975) (corporate managers do not seriously expect their antitrust violations to have an impact on their own financial status); Note, *Decisionmaking Models and the Control of Corporate Crime*, 85 YALE L.J. 1091, 1100-01, 1106-12 (1976) (prevention of corporate antitrust violations should focus on the individual corporate decisionmakers).

125. For example, a firm might cease or reduce the participation of its employees in trade or professional associations as a means of controlling exposure to price-fixing opportunities. A firm would take that step only if the costs of reduced participation did not exceed

introduces inefficiency by causing firms to expend resources that are not reasonably related to the social harm of price-fixing in monitoring firm behavior to avoid price-fixing.[126] Although the effects of a price-fixing conspiracy may reverberate throughout the economy,[127] it seems most reasonable to limit the type of injury for which treble damages may be awarded to injuries that are serious and definitely attributable to the defendant's actions. That approach would more effectively deter the firm and signal more precisely the proper level of expenditure and vigilance necessary to avoid collusive price-fixing activity.[128]

---

the perceived benefits of reduced exposure to liability. That perception would be a function both of the extent of the potential liability and the likelihood that participation would lead to liability; however, if the potential liability of the firm is disproportionately greater than the social injury caused by price-fixing, the firm may incur unnecessary and wasteful costs.

This problem should not be confused with the related problem of the firm abstaining from efficient behavior because of uncertainty in the law. *See* United States v. United States Gypsum Co., 438 U.S. 422 (1978); Breit & Elzinga, *Antitrust Penalties and Attitudes Toward Risk: An Economic Analysis*, 86 HARV. L. REV. 693 (1973); Page, *supra* note 16, at 472 ("[A]n excessive penalty may deter . . . efficient activities that firms only perceive as violations because the law is unclear. . . . Deterrence of the conduct by penalties unrelated to the social cost would itself be inefficient."). But this concern does not pertain to price-fixing, which has no efficiency-based redeeming virtue.

126. *See* Page, *supra* note 16, at 475:

> Treble damage awards will deter the conduct penalized regardless of the rationale for imposing the award. If the size of the penalty is unrelated to the anticompetitive effects of the conduct involved, the damage award will either leave incentives to engage in anticompetitive conduct or create deterrents to efficient conduct.

Note that the injury to the individual plaintiff in a section 4 action bears no necessary relationship to the social harm of the unlawful activity, except that the social harm is, presumably, no less than the injury to the plaintiff. Penalties disproportionate to the plaintiff's injury could violate this principle.

127. In a general equilibrium sense, a conspiracy to fix the price of any product will have repercussions throughout the economy. Obviously, considerations of judicial manageability forbid including everyone as a potential plaintiff. As a very simple example, suppose that the price of footballs were fixed, reducing the demand for pigskin and leading to a higher price for ham. As a consequence, the demand for beef and chicken would rise, thereby leading to higher prices. The demand for mustard would fall along with its price. Should consumers of beef and chicken and sellers of mustard have standing to sue the football price-fixers? Practical grounds preclude the maintenance of such actions, although one cannot deny that those groups have been injured. For a discussion of demand and supply of joint products, see M. FRIEDMAN, PRICE THEORY 153-65 (1976).

128. The approach is consistent with efficiency analysis that seeks an equilibrium between the expected loss to the defendant (as measured by the expected damage recovery (cost) times the probability of recovery) and the cost of avoidance. This analysis was formulated by Judge Learned Hand in United States v. Carroll Towing Co., 159 F.2d 169 (2d Cir. 1947). The analysis has also been developed extensively in the modern law and economics literature.

VI. Conclusion

In spite of efforts of the lower federal courts to develop comprehensive doctrines of antitrust standing, it is uncertain whether a purchaser under a price umbrella has section 4 standing to sue the price-fixers responsible for the umbrella. From the standpoint of both antitrust policy and widely accepted economic theory, it seems clear that purchasers from fringe competitors under a price umbrella are indistinguishable in theory, and sometimes in fact, from purchasers from the price-fixers. Principles of antitrust policy do not require that they be treated differently. In fact, important policies of section 4 and antitrust law would be advanced by granting standing to these plaintiffs. When analyzed properly, the most frequent objections to granting standing to the plaintiffs—the prospect of complex and tenuous proof of causation and the potential for ruinous damages—provide an insubstantial basis for excluding this group of plaintiffs. Courts should look closely at the substance of the "umbrella" plaintiffs' claims and at the infirmities in economic and antitrust theory that are reflected in the principal objections to the granting of standing to these plaintiffs.

Constitutionality of No Fault Jurisprudence*

*Josephine Y. King***

TABLE OF CONTENTS

** Professor of Law, Pace University School of Law.

A. THE CONTEXT

I. INTRODUCTION

Major constitutional challenges to no fault insurance laws were decided by the highest state court in approximately one-half of the nation's no fault jurisdictions in the 1970's. In some instances, litigation was initiated before the effective date of the statute or within a short period of time after it became effective. The speed with which the challenges were launched indicates the correct perception of the legal profession that the development of no fault insurance represents a profound change in traditional tort law principles and litigation.

To assert that the constitutionality of no fault law is completely settled is to concede or to claim too much. Thus far, the cases in most instances have been broadside attacks. Plaintiffs have invoked federal and state constitutional guarantees—due process, equal protection, jury trial, access to courts—often in a blunt, generalized fashion. This approach characterizes the first wave of decisions in the period between 1971 and 1976. A second wave, or second generation of challenges, however, appears to have been initiated by *Shavers v. Kelley*[1] in Michigan, in which the court pursued more subtle issues with reverberations affecting the insurance industry and its supervision by state administrative agencies.

Assessing constitutional developments in no fault jurisprudence for the future as well as the present requires reference to the broader area of insurance law of which no fault is a part. Constitutional challenges of insurance industry practices, and of state insurance regulation, have proceeded along two tracks.

The first and earlier track engages the commerce power and meets issues such as whether insurance involves interstate commerce, and the nature and extent of federal regulation of the insurance industry. That track is almost dormant today. The second track opens state legislative and, secondarily, administrative insurance regulation to penetration by the fourteenth amendment requirements of due process and equal protection. Two provocative queries surface: Will the commerce track be reactivated in the future? Has the second, individual rights track been extended to probe new territories?

A brief examination of the constitutional history of both the

1. 402 Mich. 554, 267 N.W.2d 72 (1978), *cert. denied*, 442 U.S. 934 (1979).

commerce clause track and the individual rights track is necessary before turning to a detailed analysis of individual rights track litigation in the States. Analysis of the individual rights track litigation will focus on the arguments that have recently been utilized to attack the validity of separate no fault legislation in Illinois, New Hampshire, Massachusetts, Florida, Kansas, Connecticut, Kentucky, Pennsylvania, New Jersey, New York, and Michigan. Notwithstanding the variations in statutory language, the question in each case is whether the state no fault legislation has infringed on federal due process or equal protection rights. Major issues raised by these cases include: (1) whether denying a cause of action to plaintiffs who have not suffered the requisite monetary damage or personal injuries violates constitutional equal protection guarantees by setting up unreasonable and arbitrary classifications; (2) whether the legislature must supply a reasonable alternative if it abolishes a right of access to the courts; (3) whether the no fault statutes deprive plaintiffs of rights protected by procedural or substantive due process; and (4) whether the no fault statutes violate various state constitutional provisions. In examining the Michigan no fault experience, the analytic approach must be broadened because of the challenge to that state's ratemaking regulations and practices. In analyzing the constitutional attacks on state no fault statutes, this article will also examine the various standards of constitutional review employed by individual state courts.

II. The Commerce Power and Insurance Regulation

A. *The* South-Eastern Underwriters *Case (1944)*

Over a century ago, the Supreme Court held in *Paul v. Virginia*[2] that insurance did not involve interstate commerce and that state regulation of insurance companies was not pre-empted by the federal commerce power. In *Paul*, the Court upheld a Virginia statute regulating foreign insurance companies,[3] resting its decision on a legal concept that an insurance policy was a simple contract of indemnity and not a commodity traded in the market-place.[4] The sociologic and economic explanation for the *Paul* decision, however, could be found in the need to correct abuses in the insurance business; since Congress had not regulated the industry, state in-

2. 75 U.S. (8 Wall.) 168 (1869).
3. *Id.* at 185.
4. *Id.* at 183.

tervention was a necessity.[5]

Paul remained firmly established until 1944, when the commerce line of analysis suddenly emerged from seventy-five years of darkness in *United States v. South-Eastern Underwriters Association.*[6] In *South-Eastern*, the Justice Department obtained an indictment against the defendant and its members for violations of the Sherman Antitrust Act.[7] The defendants were charged with restraining trade and commerce by fixing noncompetitive premium rates for fire and allied lines of insurance, and monopolizing trade and commerce in those insurance lines.[8] Justice Black, writing for the Court, framed two questions: Did the Sherman Act prohibit conduct that restrained or monopolized interstate trade in fire insurance? If so, did the defendant's transactions constitute "commerce among the several states"?[9]

Addressing the second issue, the Court characterized the insurance business as a "continuous and indivisible stream of intercourse among the states . . . ."[10] The Court based this conclusion on the express congressional intent embodied in the Sherman Act. It held that the Sherman Act's comprehensive language rendered the Act applicable "to all combinations of business and capital organized to suppress commercial competition."[11] Thus, the insurance industry was abruptly exposed to the federal commerce power.

B. The McCarran-Ferguson Act

Congress responded to the *South-Eastern* decision by passing the Insurance Antitrust Moratorium Act (McCarran-Ferguson Act).[12] The Act provided that the Sherman, Clayton, Robinson-Patman and Federal Trade Commission Acts would not apply to

---

5. Stern, *The Commerce Clause and the National Economy 1933-1946* (Pt. 2), 59 HARV. L. REV. 883, 909 (1946). For an intriguing account of how the insurance industry and the Court viewed regulation in the nineteenth and first half of the twentieth century, see *id.* at 909-25.

6. 322 U.S. 533 (1944).

7. Ch. 647, 26 Stat. 209 (1890) (current version at 15 U.S.C. §§ 1,2 (1976)).

8. 322 U.S. at 534-35. It was alleged that through concert of action the member companies controlled 90% of fire insurance in the Southeast, fixed premiums, and engaged in boycotts and coercive actions against non-member companies. *Id.* at 535.

9. *Id.* at 538.

10. *Id.* at 541.

11. *Id.* at 553.

12. Ch. 20, §§ 1-5, 59 Stat. 33, 34 (1945) (current version at 15 U.S.C. §§ 1011-1015 (1976)).

the insurance industry until January 1, 1948.[13] After that date, those Acts would "be applicable to the business of insurance to the extent that such business is not regulated by State law."[14]

The National Association of Insurance Commissioners and the All-Industry Committee promptly drafted All-Industry Bills establishing state regulatory mechanisms for ratemaking. Almost all states adopted some form of the All-Industry Bill that included a rate approval standard for state agencies.[15] Those regulations usually required that insurance rates not be "excessive, inadequate, or unfairly discriminatory."[16] The speed and solidarity with which the insurance companies responded to the *South-Eastern* decision reflected their aversion to federal regulation.[17]

Prevention of monopolies by enacting antitrust laws is only one facet of congressional power under the commerce clause. That power has flourished and enjoyed expansive construction by the Supreme Court. In light of *South-Eastern*'s holding that insurance involved interstate commerce, the subsequent antitrust moratorium would not appear to be a *general exemption* from the expansive federal commerce power. Many practices of the industry and also many state regulations could affect or burden interstate commerce, and such practices and regulations might be open to attack on constitutional grounds. Yet significant constitutional challenges

---

13. In 1947 the moratorium was extended to June 30, 1948. *See* Act of Mar. 9, 1945 ch. 326, 61 Stat. 448 (1947) (current version at 15 U.S.C. §§ 1012(b), 1013(a) (1976)).

14. McCarran-Ferguson Act, ch. 20, §§ 2, 3, 59 Stat. 33, 34 (1945) (current version at 15 U.S.C. §§ 1011-1015 (1976)). *See supra*, note 13. For a recent interpretation of the McCarran-Ferguson Act's exemption of the "business of insurance," see Group Life & Health Ins. Co. v. Blue Shield, 440 U.S. 205 (1979). In *Blue Shield*, the agreements between Blue Shield and pharmacies, which required insureds to pay $2 for prescription drugs with the remainder of the cost to be paid by Blue Shield to participating pharmacies, were held to be not exempt from antitrust laws. The Court stated that such agreements do not involve underwriting of risks "but are merely arrangements for the purchase of goods and services by Blue Shield." *Id.* at 214.

15. *See* Kimball & Boyce, *The Adequacy of State Insurance Rate Regulation: The McCarran-Ferguson Act in Historical Perspective*, 56 MICH. L. REV. 545, 555-56, (1958). For an account of this period and valuable information and insights relating to the conduct of the insurance business, see ANTITRUST SUBCOMM. OF THE HOUSE COMM. ON THE JUDICIARY, AUTOMOBILE INSURANCE STUDY, H. REP. No. 815, 90th Cong., 1st Sess. 19-24 (1967).

16. *See* Gardner, *Insurance and the Anti-Trust Laws—A Problem in Synthesis*, 61 HARV. L. REV. 246, 260-65 (1948).

17. A parallel may be observed in the more recent espousal by many insurers of state no fault legislation under threat of possible Congressional intervention. *See Investigation of Auto Insurance: Hearings Before the Subcomm. of Consumer Affairs of the Senate Comm. on Commerce*, 90th Cong., 2d Sess. 111-15 (1968) (statement of Professor Jeffrey O'Connell); King, *The Insurance Industry and Compensation Plans*, 43 N.Y.U. L. REV. 1137, 1146-47 (1968).

have not developed. The mere presence of state regulation triggers the McCarran-Ferguson Act's immunity, which has insulated the insurance business not only from the antitrust laws but also from the broader dimensions of the federal commerce power.

III. The Individual Rights Track

A. *Early History*

Although the commerce track of challenges to state insurance regulation appears abandoned and dull from disuse, the individual rights track utilized in early workers' compensation cases has acquired additional popularity in recent no fault litigation. Workers' compensation acts effected a major change in the common law by substituting a system of liability without fault in place of the traditional tort remedy. Early opponents of this legislatively imposed system of no fault employee compensation for work-related injuries relied on the fourteenth amendment's incorporation of various constitutional rights, as well as the basic argument that state legislatures may not abridge or abolish traditional common law rights.

B. *The* White *Case*

In *New York Central Railroad Company v. White*,[18] an employer objected to the obligations imposed by New York's Workmen's Compensation Act. New York Central argued that the act deprived it of property without due process, violated equal protection by exempting certain classes of workers, deprived both employer and employee of property and liberty to make an employment contract incorporating terms of their own choosing, infringed employees' liberty to sue in tort, violated the right to trial by jury, and exceeded the state's police power.[19]

The United States Supreme Court upheld the New York statute after an intensive analysis and formulated principles of profound significance for the exercise of legislative regulation. The Court decided that New York's statute was a reasonable exercise of state police power.[20] The Court also ruled that both the states and Congress had authority to depart from common law rules affecting

18. 243 U.S. 188 (1917).
19. *See id.* at 196-208.
20. *Id.* at 206.

employers' liability:[21] "No person has a vested interest in any rule of law entitling him to insist that it shall remain unchanged for his benefit."[22] Although both employer and employee ceded certain rights, they obtained new benefits in return.[23] Moreover, the statute's denial of trial by jury did not violate due process and its exclusion of farm laborers and domestic servants did not violate equal protection guarantees by establishing arbitrary classifications.[24]

This sweeping affirmation of legislative power to alter common law negligence principles was repeated two years later in *Arizona Employers' Liability Cases*,[25] where the Court declared that "[t]he States are left with a wide range of legislative discretion, notwithstanding the provisions of the Fourteenth Amendment; and their conclusions respecting the wisdom of their legislative acts are not reviewable by the courts."[26]

C. *Other Relevant Decisions*

In *Silver v. Silver*,[27] the Court considered a different variety of regulatory statute—Connecticut's guest passenger law.[28] That statute denied a guest passenger a cause of action against the owner or operator of a vehicle unless the guest was injured as a result of the host's intentional or reckless conduct.[29] The Court found that the fourteenth amendment equal protection guarantee was not violated by distinguishing between passengers in automobiles and those occupying other types of vehicles. The Court deemed it constitutionally sufficient that Connecticut perceived an evil to be corrected—vexatious or collusive litigation—and chose to address it.[30]

A motorist's individual rights in maintaining an operator's license was the subject of another legislative regulatory decision in

---

21. *Id.* at 200 (citing Second Employers' Liability Cases, 223 U.S. 1 (1912)).
22. *Id.* at 198.
23. *Id.* at 201.
24. *Id.*
25. 250 U.S. 400 (1919).
26. *Id.* at 419.
27. 280 U.S. 117 (1929).
28. *Id.* at 121-22 (citing 1927 Conn. Pub. Acts, ch. 308).
29. *See id.* at 119.
30. *See id.* at 122-23. The Court stated, "[T]here is no constitutional requirement that a regulation, in other respects permissible, must reach every class to which it might be applied—that the legislature must be held rigidly to the choice of regulating all or none." *Id.* at 123.

Bell v. Burson.[31] *Bell* involved provisions of Georgia's Motor Vehicle Safety Responsibility Act[32] that imposed suspension of the operator's license and of the vehicle registration of the operator and owner of any vehicle involved in an accident unless security was furnished to satisfy any possible judgment for damages, and proof of future financial responsibility was established.[33] Bell insisted that his license should not have been suspended without an administrative hearing affording him an opportunity to establish his freedom from fault. The United States Supreme Court agreed, holding that suspension of driving privileges without considering the licensee's evidence on fault was a denial of procedural due process.[34] The Court indicated that due process could be satisfied by an inquiry into whether or not a reasonable possibility existed of a judgment against the licensee. The opportunity to determine this question, however, would have to precede suspension of a license or registration.[35]

The rationale of *Bell* has not been adapted to no fault litigation. *Bell* was decided in an era of sensitivity to the needs of economically disadvantaged classes;[36] it dealt with an interest related to subsistence and earning a livelihood—the operation of a motor vehicle. The reasoning appears to be that once a license is granted, the licensee possesses an entitlement—an important interest in maintaining his legal ability to drive. The Court ruled that imposition of a penalty for failure to comply with the condition of licensure must be attended by due process. Although *Bell* involved a condition subsequent to licensure, its reasoning raises the question whether the same due process requirements operate when conditions precedent to lawful operation or ownership of a motor vehicle are imposed, such as compulsory liability and no fault insurance.[37]

---

31. 402 U.S. 535 (1971).

32. GA. CODE ANN. § 92A-601 to -621 (1958) (repealed 1977).

33. 402 U.S. at 536 n.1 (citing GA. CODE ANN. § 92A-605 (Supp. 1970)).

34. *Id.* at 542-43. Justice Brennan wrote for the Court that a driver's license, may become essential in the pursuit of a livelihood. Suspension of issued licenses thus involves state action that adjudicates important interests of the licensees. The licenses are not to be taken away without that procedural due process required by the fourteenth amendment. This is but an application of the general proposition that relevant constitutional restraints limit state power to terminate an entitlement whether the entitlement is denominated a "right" or a "privilege."

Id. at 539 (citations omitted).

35. *See id.* at 540, 542.

36. *See, e.g.*, Goldberg v. Kelly, 397 U.S. 254 (1970); Sniadach v. Family Finance Corp., 395 U.S. 337 (1969).

37. Michigan comes closest to using *Bell* as a due process wedge into entrenched in-

D. Standard of Constitutional Review

While state courts possess the power to decide federal constitutional questions, it is the United States Supreme Court that ultimately decrees the federal constitutional principles governing judicial oversight of the legislative branch. Since no fault contests have persistently involved the due process and equal protection clauses of the fourteenth amendment (as well as counterparts in state constitutions), state courts have had to follow recent developments in constitutional law to identify the appropriate standard for reviewing state legislation regulating the rights of individuals and industries. Current constitutional doctrine recognizes three levels of review that have evolved through a long series of decisions dating from the post-Civil War period.[38]

For a few decades following the Civil War the Supreme Court upheld state economic regulation.[39] A dramatic reversal occurred towards the end of the nineteenth century and continued into the mid-1930's. During those decades, the Court repeatedly invalidated federal and state economic regulation under the guise of substantive due process.[40] In that era, the Court disagreed not only with the means selected by legislatures for effecting certain public purposes, but also with the purposes themselves.[41]

The discrediting of substantive due process initiated a period of tolerance and deference in which the Court applied a "rationality" test in judging due process and equal protection attacks on legislative regulation.[42] This still valid test requires that legislative means be rationally related to a permissible governmental objective. A presumption of constitutionality cloaks legislative acts under the rationality standard, and the court generally forbears evaluation of the wisdom of the goals or the efficacy of the

---

surance industry and administrative agency practices. *See* notes *infra* 617-22 and accompanying text.

38. The history of the Court's role in reviewing legislation up to the Second World War is brilliantly recounted by the late Justice Robert H. Jackson. *See* R. JACKSON, THE STRUGGLE FOR JUDICIAL SUPREMACY (1949).

39. *See, e.g.*, Munn v. Illinois, 94 U.S. 113 (1876) (state regulation of the warehouse industry).

40. This period is known as the "Lochner Era," memorializing Lochner v. New York, 198 U.S. 45 (1905).

41. In effect, the court substituted its social values for the judgment of legislators. B. F. WRIGHT, THE GROWTH OF AMERICAN CONSTITUTIONAL LAW 154 (1942); *see* W. LOCKHART, Y. KAMISAR & J. CHOPER, CONSTITUTIONAL LAW 527-29 (5th ed. 1975).

42. McGowan v. Maryland, 366 U.S. 420, 425-26 (1961); Williamson v. Lee Optical, 348 U.S. 483, 488-91 (1955).

means.[43]

As a further reaction to substantive due process, doctrinal development in the arena of individual rights has focused on substantive equal protection. The Warren Court in the 1960's emphasized a "new" equal protection demanding a distinctly elevated level of review where fundamental rights or suspect classifications were at issue. The Court defined certain rights as fundamental[44] and certain classifications as suspect.[45] Legislation infringing on fundamental rights or employing suspect classifications was subject to a "strict scrutiny" test which required that the selected means be necessary to the accomplishment of a compelling state interest.[46]

The third tier of review has been described as the "newer" equal protection.[47] It involves an intermediate level of scrutiny focusing on *means* when the issue is discrimination based on gender, illegitimacy or alienage.[48] But there is difficulty in separating all due process and equal protection challenges into two or three discrete compartments, and in finding consistency in the content and

---

43. *See* Bennett, *"Mere" Rationality in Constitutional Law: Judicial Review and Democratic Theory*, 67 CALIF. L. REV. 1049 (1979); Note, *Legislative Purpose, Rationality, and Equal Protection*, 82 YALE L. J. 123 (1972). For an example of the "hands off" policy of the Supreme Court, see United States v. Carolene Products Co., 304 U.S. 144, 152-54 (1938). More recently, in Ferguson v. Skrupa, 372 U. S. 726 (1963), the Supreme Court declared:

> The doctrine that prevailed in *Lochner, Coppage, Adkins, Burns* and like cases—that due process authorizes courts to hold laws unconstitutional when they believe the legislature has acted unwisely—has long been discarded. We have returned to the original constitutional proposition that courts do not substitute their social and economic beliefs for the judgment of legislative bodies, who are elected to pass laws.

Id. at 730.

44. *See* Shapiro v. Thompson, 394 U.S. 618, 629-31 (1969) (the right to travel); Harper v. Virginia Bd. of Elections, 383 U.S. 663 (1966) (the right to vote). For an analysis of various theories treating the source and reviewability of fundamental rights, see Brest, *The Fundamental Rights Controversies: The Essential Contradiction of Normative Constitutional Scholarship*, 90 YALE L.J. 1063 (1981).

45. Race, for example, was indisputably a suspect classification. *See, e.g.*, Loving v. Virginia, 388 U.S. 1, 11 (1967); McLaughlin v. Florida, 379 U.S. 184 (1964); Strauder v. West Virginia, 100 U.S. 303, 310 (1879).

46. *See* Gunther, *Forward: In Search of Evolving Doctrine on a Changing Court: A Model for a Newer Equal Protection*, 86 HARV. L. REV. 1, 8, 21 (1972) (such scrutiny was "'strict' in theory and fatal in fact."

47. *Id.* at 12.

48. *See* Lalli v. Lalli, 439 U.S. 259 (1978) (Trimble v. Gordon, 430 U.S. 762 (1977); Mathews v. Lucas, 427 U.S. 495 (1976) (illegitimacy); Craig v. Boren, 429 U.S. 190 (1976) (gender). *Contra* Ambach v. Norwick, 441 U. S. 68 (1979) (alienage); Sugarman v. Dougall, 413 U.S. 634 (1973). This doctrine of rationality with a "bite" has emerged from the Burger Court as a third standard in reviewing the constitutionality of legislation. *See* Gunther, *supra* note 46.

application of the various standards.[49]

Thus, a state or federal court in reviewing fourteenth amendment objections to state no fault legislation may draw upon any of the above standards. The courts have generally applied a *mere rationality* test, in some instances with a decisional gloss drawn from state law.[50] Each of the major cases is therefore examined in its own environment of state constitutional and statutory law under the overarching federal standards of judicial review.

B. THE STATES AND NO FAULT JURISPRUDENCE

IV. ILLINOIS

A. *Introduction*

The Illinois Plan enacted in 1971[51] was the earliest no fault scheme wholly rejected by the courts.[52] The act did not mandate purchase of liability insurance by motor vehicle owners; it required instead that all insurers issuing automobile liability policies in Illinois include first party coverage for certain benefits.[53] Those bene-

---

49. *See* Rostker v. Goldberg, 453 U. S. 57 (1981) (involving the registration of males but not females under the Military Selective Service Act). While acknowledging the heightened scrutiny standard for gender-based classification in *Rostker*, Justice Rehnquist appears to apply the merest rationality test with great deference to Congress' powers in military affairs. *Id.* at 64-65. These anomalies and inconsistencies have drawn sharp criticism from Justice Marshall who has advocated a "sliding scale" approach to determine the degree of scrutiny necessary for the protection of those interests which are not textual, constitutional rights. The sliding scale or nexus theory, however, has not been followed by a majority of the Supreme Court. *See* San Antonio School Dist. v. Rodriguez, 411 U.S. 1, 102-03 (1973) (Marshall, J. dissenting). *See also* Richardson v. Belcher, 404 U.S. 78, 90 (1971) (Marshall, J. dissenting); Dandridge v. Williams, 397 U.S. 471, 519-21 (1970) (Marshall, J. dissenting).

50. There is occasional reference in state court decisions to the less restrictive alternative principle as a test to be applied to state regulation. The principle means that "an economic regulation violates due process if the government has a less restrictive alternative—that is, if the government can achieve the purposes of the challenged regulation equally effectively by one or more narrower regulations." Struve, *The Less-Restrictive-Alternative Principle and Economic Due Process*, 80 HARV. L. REV. 1463, 1463 (1967). Application of this principle would effect a tightening of the rationality requirement, allowing courts to evaluate substantively the possible choices available to the legislature.

51. Compensation of Automobile Accident Victims Act, Pub. L. No. 77-1430, 1971 Ill. Laws 2542 (repealed 1975).

52. *See* Grace V. Howlett, 51 Ill. 2d 478, 283 N.E.2d 474, 478 (1972). *See infra* notes 62-77 and accompanying text.

53. 1971 Ill. Laws at 2542. The coverage included medical, hospital and funeral expenses incurred within one year of an accident, with a maximum of $2000 for each individual; income loss calculated as 85% of earnings, but not to exceed $150 a week for 52 weeks; replacement household services where the injured was not an employee, the maximum reim-

fits applied to the named insured, family household members, occupants of the insured vehicle and pedestrians struck by the insured vehicle.[54] With minor exceptions, such as an offset for state and federal workers' compensation act payments, the no fault law did not disturb the operation of the collateral source rule.[55]

The uniqueness of the Illinois Plan did not inhere in its benefit provisions, but rather in its limitations on damages, and in the inclusion of mandatory arbitration for small claims. In any tort action brought for bodily injury or death arising out of an accident involving the ownership, operation, maintenance or use of a motor vehicle, damages for pain, suffering and mental anguish could not exceed fifty percent of reasonable medical expenses if those expenses were $500 or less, or a sum equal to those expenses if they exceeded $500.[56] These limitations were inapplicable if the accident resulted in death, dismemberment, serious disfigurement or permanent disability.[57]

The Illinois act also mandated arbitration of disputed no fault claims in counties with a population of 200,000 or more, and permitted arbitration in smaller counties of all cases where the amount in controversy did not exceed $3,000.[58] Either party could appeal the arbitration award.[59]

The legislature viewed both the benefit provisions of the act and its limitations on tort damages as integral components of a unitary program to correct inefficiencies and inequities in the existing system.[60] The severability section of the act expressly provided that invalidation of the limitations on tort damages would nullify the entire no fault statute.[61] The tort liability limitations, therefore, were an obvious focal point for constitutional challenge.

---

bursement to be $12 a day for 365 days; survivors' benefits equal to income loss benefits if the injured person dies within one year of the accident. *See id.* at 2542-43. Every insurer subject to the act was required to offer excess loss coverage with minimum limits of $50,000 per person and $100,000 per accident that the insured could reject. *Id.* at 2543.

54. *Id.* at 2542-43.

55. For a concise description of collateral source recovery and the problems it causes, see R. KEETON, BASIC TEXT ON INSURANCE LAW §§ 5.5(c), 5.9 (1971).

56. 1971 Ill. Laws at 2548.

57. *Id.*

58. *Id.*

59. The appeal would be heard de novo in the court in which the case was pending. *See id.* at 2549.

60. *See* Grace v. Howlett, 51 Ill. 2d 478, 283 N.E.2d 474, 478 (1972).

61. 1971 Ill. Laws at 2551.

B. Grace v. Howlett

In *Grace v. Howlett*,[62] a deluge of constitutional objections descended upon the Illinois Plan. The plaintiff charged that the plan violated his due process and equal protection rights, deprived him of his right to trial by jury, and also violated the separation of powers doctrine and the Illinois proscription against special legislation.[63] The Illinois Supreme Court examined each of these objections.

1. Equal Protection—The plaintiff challenged the act's classifications—inclusions and exclusions—and the inequities of the measure of damages.[64] The general theory of the plaintiff's case appeared to be a claim of denial of equal protection. The act created two classes of accident victims distinguished by the type of vehicle they encountered. Only automobiles and small utility vehicles used as private passenger automobiles[65] were required to carry first party insurance.[66] Because Illinois was not a compulsory insurance state, the availability of no fault benefits to some accident victims and not to others appeared discriminatory. A pedestrian hit by a private passenger vehicle would get "prompt payment of his medical and other expenses," but a pedestrian injured by a rented car, bus, lawyer's, physician's or salesman's car would have no assurance of receiving first party benefits because owners of vehicles used for commercial and professional purposes could elect to

---

62. 51 Ill. 2d 478, 283 N.E.2d 474 (1972). The case was brought as a taxpayer's suit before the effective date (January 1, 1972) of the no fault statute. The complaint presented a facial constitutional challenge that sought to enjoin the expenditure of public funds for the administration of the no fault act. The Supreme Court of Illinois declared provisions of the statute in violation of the federal and state constitutions. 283 N.E.2d at 474.

63. 283 N.E.2d at 476-78. Because *Grace* was a taxpayer's suit, the plaintiff was permitted extremely liberal pleading. The plaintiff was clothed with standing despite no allegation of direct injury in fact. *See* Ring, *Insight Into a Successful Constitutional Trial*, 8 TRIAL 49, Mar.—Apr. 1972, at 49-50. It is questionable whether other state courts would grant a plaintiff that advantage. *See, e.g.*, Shavers v. Kelley, 402 Mich. 554, 267 N.W.2d 72 (1978), *cert. denied*, 442 U.S. 934 (1979). Standing in the federal courts under like circumstances would also be extremely doubtful. *See, e.g.*, Warth v. Seldin, 422 U.S. 490 (1975).

64. 283 N.E.2d at 478-80.

65. "Motor Vehicles" were defined as:

> A sedan, station wagon or jeep type automobile not used as a livery conveyance for passengers, nor rented to others, and includes any other 4-wheel vehicle used as a utility automobile, pick up truck, sedan delivery truck or panel truck which is not used primarily in the occupation, profession or business of the insured.

1971 Ill. Laws at 2543.

66. *Id.* at 2542.

forego automobile liability and no fault insurance.[67]

Criticism was also directed at the tort liability damages prescribed in the act. Again two classes were created: those whose injuries were serious and who were eligible to sue for pain and suffering without statutory limitation, and those whose injuries were not serious and who were confined to the act's "formula" damages. Within the nonserious injury category, two sub-classifications emerged: those with medical expenses below and those with medical expenses above $500.[68]

Substantial differences in the costs of medical care could predictably result in widely varying general damages for similarly situated victims treated at different times and places.[69] Furthermore, the "formula" for general damages applied to all personal injury tort actions arising out of the use, operation and maintenance of motor vehicles and not just to actions brought by victims entitled to no fault benefits. The anticipated discriminatory impact of the act's classifications, coupled with the act's inseverability section, sealed the fate of the entire no fault article.[70]

2. *Special Legislation Issue*—Despite its discussion of the discriminatory nature of the act's classifications, the *Grace* court did not resolve the equal protection issues. Instead, it determined that the no fault statute ran afoul of the state's constitutional prohibition against "special legislation."[71] The Illinois Constitution states: "The General Assembly shall pass no special or local law when a general law is or can be made applicable. Whether a general law is or can be made applicable shall be a matter for judicial determination."[72] This proscription not only establishes judicial primacy in identifying what laws are "special," but also appears to

---

67. 283 N.E.2d at 478.

68. 1971 Ill. Laws at 2548. In the first group, general damages could not exceed 50% of medical expenses; for the second group general damages encompassed an additional amount equal to medical expenses in excess of $500. *Id.*

69. *See* Ring, *supra* note 63, at 49, 50. Wide variance in hospital room costs, fees for surgical procedures and physicians' charges for office visits was introduced at the trial. *See* 8 TRIAL 10, Jan.-Feb. 1972, 11-13 (text of Memorandum Opinion). Plaintiff's evidence at the trial revealed that in ten predominantly black, low-income wards in Chicago the majority of the population owned no car or, if owning a car, carried no liability—and hence no first party—insurance. *See* Ring, *supra* note 63, at 49-50. If injured by an uninsured vehicle, these individuals could obviously not look to their own insurance for no fault benefits.

70. The discretionary impact of the act's classifications persuaded the trial court that those classifications violated the equal protection clauses of the federal and state constitutions. 8 TRIAL 10, Jan.—Feb. 1972, 11-13 (text of Memorandum Opinion).

71. 283 N.E.2d at 479.

72. ILL. CONST. art. IV, § 13.

function as an additional weight in the equal protection scale. Under minimum federal scrutiny, state legislative reform may proceed "one step at a time" and is not necessarily open to an equal protection challenge simply because the statutory classification is underinclusive.[73] This does not appear to be the approach of the Illinois Supreme Court in *Grace*, however. The state constitutional prohibition on special laws was construed to impose a more rigorous standard of review for legislative experimentation.[74]

3. Right to Jury Trial—The Illinois Supreme Court also held that the act's provision for compulsory arbitration of small claims violated the state's constitution.[75] Specifically, the court ruled that assigning to the circuit courts the *review* of arbitration awards entered as judgments deprived the circuit court of *original* jurisdiction.[76] Further, the appeals process violated the state constitutional objective of abolishing trials de novo, and mandatory arbitration deprived the parties of the right to trial by jury.[77]

4. Conclusion—The Illinois experience may alert practitioners seeking to challenge or defend state no fault legislation to explore several avenues:

(1) What are the statutory parameters for taxpayer's suits?
(2) Does the state constitution include a rigorous proscription against special legislation?
(3) What standards have the state courts applied to constitutional guarantees of equal protection and right to jury trial?

Barring a radical change in the Illinois Supreme Court's philosophy, a new attempt at comprehensive no fault legislation in Illinois would not likely survive judicial review. The special legislation prohibition superimposed on equal protection analysis poses a formidable barrier to the constitutionality of no fault legislation.

---

73. *See* Williamson v. Lee Optical, 348 U.S. 483, 489 (1955).

74. 283 N.E.2d at 479. The Illinois court has interpreted the prohibition on special legislation in other contexts. *See, e.g.*, Friedman & Rochester, Ltd. v. Walsh, 67 Ill. 2d 413, 367 N.E.2d 1325, 1329 (1977) (upholding an exemption from garnishment annuities and pensions paid by the Firemen's Annuity and Benefit Fund); Wright v. Central DuPage Hosp. Ass'n, 63 Ill. 2d 313, 347 N.E.2d 736, 742-43 (1976)(declaring unconstitutional a law that provided limiting recovery for injuries caused by medical or hospital malpractice).

75. 283 N.E.2d at 474.

76. 283 N.E.2d at 480-81.

77. *Id.* The court distinguished a Pennsylvania case which sustained compulsory small claims arbitration on the ground of differences in the applicable Pennsylvania law. *Id.* at 481 (citing Application of Smith, 381 Pa. 223, 112 A.2d 625, *appeal dismissed sub. nom.* Smith v. Wissler, 350 U.S. 858 (1955)).

V. New Hampshire

Because the no fault plan proposed in New Hamphire[78] did not achieve enactment, this section will not examine in detail a state supreme court advisory opinion on the proposed plan. A brief analysis of the opinion is useful, however, to illustrate the attitude of that state's highest court to modifications of traditional tort rights of action. In *Opinion of the Justices*,[79] the Supreme Court of New Hampshire responded to a request by the state's House of Representatives to render an advisory opinion on the proposed New Hampshire no fault plan. The plan provided that accident victims must meet a monetary threshold of $1000 in medical expenses before bringing an action at law to recover damages for economic loss beyond basic no fault reparations and for pain and suffering.[80] Upholding this proposed limitation on tort recovery, the court commented, "[t]he test is whether a $1000 threshold would have a fair and substantial relation to the object of legislation. . . . the judgment of the legislature must be accepted unless it is very wide of any reasonable line of demarcation."[81]

The court took a different view, however, on the issue of the legislature's power to authorize arbitration of claims under $3000 and to require prepayment of arbitration costs as a prerequisite to appeal. The court held that, if enacted, those provisions would violate the state constitution because they infringed on the rights of jury trial and of appeal.[82] The dissenting opinion in the 3-2 decision found the tort threshold as well as the arbitration provisions unconstitutional.[83] The slim majority favoring the tort action limitation does not permit a firm prediction of the New Hampshire Supreme Court's attitude toward any future no fault plan.

VI. Massachusetts

A. *Introduction*

Massachusetts, the first state to enact compulsory motor vehicle liability insurance,[84] was also the first to adopt a no fault stat-

78. H.R. 79, 1973 Leg., 1st Sess. (1973).
79. 113 N.H. 205, 304 A.2d 881 (1973).
80. 304 A.2d at 885.
81. *Id.* at 887.
82. *Id.*
83. *Id.* at 888 (Duncan, J., Grimes, J., dissenting).
84. Act of May 24, 1926, ch. 368, 1926 Mass. Acts 423.

ute, which became effective January 1, 1971.[85] The original statute, confined to personal injury protection, was amended to add compulsory liability and no fault property damage protection, effective January 1, 1972.[86] With the subsequent repeal of the no fault property damage coverage, the Massachusetts no fault plan returned to its original scope—minimal personal injury protection.[87]

The Massachusetts no fault plan provided for $2000 of first party benefits, including necessary medical expenses, funeral expenses, substitute services and seventy-five percent of actual wage loss.[88] Every owner or operator of a motor vehicle to which personal injury benefits apply is exempt from tort liability for bodily injury or death to the extent that the victim can recover economic loss under the no fault scheme.[89] Victims seeking damages for pain and suffering must meet either a monetary or injury threshold to maintain tort actions. Either the reasonable and necessary health care expenses of those victims must exceed $500 each,[90] or they must have sustained an injury that causes death, loss of a body member, permanent disfigurement, loss of sight or hearing, or a fracture.[91] A victim may also sue in tort for expenses above $2000, and for the difference between reduced earning capacity and actual lost wages recoverable under no fault.[92] This modest no fault scheme continues in operation, having survived an early and multifaceted attack in the case of *Pinnick v. Cleary*.[93]

B. Pinnick v. Cleary

The *Pinnick* cause of action arose only two days after the effective date of the Massachusetts no fault statute.[94] Two cars in-

---

85. 1970 Mass. Acts (currently codified at MASS. ANN. LAWS ch. 90, §§ 34A, 34D, 34M, 34N; ch. 175, §§ 22E-22H, 113B, 113C; ch. 231, § 6D (Michie/Law Co-op. 1975)).

86. MASS. ANN. LAWS ch. 978, § 34(O) (Michie/Law Co-op. 1975)).

87. The compulsory $5000 property damage liability coverage was continued. *See id.* ch. 90 § 34(O) (Michie/Law Co-op. Supp. 1982).

88. *Id.* ch. 90, § 34A. Even these minimal benefits may be reduced or eliminated for the insured and household members by election of deductibles up to $2000. *Id.* ch. 90, § 34M (Michie/Law Co-op. 1975).

89. *Id.* § 34M.

90. *Id.* ch. 231, § 6D.

91. *Id.*

92. *See id.* ch. 90, § 34M.

93. 360 Mass. 1, 271 N.E.2d 592 (1971).

94. *See* 271 N.E.2d at 595. The heated battle that developed in Massachusetts over the issue of no fault insurance can be discerned both from the impressive representation of amici curiae in *Pinnick*, as well as from the literature of the day. *See id.* at 596; Coombs, *The Massachusetts Experience Under No-Fault*, 601 INS. L.J. 69 (1973); Ryan, *Massachu-*

volved in an accident were insured under policies that included personal injury protection coverage entitling both vehicle owners to no fault benefits and tort exemption up to $2000. The plaintiff sustained injuries in the amount of $115 in reasonable and necessary medical expenses.[95] The court calculated his total potential recovery, absent no fault, as $1,565.[96] In response to the plaintiff's demand for these common law damages, the defendant raised as defenses the act's tort liability exemption and its bar to general damages.[97] The plaintiff challenged the Massachusetts no fault plan on a variety of constitutional grounds, asserting that a common law tort action represented a vested property right shielded by the federal and state constitutions from interference by state legislation, and that the plan violated the due process and equal protection guarantees of the United States Constitution.[98]

1. Vested Property Right—The Supreme Judicial Court of Massachusetts found the plaintiff's characterization of the common law personal injury action as a vested and immutable property right unpersuasive.[99] Rejecting the plaintiff's contention that the Massachusetts Constitution required the preservation of common law remedies, the court stated:

> [P]laintiff seems to ignore the distinction between a cause of action which has accrued and the expectation which every citizen has if a legal wrong should occur to find redress according to the rules of statutory and common law applicable at that time. The Legislature is admittedly restricted in the extent to which it can retroactively affect common law rights of redress which have already accrued.

---

setts Tries No-Fault, 57 A.B.A.J. 431 (1971); Note, *The Massachusetts "No-Fault" Automobile Insurance Law: An Analysis and Proposed Revision*, 8 HARV. J. ON LEGIS. 455 (1971).

95. 271 N.E.2d at 596-97.

96. *Id.* at 596. That figure was composed entirely of medical expenses, wage loss and pain and suffering. *Id.*

97. *Id.* at 597. *See* MASS. ANN. LAWS ch. 90 § 34M (Michie/Law Co-op. Supp. 1982).

98. 271 N.E.2d at 594-96.

99. Specifically, the plaintiff argued that the Massachusetts Constitution, MASS. CONST. pt. II, ch. 6, art. 6, forbade the destruction of common law remedies:

> All the laws which have heretofore been adopted, used and approved in the Province, Colony or State of Massachusetts Bay, and usually practised on in the courts of law, shall still remain and be in full force, until altered or repealed by the legislature; such parts only excepted as are repugnant to the rights and liberties contained in this constitution.

Id. He further argued that the Massachusetts Declaration of Rights, MASS. CONST. pt. I, art. 11, § 12, guaranteed a remedy for any injury. The Declaration provides: "Every subject of the commonwealth ought to find a certain remedy, by having recourse to the laws, for all injuries and wrongs which he may receive in his person, property, or character." *Id.*

> However, there is authority in abundance for the proposition that [n]o person has a vested interest in any rule of law entitling him to insist that it shall remain unchanged for his benefit.[100]

The court also decided that the guarantee of a "remedy for all wrongs" expressed in the State Declaration of Rights[101] was "clearly directed toward the preservation of procedural rights."[102] Thus, the state court's definition, rather than the constitutional language securing redress for all wrongs, determined the remedies available to an injured party.

2. *Due Process and Equal Protection*—The plaintiff also attempted to elevate the personal injury tort action to a fundamental right protected by the United States Constitution. The court found that the fundamental right of personal security and bodily integrity claimed by the plaintiff was unaffected by the act.[103]

Having dismissed the plaintiff's attempt to identify a fundamental right calling for strict scrutiny review, the court instead employed a rationality test to judge the due process and equal protection challenges to the plan. The court decided that the statute was rationally related to a legitimate legislative objective.[104] Thus, the minimum threshold required to uphold the statute's constitutionality was present. The court fully described the evils and burdens of excessive motor vehicle accident litigation in Massachusetts including: court calendar congestion, delay, high cost of automobile insurance and inequities in reparations. It found the no fault statute "a rational approach to the solution of these patent inefficiencies and inequities."[105]

The general theory of plaintiff's challenge was that the legislature had violated due process by abolishing the right of tort recovery, thus leaving the victim without legal redress.[106] The Supreme

---

100. 271 N.E.2d at 599. (citation omitted) (quoting New York Central R.R. v. White, 243 U.S. 188, 198 (1917)).

101. *See* MASS. CONST. pt. I, art. 11, § 12; *supra* note 99.

102. 271 N.E.2d at 600. *See* Commonwealth v. Hanley, 337 Mass. 384, 387, 149 N.E.2d 608, 610 (1958). Massachusetts' Declaration of Rights is similar to the Florida access to court guarantee, but Florida's highest court has attached a substantive due process gloss to that provision, rendering it more useful in constitutional litigation of this type. *See infra* notes 165-176 and accompanying text.

103. 271 N.E.2d at 600. The court found the purported analogy to Griswold v. Connecticut, 381 U.S. 479 (1965), implausible.

104. 271 N.E.2d at 602. *See supra* notes 42-43 and accompanying text.

105. 271 N.E.2d at 605.

106. *Id.* at 601-02. Two other specific due process complaints were also raised. First, that requiring individuals to obtain insurance through private, profit making corporations violated due process. The court responded that first party insurance is optional because a

Judicial Court of Massachusetts rejected this contention and refused to concede any constitutional imperative that the legislature must offer a quid pro quo when it altered a common law tort remedy. Nonetheless, the court recalled the United States Supreme Court's statement in *New York Central Railroad Company v. White*[107] that a legislature could not eliminate rights of action without substituting "something adequate in their stead."[108] The court, however, found an acceptable "exchange of rights."[109] In return for restrictions imposed on noneconomic recoveries, the plaintiff enjoyed prompt payment of out-of-pocket loss without litigation expense. Even the negligent defendant could be compensated for injuries and be partially protected from tort liability. These substituted rights appeared at least as adequate as those approved by the Supreme Court in upholding the workers' compensation law in *White*.[110]

The plaintiff also charged that certain threshold criteria for instituting a tort action for pain and suffering were unreasonable and arbitrary, and thus violated his equal protection rights. Specifically, the plaintiff attacked the $500 medical expenses "floor," and the inclusion of fractures as a per se qualifying injury. Although neither of these requirements guaranteed against misuse of the right to sue, the court concluded that both were within the range of legislative judgment.[111] The more pointed objection—that a floor invidiously discriminated against the poor because medical services cost less to them—was dismissed for lack of evidence, and not addressed on the merits.[112]

---

deductible up to $2000 may be elected, and even if insurance were compulsory, it was settled that the legislature may require one to take measures for his own benefit if the public good was thereby served. *Id.* at 607. Second, that due process was violated because each vehicle owner in a class and territory paid the same premium for "compulsory" insurance regardless of his other sources of recovery; thus some would benefit more than others from a no fault system. The court rejected that contention, observing the rates fixed by the Commissioner of Insurance were reviewable upon request by an aggrieved party, and also that any person could tailor his coverage to his collateral sources of recovery. It was not administratively feasible to engage in individualized, precise ratemaking. *Id.* at 608. The court did not entertain two additional due process objections because they were not raised by the facts: (1) a policyholder can elect a deductible which applies to members of his household; (2) the pure pedestrian (who is not an owner or member of an owner's household) is subject to personal injury protection ("PIP") benefits and disabilities. *Id.* at 609.

107. 243 U.S. 188 (1917). *See supra* notes 18-26 and accompanying text.

108. 271 N.E.2d at 605 (quoting *White*. 243 U.S. at 201).

109. *Id.* at 606-07.

110. *See id.*; *supra* notes 21-22 and accompanying text.

111. 271 N.E.2d at 609-11.

112. *Id.* at 611.

In assessing the plaintiff's equal protection argument, the court ruled that legislative classifications were to be judged by principles analogous to those applicable to due process issues. The court employed the most lenient, permissive equal protection standard,[113] holding that the classifications would be sustained if they could be supported by any reasonably conceivable facts; no legislative findings were required as a predicate.[114] Under this standard, only arbitrary and irrational classifications would be beyond the broad scope of permissible legislative discretion.[115]

Accordingly, the court found that a proper legislative objective—eliminating minor (often exaggerated) claims—justified restrictions upon tort actions for pain and suffering damages.[116] The means chosen to accomplish the objective were rationally related to the legislative goal, even though some inequalities might result from their application.[117] The reasoning of the court in determining both the due process and equal protection challenges expressed extreme deference for legislative judgment in devising the no fault regulatory system.

C. *Chipman v. Massachusetts Bay Transportation Authority*

In several cases subsequent to *Pinnick*, the Supreme Judicial Court of Massachusetts considered tort recovery issues in the context of particular factual situations. The first case, *Chipman v. Massachusetts Bay Transportation Authority*,[118] involved a passenger injured while boarding the defendant's bus. The plaintiff sustained medical expenses below the $500 threshold and did not suffer any of the injuries that qualified as a serious injury.[119] The

---

113. *See supra* notes 42-43 and accompanying text.

114. 271 N.E.2d at 609. *See* Marshal House, Inc. v. Rent Control Bd., 358 Mass. 686, 266 N.E.2d 876, 882-83 (1971). Among the authorities cited by the court for this standard were: McDonald v. Board of Election Comm'rs, 394 U.S. 802, 809 (1969); McGowan v. Maryland, 366 U.S. 420, 425-26, (1961); Hall-Omar Baking Co. v. Comm'r of Labor and Industry, 344 Mass. 695, 700, 184 N.E.2d 344, 351 (1962); Maher v. Brookline, 339 Mass. 209, 213, 158 N.E.2d 320, 322 (1959).

115. *See* 271 N.E.2d at 609 (citing Opinion of the Justices, 251 Mass. 569, 601, 147 N.E. 681, 695 (1925)).

116. 271 N.E.2d at 610-11.

117. *Id.* In upholding the $500 threshold for medical expenses, the court refered to other cases in which numbers were employed in demarcating classes. It cited the Unemployment Compensation Act exemption for employers of less than 8 persons. *See id.* at 610-11 (also citing as an example Carmichael v. Southern Coal & Coke Co., 301 U.S. 495, 510-12 (1937)).

118. 366 Mass. 253, 316 N.E.2d 725 (1974).

119. 316 N.E.2d at 726.

plaintiff had no personal injury protection insurance because neither she nor any member of her household owned a motor vehicle. The issue before the court was whether the plaintiff was barred by the no fault statute from suing the defendant for pain and suffering damages.[120]

The *Chipman* court held that "a plaintiff who has no recourse to personal injury protection benefits is not barred . . . from recovering damages for pain and suffering when the uninsured defendant is expressly exempted from the no-fault scheme."[121] To rule otherwise would have resulted in plaintiff's total loss of the right of recovery. The court implied that the *Pinnick* rationale required a quid pro quo before the plaintiff could be deprived of the right to sue for out of pocket expenses.[122]

The court appears to have reasoned that a no fault exemption for a tortfeasor must be matched by first party benefits available to a plaintiff.[123] Permitting the plaintiff to recover pain and suffering when neither party was a "covered person," did not violate the no fault act's principal purpose of reducing liability insurance rates for Massachusetts motor vehicle owners. Because the defendant was self-insured, its loss experience would have no effect on liability insurance rates.[124] The *Chipman* court reached its conclusion purely on the basis of statutory construction. It cautioned, however, that any other result would raise doubts about the constitutionality of denying damages to a "pure pedestrian" who did not receive the benefits of no fault personal injury protection in return.[125]

D. Cyr v. Farias

The next challenge to Massachusetts' no fault plan came in *Cyr v. Farias*.[126] In *Cyr*, nonresident plaintiffs sued for injuries that resulted from an accident in Massachusetts in which the defendant, a Massachusetts resident, was a no fault insured party. The medical expenses of each of the plaintiffs did not reach the

120. *Id.* at 727-28. The defendant claimed the benefit of the act's exemption from liability for those damages even though it was exempted from the requirements of the no fault statute. *Id.*

121. *Id.* at 729.

122. *Id.* at 728-29. *Cf. supra* note 108 and accompanying text.

123. *See id.* at 728.

124. *Id.*

125. *Id.* at 728-29. The *Chapman* court implied that denying damages to a "pure pedestrian" under these circumstances would raise the issue of deprivation of due process. *Id.*

126. 367 Mass. 720, 327 N.E.2d 890 (1975).

$500 threshold. Both plaintiffs claimed pain and suffering damages.[127] The issues in *Cyr* were whether the threshold requirement for pain and suffering damages[128] applied to nonresidents not covered by no fault benefits, and whether denial of damages for pain and suffering to nonresidents violated the due process and equal protection clauses of the federal and state constitutions.

The *Cyr* court held that the statute could exclude nonresident plaintiffs from the right to seek pain and suffering damages if their medical expenses were less than $500.[129] Like *Chipman*, this decision too was based strictly upon statutory construction. The court stated that there was no invidious discrimination against nonresidents; the threshold limitations operated with equal force on resident plaintiffs. Under the Massachusetts statute, nonresidents could sue for the same amounts available to those covered by no fault.[130] The *Cyr* court concluded, "[t]hat Massachusetts does not choose to provide, free of cost, no-fault payments to injured nonresidents is not a reason for concluding that it must allow nonresidents to recover for pain and suffering in cases where its own citizens could recover only their actual losses."[131]

E. Conclusion

Massachusetts has maintained the monetary and qualifying injuries thresholds challenged in *Pinnick* in 1971. Other states, such as New York, Florida and Michigan have abandoned or substantially amended original criteria for tort actions.[132] There are several possible explanations for the continued vitality of Massachusetts' threshold test.

From an insurer's point of view it might be advantageous to raise substantially the monetary threshold. But given the Massachusetts court's extremely deferential attitude to legislative classifications, it would be difficult to mount a convincing argument against the $500 line drawn in the statute.[133] Statistical evidence

---

127. 327 N.E.2d at 891.

128. MASS. ANN. LAWS ch. 231, § 6D (Michie/Law. Co-op. 1975).

129. 327 N.E.2d at 893-94.

130. *Id.* at 893. In addition, lost residents could sue for 100% of lost wages; residents were restricted to a 75% recovery. *Id.* Nonresidents could thus recover their out-of-pocket losses.

131. *Id.* at 894.

132. *See supra infra* notes 147-52, 464-72, 522-23 and accompanying text.

133. *See* 271 N.E.2d at 610-11. Commenting on the $500 threshold requirement for medical expenses, the court quoted Justice Holmes:

When a legal distinction is determined, as no one doubts that it may be, between

might be available, however, to show that the legislative objective of eliminating minor and exaggerated claims was demonstrably disserved by so low a threshold.

From the perspective of plaintiffs and their attorneys, an attempt to eliminate the $500 gateway might not be worth the effort. Since the *Pinnick* decision in 1971, health care costs have mounted in the general inflationary spiral and it is not difficult today to catalog medical expenses in excess of the $500 minimum. Other states have found the threshold to be an almost meaningless barrier; thus there is little incentive for plaintiffs to challenge the Massachusetts provision.

The due process rights of the pure pedestrian, when confronting a defendant who is not a no fault insured, have been cautiously and limitedly exposed in the *Chipman* case. Some innovative extrapolations from this decision might produce constitutional issues, but the case as it stands is confined to its facts. It could be precedent for other lawsuits involving governmental bodies or persons exempted from no fault insurance requirements.

In *Pinnick*, the Supreme Judicial Court of Massachusetts also found no merit in the plaintiff's argument that no fault ratemaking procedures worked a denial of due process. That issue, however, was the most significant question in the Michigan case of *Shavers v. Kelley*,[134] and might be worthy of further examination in Massachusetts. The likelihood of a successful due process challenge to Massachusetts' no fault ratemaking depends on the philosophy of the Massachusetts court,[135] the presence or absence of adequate statutory protection for the insured and the energy of the bar in exploring a complex regulatory mechanism.

---

night and day, childhood and maturity, or any other extremes, a point has to be fixed . . . [W]hen it is seen that a line or point there must be, and that there is no mathematical or logical way of fixing it precisely, the decision of the Legislature must be accepted unless we can say that it is very wide of any reasonable mark.

Id. (quoting Louisville Gas & Elec. Co. v. Coleman, 277 U.S. 32, 41 (1928)).

134. 402 Mich. 554, 267 N.W.2d 72 (1978). *See infra* notes 591-657 and accompanying text.

135. For example, if the Massachusetts Supreme Judicial Court adopted the viewpoint of the Michigan Supreme Court, the Massachusetts no fault rate schedule might be vulnerable to a due process attack.

VII. FLORIDA

A. *Introduction*

Because the Florida no fault statute has been substantially modified since its enactment in 1971,[136] it is helpful to summarize both its current and original provisions before analyzing challenges to the statute. Claims that are presently outstanding may be based on a former or the current version of the statute. The evolution of the present Florida act also presents an illustrative example of a successful attack on no fault jurisprudence.

The original 1971 act stated that every vehicle required to be registered and licensed in Florida must maintain liability security by insurance or other approved method,[137] as well as no fault coverage that complied with the Florida Automobile Reparations Reform Act.[138] Noncompliance exposed the motor vehicle owner to personal liability for payment of no fault benefits.[139]

The benefits provided by the act included all reasonable medical expenses,[140] disability benefits for loss of income and earning capacity,[141] funeral expenses up to $1000, and all expenses reasonably incurred in obtaining substitute services for those that the injured person would have performed for his household.[142] The aggregate of benefits for each individual was $5000.[143] Persons

---

136. Florida Automobile Reparation Reform Act, Pub. L. No. 71-252, 1971 Fla. Laws 1355, 1355-71, *repealed by* Regulatory Sunset Act, Pub. L. No. 81-318, 1981 Fla. Laws 1493 (amending FLA. STAT. ANN. §§ 11.61-.6115 (West Supp. 1982)). The Florida legislature substantially reenacted the no fault act in 1982. Act of June 21, 1982, Pub. L. No. 82-243, 1982 Fla. Laws 1553-70 (1982). The new act provides for a maximum deductible of $2000, eliminating the previous options which permitted a maximum deductible of $8000. 1982 Fla. Laws at 1568.

137. 1971 Fla Laws at 1356.

138. *Id.*

139. *Id.* at 1357.

140. *Id.* at 1359. Those expenses included medical, surgical, dental, rehabilitative, hospital and nursing services. The act specified that charges for medical treatment must be "reasonable" and not in excess of customary charges in cases not involving insurance. *Id.* at 1359.

141. The act provided for 100% recovery if the benefits were includable as gross income for federal income tax purposes, and 85% recovery if they were not. *Id.*

142. *Id.*

143. *Id.* Effective January 1, 1979, the aggregate limit of personal injury benefits was increased to $10,000. FLA. STAT. ANN. § 627.736(1) (West Supp. 1982). That increase covered 80% (instead of the previous 100%) of reasonable expenses for necessary medical services; 80% (instead of 100%) of loss of gross income, and 60% (instead of 85%) of income not includable in gross income for federal income tax purposes; the replacement services expenses and the original $1000 maximum for funeral expenses were retained. *Id.* §

qualifying for benefits were the named insured, relatives residing in the same household, operators and other occupants of the insured vehicle,[144] and persons struck by the insured vehicle who were not occupants of a motor vehicle or motorcycle.[145] The named insured and relatives could also recover no fault benefits for accidental bodily injury sustained in Florida if they were occupants of any motor vehicle, or if they were nonoccupants whose injuries were caused by physical contact with a motor vehicle. They could also recover for injury sustained elsewhere in the United States or Canada if they were occupants of the owner's vehicle.[146]

The 1971 statute granted tort exemptions, to the extent of the no fault benefits described above, to any properly insured owner or operator.[147] An injured person, however, could bring a tort action against an owner or operator for pain, suffering, mental anguish and inconvenience if medical expenses exceeded $1000, or the injury involved permanent disfigurement, fracture of a weight bearing bone, loss of body function or member, or permanent injury or death.[148] The insurer had a right of reimbursement out of any tort recovery that duplicated no fault benefits.[149]

The 1971 act included *optional* full or basic coverages for accidental property damage to the owner's motor vehicle.[150] It also granted a tort exemption up to $550 to owners and operators of properly insured motor vehicles for accidental property damage to

---

627.736(1)(a)(b).

144. Motor vehicles, defined generally in the original act, were defined in greater detail in 1978 to include any four-wheel vehicle not used in a business. Act of June 20, 1978, Pub. L. No. 78-374, 1978 Fla. Laws 1042 (codified at FLA. STAT. ANN. § 627.732(1) (West Supp. 1982)).

145. 1971 Fla. Laws at 1359.

146. *Id.* at 1362.

147. *Id.* at 1365-66.

148. *Id.* at 1366. This area of the Florida no fault plan was amended in 1976, 1978, and 1979. Tort exemption to the extent of mandated personal injury limits ($10,000) was maintained for owners and operators complying with the law's security requirements; but the amendments altered the definition of injuries that would sustain a right of action for noneconomic loss, and abolished the $1000 threshold alternative effective October 1, 1976. Act of June 20, 1978, Pub. L. No. 78-374, 1978 Fla. Laws 1043; see Act of June 27, 1976, Pub. L. No. 76-266, 1976 Fla. Laws 722. Effective January 1, 1979, pain, mental anguish and inconvenience damages are recoverable only if the injury consists of (1) permanent loss of important body function, or (2) permanent injury, or (3) permanent disfigurement, or (4) death. 1978 Fla. Laws at 1043-44.

149. 1971 Fla. Laws at 1366.

150. *Id.* at 1366. Full coverage provided insurance without regard to fault for accidents occurring in the United States or Canada. Basic coverage was limited to damage caused by the fault of another, and contact between the owner's and another vehicle was required. *Id.* at 1367.

other (except parked) motor vehicles that were required to be insured by the act.[151] Owners who did not elect to purchase property damage insurance for their vehicles could bring a tort action against the persons responsible for the vehicle that caused damage, but only if the damage exceeded $550.[152]

In addition to prescribing a basic reparations scale, the act addressed the important subject of *offsets*. Originally, only workers' compensation benefits were required to be credited against benefits due from an insurer. In 1977, however, the legislature required that Medicaid payments be included in those offsets.[153] The legislature also directed courts to admit into evidence in personal injury or wrongful death automobile accident cases, the total amount of collateral sources of recovery paid to the claimant prior to trial.[154] A 1978 amendment to the Florida act considerably strengthened and amplified this collateral source provision by requiring courts to instruct juries to deduct from their verdicts the value of all benefits received from any collateral source.[155] This sweeping abrogation of the collateral source rule went far beyond other state statutes, and applied to all tort actions arising out of the operation of a motor vehicle.

B. *Kluger v. White*

In *Kluger v. White*,[156] the Florida Supreme Court declared unconstitutional the property damage provision of Florida's original no fault law.[157] The property damage section provided that an

---

151. *Id.*

152. *Id.* It was this no-recovery zone that influenced the court to declare the property provisions unconstitutional in Kluger v. White, 381 So. 2d 1 (Fla. 1973). *See infra* notes 156-76 and accompanying text.

153. Act of June 23, 1977, Pub. L. No. 77-468, 1977 Fla. Laws 2078 (codified at FLA. STAT. ANN. § 627.736(4) (West Supp. 1982)). For the definition of medicaid payments, see 42 U.S.C. § 1396ff (Supp. IV 1980).

154. FLA. STAT. ANN. § 627.737(2) (West Supp. 1982).

155. Collateral sources were defined as payments made to the claimant or on his behalf pursuant to the Social Security Act; any federal, state or local income disability act or other public program providing medical expenses or disability benefits, any health, sickness or income disability insurance or similar insurance (other than life insurance) whether purchased by the claimant or others; any agreement of any group, corporation or organization providing payments for medical care, a contractual or voluntary wage continuation plan provided by employers. Act of June 20, 1978, Pub. L. No. 78-374, 1978 Fla. Laws 1043 (codified at FLA. STAT. ANN. § 627.7372 (West Supp. 1982)).

156. 281 So. 2d 1 (Fla. 1973).

157. The property damage provision was repealed in 1976 because of the *Kluger* opinion. Act of June 27, 1976, Pub. L. No. 76-266, 1976 Fla. Laws 726. For an approving discussion of *Kluger*, see Comment, 2 FLA. ST. U. L. REV. 178 (1974).

owner of a motor vehicle could elect to purchase full or basic protection coverage for accidental damage to his vehicle. Full coverage afforded protection without regard to fault; basic protection coverage reimbursed damages caused by another's fault. Owners of vehicles maintaining the requisite no fault security were exempt from tort liability for accidental property damage to a motor vehicle in Florida.[158] Notwithstanding this exemption, an owner-plaintiff who rejected the property damage insurance option could sue the at-fault owner or operator of another vehicle if the damage exceeded $550. Furthermore, the insurer of a car with property damage coverage could recover from the insurer of the at-fault owner or operator any benefits that it had paid its insured if the damage exceeded $550.[159] Thus, the common law action for property damage arising from automobile accidents was abolished unless the plaintiff had not purchased property protection insurance and the damages exceeded $550.

In *Kluger*, the plaintiff sustained $250 property damage to her car as a result of the defendant's allegedly negligent driving. The plaintiff's insurer refused to pay because her insurance did not include coverage for property damage[160] or collision damage[161] to her vehicle. The defendant's insurer refused to pay for the loss on the ground that the statute conferred tort immunity for vehicle damages that did not exceed $550.

Following the circuit court's dismissal of the action,[162] the plaintiff appealed, contending that the statute's abolition of her right to sue to recover for property damage was unconstitutional as a denial of equal protection and of jury trial, and as a deprivation of property without due process. The equal protection claim was based on the act's differential treatment of owners with property damage insurance, as opposed to uninsured owners, and of owners with damages exceeding $550 as compared with those suffering damages of a lesser amount.[163] Due process was implicated because the act afforded the plaintiff no opportunity to be heard.[164] The court did not reach those contentions. It concentrated on the

158. FLA. STAT. ANN. § 627.738 (West Supp. 1982).

159. *Id.*

160. The coverage had been offered the plaintiff in compliance with the statute; plaintiff had executed a rejection form. 281 So. 2d at 5 (Boyd, J., dissenting).

161. *Id.* at 2.

162. 37 Fla. Supp. 183 (Dade Co. Cir. Ct. 1972).

163. 281 So. 2d at 3.

164. See Comment, *supra* note 157 at 180 n.10.

plaintiff's chief complaint—denial of access to the courts.

1. Access to the Courts—Article I, section 21 of the Florida Constitution reads: "The courts shall be open to every person for redress of any injury, and justice shall be administered without sale, denial or delay."[165] In light of that provision, the *Kluger* court viewed the central issue to be whether the legislature had authority to abolish a common law or statutory right without providing an adequate alternative.[166] A closely divided court ruled that under the state constitution the legislature lacked authority to bar property damage claimants like the plaintiff from access to the courts.[167]

The Florida Supreme Court held that the legislature could not abolish an established right of action without providing a reasonable alternative:

> Where a right of access to the courts for redress for a particular injury has been provided by statutory law predating the adoption of the Declaration of Rights of the Constitution of the State of Florida, or where such right has become a part of the common law of the State pursuant to Fla. Stat. § 2.01, F.S.A., the Legislature is without power to abolish such a right without providing a reasonable alternative to protect the rights of the people of the State to redress for injuries, unless the Legislature can show an overpowering public necessity for the abolishment of such right, and no alternative method of meeting such public necessity can be shown.[168]

Further, the majority advised:

> Had the Legislature chosen to require that appellant be insured against property damage loss—as is, in effect, required by Fla. Stat. § 627.733, F.S.A., with respect to other possible damages—the issues would be different. A reasonable alternative to an action in tort would have been provided and the issue would have been whether or not the requirement of insurance for all motorists was reasonable. That issue is not before us.[169]

Several provocative extrapolations are suggested by the *Kluger* holding and dictum. First, the right to "open courts," as the Florida Supreme Court construes article I, section 21, is not

---

165. FLA. CONST. art. I, § 21.

166. 281 So. 2d at 3.

167. *Id.* at 4, 5. The *Kluger* court split four justices to three.

168. 281 So. 2d at 4. Of course, a cause of action for vehicle property damage was recognized in Florida before adoption of the 1968 Constitution. *Id.*

169. *Id.* at 5.

necessarily congruent with procedural due process.[170] Although section 21 may include protection against certain procedural barriers, the court does not distinguish between that section's procedural and substantive aspects. Section 21 apparently protects more than mechanical procedures, and preserves substantial remedial rights.[171] Therefore, it would appear that every traditional right of action remains inviolate in Florida unless one of two tests is met.

The first test prescribed by the *Kluger* court is that the legislature may abolish an established "right of access to the courts" if it substitutes a "reasonable alternative."[172] Whether the substitute is "reasonable" and whether it fulfills the role of an "alternative" must be determined by the courts. They would at the very least be obligated to balance detriments against gains in the exchange of rights.[173] At the extreme, this approach might lead to a revival of the substantive due process review of the early twentieth century in which some judges imposed their philosophic, economic and social beliefs as criteria for legislative validity.

The second test established by the *Kluger* court is that if the legislature abolishes a "right of access to the courts" without supplying a "reasonable alternative", it will be acting beyond its constitutional power unless the abolition is justified by "an overpowering public necessity", and there is no alternative way to meet that necessity.[174] This standard of judicial review rises to the level of strict scrutiny adopted by the United States Supreme Court in cases that involve fundamental rights or suspect classifications.[175] Such a high standard of review imposes a severe burden on the legislature whenever it attempts to abrogate existing remedies.

---

170. *See* Boddie v. Connecticut, 401 U.S. 371 (1971). In *Boddie*, indigent persons seeking to file suits for divorce in state courts claimed inability to pay the cost of service of process. which the state would not waive. The Supreme Court held that because only the state could legally dissolve marriages, plaintiffs' due process rights of access to courts were violated by the state imposed financial barrier. *Id.* at 382-83.

171. The predecessor of article I, section 21, as noted in the dissenting opinion in *Kluger*, arguably not only guarantees a remedy for a wide range of "wrongs," but also recognizes the substantive rights that those remedies protect: "All courts in this State shall be open, so that every person for any injury done him in his lands, goods, person or reputation shall have remedy, by due course of law, and right and justice shall be administired [sic] without sale, denial or delay." 281 So. 2d at 6, n.3 (Boyd, J., dissenting). *Compare* the Florida Supreme Court's holding of "access to the courts" with the Massachusetts Supreme Judicial Court's treatment of the "remedy for all wrongs" provision of that state's constitution, discussed *supra*, notes 100-102 and accompanying text.

172. *See* 281 So. 2d at 4.

173. *See, e.g.*, N.Y. Central R.R. v. White, 243 U.S. 188, 201 (1917).

174. *See* 281 So. 2d at 4.

175. *See supra* notes 44-46 and accompanying text.

It seems likely that the Florida Supreme Court did not intend such a drastic result. It illustrated this segment of its holding by referring to the abolition of the cause of action for alienation of affections and criminal conversation.[176] The court confirmed the legislature's powers to bar those claims without supplying a "reasonable alternative" because they had become tools for extortion and blackmail. Arguably, however, the legislature could have engrafted restrictions upon those types of action without totally barring them. How literally the Florida Supreme Court applies the broad holding of *Kluger* in the future will determine the success of any constitutional challenge of legislative authority to abrogate an established right of action.

C. *Lasky v. State Farm Insurance Co.*

In *Lasky v. State Farm Insurance Co.*,[177] the Florida Supreme Court followed its *Kluger* holding that denial of a right of action for property damage of $550 or less violated a constitutional right of access to the courts. The importance of *Lasky* lies not so much in its reiteration of the invalidity of the act's property damage provisions, as in its analysis of due process and equal protection challenges.

The plaintiff in *Lasky* was operating her husband's car when it was struck by another vehicle, resulting in personal injury to the plaintiff and total loss of her car. Its replacement value did not exceed $550, and, therefore, under the no fault act's property damage provisions, no right of action existed to recover the property loss.[178] The plaintiff's personal injuries did not match any of the injuries that would have qualified her to maintain a tort action for noneconomic loss.[179] Neither had she shown at trial that her medical expenses exceeded the $1000 threshold. Consequently, the trial court dismissed the complaint.[180]

1. Restricting Recovery for Pain and Suffering—On appeal, the Florida Supreme Court again indicated that the act's exemption from tort liability for property damage was invalid as it unconstitutionally "denied the right of access to the courts."[181] After

176. 281 So. 2d at 4, (commenting on Rotwein v. Gersten, 160 Fla. 736, 37 So. 2d 419 (1948)).

177. 296 So. 2d 9 (Fla. 1974).

178. *Id.* at 12.

179. *Id.*

180. *Id.* at 12, 13.

181. *Id.* (citing FLA. CONST. art. I, § 21).

summarily dispatching the act's property damage section, the *Lasky* court focused its attention on the act's limitation on pain and suffering damages. Unlike the property damage section, the section limiting recovery for pain and suffering *required* owners of motor vehicles to maintain security for payment of no fault benefits.[182] Tort immunity protected the properly insured owner from liability for economic loss covered by the no fault benefits; it also protected that owner from noneconomic losses where the plaintiff could not meet the threshold requirements.[183]

In approving the restrictions on pain and suffering recovery, the *Lasky* court reasoned that an "exchange of rights" justified the legislative scheme.[184] In return for the right to recover for intangible harm in a few circumstances, the injured party gained prompt payment of medical expenses and wage loss without regard to fault.[185] In this manner, the limitation on pain and suffering damages, unlike the property damage provision, provided a *reasonable alternative* to the common law tort action and did not violate the constitutional guarantee of access to the courts.[186]

2. *Due Process Challenge*—The plaintiffs also contended that Florida's no fault insurance law contravened state and federal due process requirements. In assessing this allegation, the court adopted the rationality test.[187] The *Lasky* court concluded that the means selected by the legislature in the no fault act bore a reasonable relationship to permissible objectives and did not violate due process rights.[188] The deferential attitude of the court was expressed in its statement that "we do not concern ourselves with the wisdom of the legislature in choosing the means to be used, or even with whether the means chosen will in fact accomplish the intended goals."[189]

3. *Equal Protection Challenges*—The act's threshold requirements represented the focus of the plaintiff's equal protection challenges. The terms of the act, at the time of *Lasky*, permitted a

---

182. FLA. STAT. ANN. § 627.737 (West Supp. 1982).

183. *See* 296 So. 2d at 14, 15.

184. *Id.* at 15.

185. *Id.* at 14.

186. FLA. CONST. art. I, § 21.

187. 296 So. 2d at 15.

188. *Id.* at 17. The legislative objectives accepted by the court included the need for prompt payment, reduction of court congestion, reduction of insurance premiums, and elimination of some inequalities of recovery (minor claims overpaid, major claims underpaid). *Id.* at 15.

189. *Id.* at 15-16.

tort action for pain, suffering, mental anguish and inconvenience if the benefits payable for the injury exceeded the monetary threshold of $1,000; or the injury was permanent, resulted in death, or involved a fracture to a weight-bearing bone, or a compound, comminuted, displaced or compressed fracture.[190]

The standard adopted by the court in its equal protection analysis was formulated in these terms: "statutory classifications must be reasonable and non-arbitrary, and all persons in the same class must be treated alike."[191] The *Lasky* court considered the contention that the monetary threshold permitted discrimination based on wealth because poor persons who were charged less than the affluent for medical services would have greater difficulty in satisfying the $1000 requirement. The court rejected this argument on two grounds. First, the plaintiffs had not tendered the necessary proof that poor persons actually paid less for their medical care, or that wide geographical variations existed in the cost of medical care.[192] Second, the statute spoke in terms of *reasonable* and necessary medical expenses, and further directed that health care providers could charge only *reasonable* amounts for their services.[193] The court concluded, therefore, that the monetary threshold for tort actions did not violate equal protection.[194]

The court also found that the legislative classification that allowed only those suffering death or permanent injury to recover for pain and suffering was not arbitrary or unreasonable and therefore was not violative of equal protection.[195] The court explained that this classification served the legislative purposes of diminishing court congestion and delay by reducing the number of claims filed, and of providing more equitable reparations for injuries that caused long-term suffering.[196] However, the *Lasky* court declared the fracture threshold unconstitutional.[197] It viewed that provision's discriminatory impact on persons who sustained serious in-

---

190. *See* Florida Automobile Reparations Reform Act, Pub. L. No. 71-252, 1971 Fla. Laws 1366.

191. 296 So. 2d at 18.

192. *Id.*

193. *Id.* The court noted: "It is not therefore the amount of the actual bill . . . which determines if the $1,000 'threshold' has been met, but rather the reasonable amount referred to in the statute and such medical expense as shall be 'necessary.'" *Id.* (Ervin, J., dissenting in part, concurring in part).

194. *Id.* at 19, 20.

195. *Id.*

196. *Id.* at 20.

197. *Id.* at 20, 21.

juries that were not among the included fractures as a denial of equal protection.[198]

4. Right to Trial by Jury—The *Lasky* court also held that the Florida act did not deprive persons of the right to jury trial under federal or state constitutional provisions.[199] The court concluded where legislation "abolishes all right of recovery of specific items of damage in specific circumstances, as to those areas, [it] leaves nothing to be tried by a jury."[200]

5. Exchange of Rights—A possibly precarious aspect of the *Lasky* decision was the court's conclusion that no fault benefits represented a reasonable quid pro quo for limitations imposed on tort actions. This "exchange of rights" or "adequate alternative" principle is a corollary of the "access to the courts" guarantee established in *Kluger*.[201] It qualifies legislative freedom to abrogate existing statutory or common law rights of action by requiring a proper substitute or alternative to replace the abolished right. The *Lasky* court articulated the act's substitute rights which justified curtailment of the traditional tort action: allaying the hardship of delay (and uncertainty) in recovery for losses by prompt payment of basic benefits without regard to fault, and by granting partial tort immunity.[202]

In making this assessment, the court attempted to balance an equation, factoring in new advantages to offset the subtracted tort rights. This analysis implies a continuing weighing process, should the act's major terms be altered significantly, rather than a static, unreviewable approval.

Material changes since the *Lasky* decision, however, have diluted Florida's no fault scheme. The legislature eliminated the $1000 threshold and revised the verbal threshold in 1976.[203] It later

---

198. *Id.* The Court stated:

> A person involved in an accident who suffered a fractured skull which is not a compound, comminuted, displaced or compressed fracture, may not maintain an action for pain and suffering under this provision, unless the fracture is considered to be a permanent injury. The same is not true as to the weight-bearing little toe. One who suffers a soft tissue injury may not seek recompense for pain and suffering unless it can be proved that the injury is permanent; yet these have been shown to be among the most serious of bodily injuries.

Id. at 20.

199. *Id.* at 22, (interpreting U.S. CONST. amends. 7, 14; FLA. CONST. art. I, § 22, and citing Mountain Timber Co. v. Washington, 243 U.S. 219).

200. 296 So. 2d at 22.

201. *See supra* notes 172-174 and accompanying text.

202. 296 So. 2d at 14.

203. 1976 Fla. Laws at 722-23.

reduced the percentage of medical expenses and lost income payable as benefits,[204] raised the limit on allowable deductibles from $1,000 to $4,000[205] and removed the bodily injury liability insurance requirement of $10,000 per person and $20,000 per accident.[206] Effective January 1, 1979, the ceiling on personal injury benefits was increased to $10,000 and the maximum allowable deductible to $8000.[207]

These amendments significantly diminished the protection of the victim who had involuntarily bartered the traditional tort action for guaranteed no fault payments. The $8000 deductible, which would likely be elected by those least able to bear accidental economic loss, would permit a minimal recovery under the statute. Following the 1977 and 1978 amendments, the residual protection provided by the act was clearly open to challenge as an unreasonable, inadequate alternative.[208]

D. Chapman v. Dillon

A challenge to the protection provided by the act materialized in *Chapman v. Dillon*,[209] in which the 1979 provisions of the no fault statute were attacked as a denial of access to the courts, a denial of due process and a violation of equal protection. Although the Florida Fifth District Court of Appeal held that the post-*Lasky* modifications lowering personal injury protection benefits (PIP)[210] and increasing permitted optional deductibles[211] were unconstitutional, the Florida Supreme Court sustained all aspects of the amended no fault law.

The court held that victims of automobile accidents were not denied access to the courts by the reduction in benefits and increased deductibles.[212] It stated that, "the changes made by the

---

204. 1978 Fla. Laws at 1043.

205. *Id.* at 1044.

206. 1976 Fla. Laws at 717-18.

207. FLA. STAT. ANN. § 627.736(1) (West Supp. 1982).

208. *See* Note, *No Fault Auto Insurance: Is Eliminating Pain and Suffering a Viable Option Under the Florida Constitution*, 30 U. FLA. L. REV. 445 (1978).

209. 415 So. 2d 12 (Fla. 1982) *rev'g* 404 So. 2d 354 (Fla. 5th Dist. Ct. App. 1981).

210. FLA. STAT. ANN. § 627.736(1) (West Supp. 1982).

211. *Id.* § 627.739(1).

212. 415 So. 2d at 17. The district court had concluded that lowering the benefits and increasing the deductibles had reduced compensation and increased litigation to the point where the no fault no longer provided a reasonable alternative for the redress of injuries. *See id.* The Supreme Court of Florida disagreed:

> Many motorists . . . are covered by some other type of insurance or benefit program that would help pay for their medical expenses and lost income. . . . The benefits

legislature [since *Lasky*] have not fundamentally changed this essential characteristic of the no-fault law."[213] The Florida Supreme Court rejected as well the lower court's conclusion that modifications of the statute were not reasonably related to permissible objectives and therefore contravened due process. "[T]he injured party still recovers most of his out-of-pocket expenses from his own insurer and is allowed to bring suit for the remainder . . . . The amount of PIP coverage . . . is sufficient to prevent a party from being forced into dire financial circumstances and accepting unduly small settlements."[214] Finally, the Florida Supreme Court indicated that the permanent injury threshold for tort actions seeking noneconomic damages did not deny equal protection to persons who have suffered serious but not permanent injury.[215] The court followed *Lasky* in concluding that distinguishing between persons permanently injured and those suffering temporary disability was reasonable and not arbitrary.

The Florida Supreme Court decision appears to be a clear affirmation of the post-*Lasky* legislative amendments of Florida's no fault plan. However, in concurring, or concurring and dissenting opinions, three of the six justices raised caveats that should not be ignored by the Florida legislature. The concurring opinion stated that the "act as amended is at the absolute outer limits of constitutional parameters" in reducing benefits for medical expenses and lost earnings.[216] The two concurring and dissenting justices observed that these same amendments reach "the outer limits of constitutional tolerance."[217] It seems probable, therefore, that any further legislative dilution of Florida's no fault coverage may be found to contravene access to the courts and due process guarantees.

---

from these collateral sources are often more than sufficient to pay for the expenses not included in the PIP coverage. Thus motorists entitled to these collateral benefits would receive full compensation without needing to file a suit.

415 So. 2d at 17.

213. *Id.*

214. *Id.* at 18. The court interpreted the purpose of raising the permissible deductible as a prevention, where car owners have other insurance, of duplicative coverage.

215. *Id.*

216. *Id.* at 19 (Overton, J., concurring).

217. *Id.* (Sundberg, C.J., and Adkins, J., concurring and dissenting).

VIII. KANSAS

A. *Introduction*

A short-lived version of a no fault plan was enacted in Kansas in 1973, repealed in 1974[218] and replaced by the present Kansas Automobile Injury Reparations Act.[219] The present act requires that every owner of a motor vehicle have liability insurance or qualify as a self-insurer, and specifies penalites for noncompliance.[220] The prescribed liability limits are $15,000 per individual in one accident, $30,000 for all individuals in one accident for bodily injury or death and $5,000 for property damage in any one accident.[221] Every insurance policy issued to a resident of Kansas must also include no fault PIP benefits.

Except for usual exclusions,[222] and for motorcycle owners who may reject PIP benefits,[223] covered persons are entitled to medical and funeral expenses, disability, rehabilitation, survivors' benefits and compensation for substituted services.[224] An accident victim may also sue for noneconomic loss such as pain and suffering only if the reasonable value[225] of medical expenses aggregate $500 or more, or if the injury resulted in death or one of the specified types of serious injury.[226] Aware that this pain and suffering recovery

---

218. No Fault Insurance Act, ch. 198, 1973 Kan. Sess. Laws 1, *repealed by* No Fault Insurance Act, ch. 193, 1974 Kan. Sess. Laws 22.

219. Automobile Injury Reparations Act, ch. 193, 1974 Kan. Sess. Laws 1 (codified as amended at KAN. STAT. ANN. §§ 40-3101 to -3121 (1981)).

220. KAN. STAT. ANN. § 40-3104 (1981).

221. *Id.* § 40-3107(e). The act disallows motor vehicle registeration in Kansas unless the owner has the required liability coverage. *Id.* § 40-3118.

222. *Id.* § 40-3107(f). An insurer can exclude coverage for a work related injury, for self-inflicted intentional injury, for an injury suffered by an unauthorized user of the vehicle and for injuries suffered by the insured or members of his household if these injuries are suffered in a vehicle that is insured under the policy. *Id.*

223. *Id.* § 40-3108.

224. *Id.* § 40-3103, -3107(f). The coverage specifically includes: reasonable medical expenses up to $2000; disability benefits (100% of monthly earnings if not income tax exempt or 85% if tax exempt) up to a maximum of $650 per month for one year from the date of incapacity; funeral expenses not to exceed $1000; rehabilitation benefits (occupational therapy, psychiatric services) up to $2000; substituted services, up to $12 a day for one year, and survivor's benefits, which include substituted services and monthly earnings benefits up to $650 for a maximum of one year after the injured person's death. *Id.*

225. Treatment charges must be reasonable, that is, they must be what the person or institution would charge in cases not involving insurance. *Id.* § 40-3111.

226. Serious injuries specified by the act are: "[P]ermanent disfigurement, a fracture to a weight bearing bone, a compound, comminuted, displaced or compressed fracture, loss of a body member, permanent injury, . . . permanent loss of a bodily function or death." *Id.* § 40-3117.

limitation was the heart of the no fault scheme, the Kansas legislature declared it to be nonseverable.[227]

B. *Manzanares v. Bell*

The Kansas no fault plan was challenged on constitutional grounds in *Manzanares v. Bell*,[228] an action against the State Commissioner of Insurance to enjoin the enforcement of the original no fault act.[229] The district court declared that statute unconstitutional.[230] While an appeal was pending, the legislature repealed the first act and adopted new legislation correcting the constitutional infirmities found by the district court.[231] Consequently, the *Manzanares* court reviewed the constitutionality of the revised statute.[232]

1. Fundamental Right to Travel—In *Manzanares*, the plaintiff claimed that the compulsory insurance provisions of the second act impermissibly burdened his fundamental right to travel, a burden that could not be justified absent a compelling state interest.[233] The court rejected that contention, concluding that the no fault legislation did not affect the *right to travel*. Rather, the court decided that the Automobile Reparations Act was directed at the *privilege*, not the right, of operating a motor vehicle on Kansas highways. Exercise of that privilege could be subject to a condition precedent—providing insurance coverage.[234]

2. Due Process—The plaintiff further argued that the no fault law deprived individuals of freedom of choice, and that the premium payments for first party coverage constituted taking of property without due process of law.[235] The plaintiff also alleged that the imposition of a $500 threshold was a taking of property without compensation.[236] These objections were grounded on the fourteenth amendment of the United States Constitution, and sec-

---

227. *Id.* § 40-3121.

228. 214 Kan. 589, 522 P.2d 1291 (1974).

229. 522 P.2d at 1301.

230. *See id.*

231. The Automobile Injury Reparations Act, ch. 193, 1974 Kan. Sess. Laws 1, 22 (codified as amended at KAN. STAT. ANN. §§ 40-3101 to -3121 (1981)), corrected the areas in the former act which the district court had found unconstitutional.

232. 522 P.2d at 1301. It was to the constitutionality of the 1974 act that the *Manzanares* court directed its attention.

233. *Id.*

234. *Id.* at 1302.

235. *Id.* at 1303.

236. *Id.* at 1299.

tion 18 of the Kansas Bill of Rights, which provides: "Justice without delay. All persons, for injuries suffered in person, reputation or property, shall have remedy by due course of law, and justice administered without delay."[237]

In assessing these due process challenges, the *Manzanares* court applied a rationality standard of review established in prior cases involving regulatory legislation.[238] In determining whether the statute bore a reasonable relation to a permissible goal, the court relied on a presumption of constitutionality.[239] That deference to legislative determination was premised on the existence of a state of facts that would justify the legislation.[240] The court considered in detail the legislative history of insurance regulation in Kansas,[241] the results of studies undertaken by state and federal governments and independent scholars establishing the shortcomings of the traditional tort reparations system,[242] and the proper scope of the police power of the state.[243] On the basis of this extensive due process analysis, the court concluded that the no fault statute "being reasonably directed toward problems that affect the public welfare, including the economic welfare of the state and its citizens, . . . represents a proper and legitimate exercise of the police power of the state."[244]

3. Equal Protection—The plaintiffs argued that the limitations on tort recovery for pain and suffering invidiously discriminated between automobile accident victims and those injured in other circumstances,[245] and that the $500 threshold would operate to discriminate against the economically disadvantaged.[246] Additionally, the plaintiffs contended that the Kansas act granted motorcycle owners an arbitrary and unjustifiable privilege by giving

---

237. KAN. CONST. Bill of Rights § 18.

238. 522 P.2d at 1303 (citing City of Colby v. Hurtt, 212 Kan. 113, 509 P.2d 1142 (1973); City of Lyons v. Sutter, 209 Kan. 735, 498 P.2d 9 (1972)).

239. *Id.* (quoting Tri-State Hotel Co. v. Londerholm, 195 Kan. 748, 760, 408 P.2d 877, 888 (1965)): "All doubts of invalidity must be resolved in favor of the law. It is not our province to weigh the desirabiltiy of social or economic policy underlying the statute or to question its wisdom."

240. *Id.* at 1303 (citing McGowan v. Maryland, 366 U.S. 420, 426-27 (1961); Munn v. Illinois, 94 U.S. 113 (1867)).

241. *Id.* at 1305-06.

242. *Id.* at 1304.

243. *Id.* at 1306.

244. *Id.* at 1307.

245. *Id.* at 1308.

246. *Id.* at 1309.

them the option to reject compulsory first party coverage.[247]

The Kansas standard for constitutional review of equal protection challenges requires that a classification bear a rational relation to the legislative goal.[248] Applying this test to the no fault law, the court found that the $500 threshold provision emanated from a proper purpose because it was intended to eliminate minor and exaggerated claims for pain and suffering.[249] Further, because the statute specified that medical expenses must be judged on the basis of "reasonable value" of medical services, courts had flexibility in preventing discriminatory application of the threshold.[250]

The *Manzanares* court observed that the legislature enjoyed a "wide range of discretion" in classifying and distinguishing those subject to the act.[251] In the exercise of its police power to regulate the use of the highway, the legislature could legitimately treat motorcycle owners and operators differently because the character and use of the vehicle, coupled with the high risk of injury to its occupants, would result in substantially higher PIP premiums than those for automobile owners.[252] The motorcycle owner, however, was not exempted from purchasing third party liability insurance.

One justice, concurring and dissenting, expressed his view that the tort action threshold violated section 18 of the Kansas Bill of Rights.[253] His reasoning paralleled the substantive access to courts analysis of the Florida Supreme Court in *Kluger v. White*,[254] that absent an overriding public necessity, the legislature lacked authority to abolish a common law right without substituting a proper alternative.[255]

C. *Conclusion*

The Supreme Court of Kansas registered clear approval in its

---

247. *Id. See* KAN. STAT. ANN. § 40-3107(f) (1981).

248. *See* Henry v. Bauder, 213 Kan. 751, 518 P.2d 362 (1974); Pinkerton v. Schwiethale, 208 Kan. 596, 493 P.2d 200 (1972).

249. 522 P.2d at 1308.

250. *Id.*

251. *Id.* at 1310 (citing Orient Ins. Co. v. Daggs, 172 U.S. 557 (1898); Shelton v. Phalen, 214 Kan. 54, 519 P.2d 754 (1974); State v. Weathers, 205 Kan. 329, 469 P.2d 292 (1970); Tri-State Hotel Co. v. Londerholm, 195 Kan. 748, 408 P.2d 877 (1965); Martin v. Davis, 187 Kan. 473, 357 P.2d 782 (1960), *appeal dismissed*, 368 U.S. 25 (1961)).

252. 522 P.2d at 1310-11. *See supra* note 223 and accompanying text.

253. *Id.* at 1318.

254. 281 So. 2d 1 (Fla. 1973). *See supra* notes 165-176 and accompanying text.

255. 522 P.2d at 1319. The three additional dissenting and concurring justices took issue with the motorcyclist's option to reject first party insurance, but did not voice objections to the the tort threshold provision.

constitutional review of the Kansas Automobile Injury Reparations Act. The *Manzanares* case focused on the crucial tort action limitation that the legislature had declared inseverable from other provisions in the act.[256] A full analysis of due process and equal protection objections grounded in federal and state constitutional law confirmed the act's validity. As in the other states, the obvious constitutional challenges to the Kansas no fault legislation were launched, and repulsed by the application of an indulgent, rationality standard of review.

IX. CONNECTICUT

A. *Introduction*

The Connecticut legislature enacted the No-Fault Motor Vehicle Insurance Law in 1972.[257] It provided mandatory security requirements for first party benefits and "residual tort liability,"[258] in contrast to the pre-existing law that required proof of financial responsibility only on the violation of certain statutes.[259]

The new mandatory security requirements applied to owners of private passenger motor vehicles required to be registered in Connecticut, and to owners of cars not required to be registered but that were operated, maintained or used in Connecticut.[260] Unlike the statutes of some other no fault jurisdictions, the Connecticut law did not incorporate a restriction upon the severability of any provision if parts of the law were declared invalid.

The Connecticut statute provided for economic loss benefits, which included medical and funeral expenses, work loss reparation and survivor's benefits.[261] The maximum economic loss benefits ap-

---

256. *See supra* note 227 and accompanying text.

257. Ch. 690, 1972 Conn. Acts 438 (Spec. Sess.), codified at CONN. GEN. STAT. §§ 38-319 to -351 (1981).

258. CONN. GEN. STAT. §§ 38-326, -327 (1981).

259. *Id.* § 14-112. Proof of financial responsibility means that the owner or operator of a motor vehicle must supply a certificate of insurance, bond or collateral to satisfy claims of $20,000 for death or injury to one person in one accident, $40,000 for death or injury to all persons in one accident and $5,000 for property damage. *Id.*

260. Compliance was evidenced by a policy of insurance or by qualification as a self-insurer. *Id.* § 38-326(a)-(c).

261. Specifically, the benefits covered by the act were reasonable charges for medical, surgical and other treatment, hospital care and rehabilitation; a maximum of $2000 for funeral expenses; reparation for work loss consisting either of loss of income of an employed person, or the equivalent of unemployment compensation for a person unemployed at the time of the accident, and substitute services expenses for work the injured person would have performed for himself or his family; however, benefits for work loss (excepting unem-

plicable to injury of one person was $5000.[262] Property damage was not included in the Connecticut no fault scheme.[263]

The act disallowed tort actions for economic loss or noneconomic detriment against an owner, registrant, operator or occupant of a private passenger vehicle that had the required security, unless either the injury or monetary threshold was met.[264] The injury threshold required that the victim suffer permanent injury, or fracture to any bone, or permanent disfigurement, or permanent loss of a bodily function, or loss of a body member, or death.[265] The monetary threshold required that medical and related charges exceed $400.[266]

Although the Connecticut Legislature enacted the no fault act on May 19, 1972, it delayed the effective date of the benefits and tort limitations provisions until January 1, 1973. It declared immediately effective, however, a section that encouraged any state resident to initiate an action to determine the constitutionality of the statute.[267] A declaratory judgment action was promptly instituted, but the opinion of the state's highest court was not rendered until 1975 in *Gentile v. Altermatt*.[268]

B. Gentile v. Altermatt

The plaintiffs in *Gentile* were Connecticut motor vehicle owners. Some of them had been in accidents and sustained injuries that were not sufficient to qualify them to bring tort actions. The defendants were the Commissioner of Insurance, other state officials and eight foreign and domestic companies that intervened.[269] The issues considered by the *Gentile* court mirrored questions addressed by other state courts in reviewing their no fault statutes.

1. Right to Redress—The plaintiffs claimed that limitation

---

ployment compensation) could not exceed 85% of lost income, and was subject to a $200 a week maximum. *Id.* §§ 38-319 to -320. The survivor's loss benefits included financial support to dependents that would have derived from decedent's income for work or unemployment compensation, subject to a maximum of $200 per week. *Id.*

262. *Id.* § 38-320(d).

263. *Id.* § 38-329.

264. *Id.* § 38-323.

265. *Id.*

266. *Id.* § 38-323.

267. *Id.* § 38-351(a), *repealed by* Pub. L. No. 76-436, 1976 Conn. Acts 722 (Reg. Sess.).

268. 169 Conn. 267, 363 A.2d 1 (1975), *appeal dismissed*, 423 U.S. 1041 (1976).

269. 363 A.2d at 4. The parties stipulated to the facts, reserving questions of law for the Connecticut Supreme Court.

of an accident victim's right to sue in tort[270] violated the directive contained in article I, section 10 of the Connecticut Constitution that the "courts shall be open, and every person, for an injury done to him in his person, property, or reputation, shall have remedy by due course of law, and right and justice administered without sale, denial or delay."[271] In refusing to construe this provision as an absolute bar to modification or abrogation of traditional common law, the court declared that while "the right of redress for injuries is constitutional in its nature . . . the nature of a specific injury is a right derived from the common law or statute. Insofar as [the no fault provision] merely redefines injury it does not infringe upon the right to redress."[272]

2. *Reasonable Alternative*—Paralleling arguments raised in other states, the plaintiffs in *Gentile* claimed that article I, section 10 had constitutionally incorporated the right of an accident victim to sue for tort damages. In responding to this assertion, the *Gentile* court's approach approximated that of the Florida Supreme Court in *Kluger v. White*[273] and the Supreme Judicial Court of Massachusetts in *Pinnick v. Cleary*,[274] that legislative abrogation of a common law right requires substitution of a reasonable alternative. The *Gentile* court held that the Connecticut Legislature had supplied a reasonable exchange of benefits for the restriction on tort actions.[275] That exchange consisted of prompt payment for out-of-pocket loss, partial tort immunity and reduced premiums for no fault coverage.

3. *Equal Protection*—The plaintiffs invoked the equal protection guarantees of the fourteenth amendment of the United States Constitution and article I, section 20 of the Connecticut Constitution for their contention that the tort action limitation applicable to some plaintiffs was invidiously discriminatory. The *Gentile* court applied the rational relationship test and declared that "the issue presented is whether this classification or discrimination bears a rational relationship to a legitimate state end and is

---

270. CONN. GEN. STAT. § 38-323 (1981). *See supra* notes 264-266 and accompanying text.

271. CONN. CONST. art. I, § 10.

272. 363 A.2d at 11.

273. 281 So. 2d 1 (Fla. 1973). *See supra* notes 156-176 and accompanying text.

274. 360 Mass. 1, 271 N.E.2d 592 (1971). *See supra* notes 94-117 and accompanying text.

275. The act must be viewed generally as to the benefits conferred, "the law requires a reasonable alternative and not an exact equation of remedies." 363 A.2d at 15.

based on reasons related to the pursuit of that goal".[276] It concluded that the legislature had found that the public interest would be served by removing minor injury claims from tort litigation. The $400 threshold delineated a classification related to the legislative purpose of improving the accident reparation system, and was therefore constitutional.[277]

4. Trial by Jury—Similarly, the state constitutional guarantee of the right to jury trial[278] did not, in the court's view, prohibit elimination or limitation of common law tort actions. The court apparently reasoned that if the legislature may constitutionally abrogate tort actions for plaintiffs who do not meet the statutory thresholds, those plaintiffs' right to jury trials is concurrently and legally obliterated.[279]

5. Conclusion—The decision in *Gentile* considers issues developed in parallel cases in other jurisdictions and involves no unique analysis. The Connecticut no fault statute itself is straightforward and appears to contain no substantive, intrinsic contradictions which would signal its vulnerability to future constitutional challenge.

X. KENTUCKY

A. *Introduction*

The Kentucky Legislature enacted the Motor Vehicle Reparations Act in 1974, effective July 1, 1975.[280] That act required every

---

276. *Id.* at 16. The court relied on Reed v. Reed, 404 U.S. 71 (1971); McDonald v. Board of Election Comm'rs, 394 U.S. 802 (1969); Railway Express Agency v. New York, 336 U.S. 106 (1949); Lindsley v. Natural Carbonic Gas Co., 220 U.S. 61 (1911); Barbier v. Connolly, 113 U.S. 27 (1885), for the proposition that the fourteenth amendment does not forbid the states from treating different classes of individuals differently. The Court added that "classifications must be based on natural and substantial differences, germane to the subject and purpose of the legislation." 363 A.2d at 16 (quoting Tough v. Ives, 162 Conn. 274, 292-93, 294 A.2d 67, 77 (1972)).

277. 363 A.2d at 17.

278. CONN. CONST. art I, § 19.

279. On this point, the court's language is somewhat obtuse and confusing. 363 A.2d at 17. The court ruled that "as to cases triable to the jury prior to the Constitution of 1818 . . . the right may not be abolished." *Id.* The opinion further explained, "had the legislature provided . . . that, rather than abolish the cause of action for nonexempted plaintiffs, factual issues . . . are to be tried to a commissioner or the court without a jury, this would offend article first, § 19." *Id.* A portion of the opinion relates to mandatory security provisions of the financial responsibility law of Connecticut, and is not directly germane to consideration of the no fault law.

280. Act of April 2, 1974, 1974 Ky. Acts 752 (codified at KY. REV. STAT. ANN. §§

owner of an automobile registered or operated in Kentucky to provide security for minimum tort liability[281] and "basic reparation benefits" (BRB) for net loss due to bodily injury resulting from an automobile accident.[282] The act provided a $10,000 overall maximum for BRB for any one person per accident.[283] Every person suffering injury arising from the maintenance or use of a motor vehicle in Kentucky is entitled to BRB,[284] unless he or she opts to forego the tort limitations specified in the act,[285] or intentionally causes injury.[286]

Superficially, Kentucky's Motor Vehicle Reparations Act appears straightforward enough, but it actually conceals hidden difficulties.[287] Relatively few of those difficulties have been judicially addressed. A problem sufficiently patent to have prompted early litigation, however, is the implied consent provision contained in the act.[288] That section provides that consent to the Motor Vehicle Reparations Act in general, and to the tort limitations section in particular, is implied on the part of "any person who registers, operates, maintains or uses a motor vehicle" in Kentucky.[289] The insured's tort liability for personal injury is eliminated to the extent basic reparations benefits are payable to the injured person.[290] Moreover, tort actions against owners, registrants or occupants of properly insured vehicles are available for noneconomic detriment

---

304.39-010 to -340 (Baldwin 1981)).

281. KY. REV. STAT. ANN. §§ 304.39-110 to -150 (Baldwin 1981). The act provided for minimum tort liability of $10,000 per person, per accident; $20,000 for all persons in a single accident and $5,000 for property damage in any one accident. *Id.*

282. The term "basic reparation benefits" (BRB) designates first party payments that reimburse an accident victim for net loss. Basic reparation benefits cover: (1) work loss, replacement services loss, survivor's economic loss, and survivor's replacement services loss—$200 weekly maximum; and (2) reasonable medical and rehabilitation expenses, including a $1000 maximum for funeral expenses. *Id.* § 304.39-020(5)(a), -130. Benefits a person receives or is entitled to receive from social security or workmen's compensation are subtracted in calculating net loss. *Id.* § 304.39-120(1). If loss of income benefits are not taxable, a subtraction of up to 15% is prescribed to arrive at net loss. *Id.* § 304.39-120(2). The reparation obligor (insurer, self-insurer or obligated governmental entity) is required to make available optional additional benefits. *Id.* § 304.39-140.

283. *Id.* § 304.39-020(2).

284. *Id.* § 304.39-030.

285. *Id.* § 304.39-060.

286. *Id.* § 304.39-200.

287. *See* Note, *Kentucky No-Fault; An Analysis and Interpretation*, 65 KY. L.J. 466 (1976).

288. KY. REV. STAT. ANN. § 304.39-060 (Baldwin 1981).

289. *Id.* § 304.39-060(1).

290. *Id.* § 304.39-060(2).

only if a threshold requirement is satisfied.[291] The insured, however, has the option to refuse consent to limitation of tort rights and liabilities if he files his refusal in writing with the Department of Insurance.[292] Disqualification for basic reparations benefits accompanies exercise of the option.[293] This unique feature of the Kentucky no fault scheme permits the insured to accept or reject partial abolition of tort liability.

Kentucky's unusual implied consent and option to reject tort liability limitations are explicable as devices to accommodate the state's constitutional prohibition against restrictions on the amount recoverable for death or injury to person or property.[294] The implied consent issue commanded the attention of the Kentucky Court of Appeals in *Fann v. McGuffey*,[295] the major constitutional decision upholding the Motor Vehicle Reparations Act.

B. Fann v. McGuffey

The Kentucky Court of Appeals did not concern itself with the factual context of the *Fann* case, but proceeded directly to the constitutional questions. Drawing on the reasoning of the Supreme Judicial Court of Massachusetts in *Pinnick v. Cleary*,[296] and the Kansas Supreme Court in *Manzanares v. Bell*,[297] the Kentucky Court of Appeals concluded that due process and equal protection

---

291. The threshold requirements are that the plaintiff's medical expenses must exceed $1000, or the plaintiff's injury must consist of "permanent disfigurement, fracture to a weight-bearing bone, compound, comminuted, displaced or compressed fracture, loss of body member, permanent injury . . ., permanent loss of body function, or death." *Id.* § 304.39-060(2)(b). These threshold requirements do not apply to passengers of a motorcycle. *Id.* § 304.39-060(2)(c).

292. A person may reject these restrictions on tort actions by filing a written form with the department of insurance prior to any motor vehicle accident. The rejection becomes effective on the date of filing and continues until revoked. Revocation must be in writing and becomes effective on the date of filing with the department of insurance. *Id.* § 304.39-060(4),(5).

293. A person who has countermanded the partial tort immunity of section 304.39-060 by his or her properly filed rejection may not collect basic reparations benefits, and is subject to liability for economic and non-economic loss. *Id.* § 304.39-060(7),(8). If a nonrejecter sues a rejecter, the nonrejecter may claim basic reparations benefits from the insurer of the vehicle, and noneconomic damages from the rejecter.

294. KY. CONST. § 54.

295. 534 S.W.2d 770 (Ky. 1975) (4-3 decision).

296. 360 Mass. 1, 271 N.E.2d 592 (1971). *See supra* notes 94-117 and accompanying text.

297. 214 Kan. 589, 522 P.2d 129. (1974). *See supra* notes 228-56 and accompanying text. *See also* Robinson v. Board of Regents, 475 F.2d 707 (6th Cir. 1973); Board of Educ. v. Board of Educ., 522 S.W. 2d 854 (Ky. 1975).

guarantees were preserved in Kentucky's no fault statute.[298]

Although the constitutional rights to jury trial,[299] access to the courts[300] and recovery for wrongful death might have been analyzed, the court focused instead on the constitutional prohibition of tort limitations.[301] Section 54 of the Kentucky Constitution states that "[t]he General Assembly shall have no power to limit the amount to be recovered for injuries resulting in death, or for injuries to person or property."[302] The *Fann* court took the position that if the no fault act satisifed section 54, it would survive the other possible due process and equal protection challenges, and observed that legislation embodying "public policy sufficient to justify an implied waiver of one constitutional protection ought to provide an equally reliable basis for the implied waiver of other constitutional protections that are of no higher value."[303]

1. Implied Consent—The court recognized that implied consent was a fiction, but found it a necessary fiction that was not defeated because a constitutional right was involved.[304] It referred to nonresident motorist laws, which have been upheld for many years on due process grounds.[305] Those laws subject nonresidents to restricted in personam jurisdiction for acts committed within the forum state. Similarly, the court relied on the analogy to the implied consent provisions of the Kentucky workmen's compensation laws, which had survived constitutional challenges in the past.[306]

Finally, the court reasoned that regulation of the use of the highways was within the state's police power.[307] Implied consent to the state's vehicle insurance law attached as a necessary corollary to an individual's use of the highways, whether or not the individual actually had an opportunity to reject coverage. Therefore, the option to reject coverage was not indispensable to the constitutionality of the Kentucky no fault law. As stated by the court:

---

298. 534 S.W.2d at 778-79.
299. KY CONST. § 7.
300. *Id.* § 14.
301. *Id.* § 241.
302. KY. CONST. § 54.
303. 534 S.W.2d at 776.
304. *Id.*
305. *See id.* at 777 (quoting Hess v. Pawlowski, 274 U.S. 352 (1927)).
306. 234 S.W.2d at 777. *See* Wells v. Jefferson County, 255 S.W.2d 462 (Ky. 1953) (upholding statute providing that employees would be deemed to consent to application of the worker's compensation act unless they rejected it in writing).
307. 534 S.W.2d at 777.

> Being of the opinion that a state's undoubted authority to place conditions upon the use of its highways includes the power to require liability insurance and the acceptance of a no-fault system of loss distribution, we do not regard the option to reject as indispensable to the legal presumption that any person, . . . who uses the highways accepts those conditions.[308]

2. *Exchange of Rights*—The *Fann* court had no difficulty in finding that the usual advantages attributed to first party benefits sufficed as adequate quid pro quo for diminution of the right to sue:

> [F]rom the standpoint of what the individual gives up, this really is a rather innocuous law. The compulsory insurance aspect, about which there seems to be no legal question, is likely to rip off more skin than the limitation of tort rights. Considering the modest extent to which the scope of tort recovery is constricted, the no-fault law gives much more to the many than it takes from the few.[309]

C. *Recent Decisions*

1. *Probus v. Sirles*—Kentucky next examined the implied consent issue in *Probus v. Sirles*.[310] In *Probus*, the owner-driver plaintiff was uninsured at the time of the accident. He was unemployed and allegedly without funds to pay insurance premiums. His wife, a passenger in the car, owned no car and no automobile insurance but she had recently obtained an operator's license. They sued the defendant in a common law negligence action, claiming damages for their personal injuries. The trial court dismissed the complaint on the ground that the plaintiffs had never rejected the no fault act's limitations on the right to sue in tort.[311]

The plaintiffs appealed, asserting that their rights were not governed by the implied consent stipulation.[312] They argued that denial of the right to sue because of indigency amounted to unconstitutional economic discrimination.[313] They contended, additionally, that the no fault system abridged the right to travel, and that it violated the commerce clause of the federal constitution.[314] Thus, *Probus* forced the Kentucky Court of Appeals to consider

---

308. *Id.* at 778.
309. *Id.* at 777.
310. 569 S.W.2d 707 (Ky. Ct. App. 1978).
311. *See id.* at 708.
312. *Id.*
313. *Id.* at 710.
314. *Id.* at 709.

some of the questions left open in *Fann*: the degree of use or operation of a motor vehicle necessary to subject an individual to statutory limitations on tort recovery, and whether an uninsured motorist is subject to those limitations.[315] Additionally, the plaintiffs, as indigents, raised several unusual state and federal constitutional issues.

(a) Implied Consent Absent Formal Rejection—The plaintiff-passenger contended that as a newly licensed operator she had not yet begun to drive, and therefore was not a *user* or *operator* of a motor vehicle to whom tort recovery limitations applied. Also, because her spouse had no insurance on his vehicle, she was not a member of an insured's household. Consequently, she asserted that the no fault statute did not apply to her, and she therefore retained an undiminished common law right to sue in tort.[316] The Kentucky Court of Appeals disagreed, and held that using the highways while learning to drive and obtaining a license qualified the plaintiff as a user or operator.[317]

The *Probus* court next addressed the plaintiff-owner's argument that because he was uninsured, he had preserved his constitutional right to seek damages without formally rejecting the statute's benefits and burdens. That contention had been rejected in an earlier case in which the Kentucky Court of Appeals stated that all automobile owners, whether or not insured, "are subject to the limitations of 'no-fault,' unless the owner actually rejects the limitation of his tort rights and liabilities."[318] The factor triggering application of the statute was ownership of a vehicle, not the purchase of insurance coverage.

These rulings appear logical. The difficulties that persist in implied consent arise from the statute's confusing "use-user" terminology, and from dictum in *Fann*.[319] For example, the Kentucky statute defines "use of a motor vehicle" as "any utilization of the motor vehicle as a vehicle."[320] This definition could encompass anyone who has ever traveled in an automobile. Similarly, "user" means a resident "in a household in which any person owns or maintains a motor vehicle."[321] Thus, a person could be a "user" of

315. *Id.* at 709-10.
316. *Id.* at 709.
317. *Id.*
318. *See id.* at 709-10 (citing Atchison v. Overcast, 563 S.W.2d 736, 737 (Ky. 1977)).
319. 534 S.W.2d 770 (Ky. 1975).
320. KY. REV. STAT. ANN. § 304.39-020(6) (Baldwin 1981).
321. *Id.* § 304.39-020 (15).

a motor vehicle without any actual use of a vehicle.[322]

The distinction between users and non-users is critical, however, because a non-user can sue for psychic damages without satisfying the monetary or verbal threshold.[323] An explanation of this distinction attempted by the Kentucky Court of Appeals in *Fann* and repeated in *Probus* reads:

> This limitation upon recovery for pain, suffering, mental anguish and "inconvenience" does not apply if the plaintiff was not an "owner, operator, maintainer or user" of an automobile. Hence it does not apply to an injured pedestrian unless at the time of the accident he owns or maintains an automobile, or is an operator or user in the sense that upon occasion he drives, uses, or has driven or used an automobile on the roadways of this state. In this special respect, one who "uses" an automobile (e.g., a passenger) is not a "user" unless he is a named insured in a policy with BRB coverage or is covered as a member of the named insured's household. Nor does it apply if either the injured claimant or the person against whom his claim is asserted has rejected it.[324]

Conceivably, the broad statutory definition of users could produce a very small class of "pure pedestrians" and a rather large class of perpetual passengers outside the pale of no fault law.

(b) Miscellaneous Constitutional Issues—Two other aspects of *Probus* are significant. The first is the economic discrimination charge. The indigent plaintiffs argued that because they were unable to afford insurance they were denied the right to bring suit for their injuries.[325] The court agreed that the right to sue may not be made contingent on plaintiffs' ability to pay the litigation costs,[326] but concluded that Kentucky's no fault statute did not have that effect because both the indigent plaintiff and his nonindigent counterparts had the opportunity to reject the no fault statute's tort limitations or be equally bound by them.[327] The state had "the

---

322. The 'deemed-to-have-accepted-no-fault-law-unless-formally-rejecting-it' clause of the Kentucky statute applies to anyone who "registers, operates, maintains or uses a motor vehicle" on public roads. *Id.* § 304.39-060. This states the coverage of the rule and is not a definition of "user."

323. *Id.* § 304.39-060(2)(c).

324. Fann v. McGuffey, 534 S.W.2d at 774 (footnotes omitted). *See* Lawrence v. Risen, 598 S.W.2d 474, 475 (Ky. 1980) (discussing the definition of "user").

325. 569 S.W.2d at 710.

326. *Id.* (citing KY. REV. STAT. § 453.190 (1979)).

327. *Id.* at 710. The statute states that "any person may refuse to consent." KY. REV. STAT. ANN. § 304,39-060(4) (Baldwin 1981). However, it is only the insurance applicant who must receive a rejection form and an explanation of it from the "reparation obligor" (insurer). *Id.* The uninsured motorist would presumably be charged with knowledge of the

authority to place conditions on the use of its highways."[328] The legislature could "properly determine that [indigents] not financially prepared to assume responsibility for the risks to which the general public is exposed due to their driving shall not be permitted to use the highways without making provision for insurance."[329] Equal protection was not offended by the indigent, uninsured plaintiff's preclusion from receiving basic reparations benefits.[330]

The plaintiffs also raised an unusual commerce clause attack on Kentucky's no fault legislation. The substantial burden on interstate commerce that some requirements of state insurance law arguably impose may eventually prove the Achilles heel of state regulation. The *Probus* court, however, swept aside the commerce clause question by a pronouncement that Kentucky's no fault statute created no substantial burden on interstate commerce and did not conflict with any federal regulation.[331]

At the time of the *Probus* decision, the commerce clause challenge to no fault was viewed as little more than a makeweight. Recent decisions of the United States Supreme Court, however, examine even state highway safety regulations with less deference than in the past, if a burden on interstate commerce is alleged.[332] A solidly mounted commerce clause attack, in the future, might not be so readily dislodged.

2. *Stinnett v. Mulquin*—In *Stinnett v. Mulquin*,[333] nonresidents sued for personal injuries suffered in an automobile accident in Kentucky. The Kentucky Court of Appeals remanded the case for a determination of whether or not the nonresident plaintiffs had an opportunity to reject the limitations on tort recovery embodied in the Kentucky no fault law. The opportunity to reject was termed the "cornerstone" of the statute in view of section 54 of the state constitution.[334] It is difficult to reconcile this statement with

---

rejection option.

328. 569 S.W.2d at 710.

329. *Id.*

330. *Id. See* Thomas v. Ferguson, 560 S.W.2d 835 (Ky. Ct. App. 1978).

331. 569 S.W.2d at 711. The court relied on the general principle that "[t]he power of the State to regulate the use of its highways is broad and pervasive." Bibb v. Navajo Freight Lines, 359 U.S. 520, 523 (1959). The Supreme Court in *Bibb* struck down an Illinois mudguard regulation, presumably enacted as a safety measure, because it imposed a "heavy burden" on interstate movement of trucks.

332. *See* Kassel v. Consolidated Freightways Corp., 450 U.S. 1039 (1981); Raymond Motor Transp., Inc. v. Rice, 434 U.S. 429 (1978).

333. 579 S.W.2d 374 (Ky. Ct. App. 1978).

334. *Id.* at 375.

the court's view in *Fann* that the option was not indispensable to the statute's validity.

The Court of Appeals stated that nonresidents must be accorded constitutional protection equal to residents.[335] But it is not clear how all nonresidents can be individually canvassed concerning their rejection or acceptance of Kentucky's no fault statutory limitations on the right to sue. Giving all nonresidents notice of Kentucky's no fault law and allowing them the opportunity to reject its application does not seem feasible.

XI. PENNSYLVANIA

A. *Introduction*

The Pennsylvania No Fault Motor Vehicle Insurance Act was enacted in 1974 and became effective in July, 1975.[336] It requires an owner of a vehicle that is registered or operated in Pennsylvania to maintain security for payment of first party benefits, and for third party liability coverage.[337] The act provides for basic loss benefits, additional loss benefits, and contains a limitation on tort liability.[338]

The Pennsylvania act extends beyond the usual details of a no fault scheme. It prescribes discovery rules for statements of earnings and medical records, empowers courts to order insurers to pay for rehabilitative training and to impose sanctions on injured persons refusing rehabilitation,[339] establishes an "accountability program" of surveillance over medical and vocational rehabilitation services,[340] treats releases and settlements,[341] collateral benefits,[342] choice of law[343] and provides for a unique, non-reimbursable tort fine.[344] The statute is exceptionally comprehensive and well-structured.

---

335. *Id.*

336. Act of July 19, 1974, Pub. L. No. 74-176, 1974 Pa. Laws 489 (codified at PA. STAT. ANN. tit. 40, §§ 1009.101 to .701 (Purdon Supp. 1982-1983).

337. The required liability coverage is $15,000 per individual and $30,000 per accident for bodily injury or death and $5,000 per accident for property damage. PA. STAT. ANN. tit. 40, § 1009.104(a) (Purdon Supp. 1982-1983).

338. *See Id.* §§ 1009.202 to .207, .301.

339. *Id.* §§ 1009.401 to .408.

340. *Id.* § 1009.109(c).

341. *Id.* § 1009.106(b).

342. *Id.* § 1009.203.

343. *Id.* § 1009.110(c).

344. *Id.* § 1009.301(b).

1. Major No Fault Provisions—The act requires "[e]very owner of a motor vehicle" registered or operated in Pennsylvania to provide security "for the payment of basic loss benefits."[345] The term "basic loss benefits" refers to net loss resulting from injury arising out of the use or maintenance of a motor vehicle.[346] Property damage is excluded .[347] Four categories of basic loss benefits are provided: (1) allowable expenses for medical, health and funeral costs,[348] (2) work loss benefits,[349] (3) replacement services loss,[350] and (4) survivors' loss.[351] There is no overall limit on allowable expenses other than the caveat that physician and hospital charges for treatment must be reasonable.[352]

The statute also mandates that insurers offer optional coverage for losses exceeding basic benefits and for damage to property and physical damage to motor vehicles.[353]

2. Tort Liability—

(a) Economic Loss. A motor vehicle owner who maintains the required security is not liable in tort for loss covered by basic loss benefits, but remains liable for economic losses that exceed those no fault benefits. Certain persons do not enjoy the partial tort immunity provided by the act, however: those whose act or omission in manufacturing or repairing a motor vehicle results in a defect that causes injury, those who intentionally injure themselves or others,[354] and those who own or operate a motorcycle.[355]

---

345. *Id.* § 1009.104(a).

346. *Id.* § 1009.103.

347. *Id.* § 1009.202.

348. Allowable expenses cover reasonable charges for "(A) professional medical treatment and care; (B) emergency health services; (C) medical and vocational rehabilitation services; (D) expenses directly related to the funeral, burial . . . of a deceased victim not to exceed . . . $1,500." *Id.* § 1009.40.

349. Work loss refers to gross income, subject to the rules for calculating "probable weekly income," for the regularly employed, the seasonally employed and the unemployed victim; maximum work loss benefit is $15,000. *Id.* §§ 1009.103, .202(b)(2), .205.

350. Replacement service loss covers "expenses incurred in obtaining ordinary and necessary services [that] the victim would have performed . . . for . . . his family." *Id.* § 1009.103. The maximum for such expense is $25 per day for an aggregate period of one year. *Id.* § 1009.202(c).

351. Survivors' loss refers to "loss of income of a deceased victim" and the expense of substitute services which the victim would have performed. *Id.* § 1009.103. Maximum recovery is $5000. *Id.* § 1009.202(d).

352. The act provides: "In no event, however, may a charge by any person or institution be in excess of the amount the person or institution customarily charges . . . in cases involving no insurance." *Id.* § 1009.407.

353. *Id.* § 1009.207(a).

354. *Id.* § 1009.103.

355. *See id.* § 1009.301.

(b) Noneconomic Loss. Under the act, a person is liable in tort and, conversely, a victim or his representative may sue if an automobile accident results in death, serious or permanent injury, permanent and irreparable disfigurement, or physical or mental incapacity preventing the victim from performing substantially all normal activities for more than sixty consecutive days.[356] A tort action may be pursued as well if a victim incurs reasonable charges for necessary medical expenses that exceed $750.[357] The Pennsylvania act thus prescribes alternative verbal and monetary thresholds sustaining a cause of action for pain and suffering damages. Potential tort liability is also preserved for persons against whom non-reimburseable tort fines are assessed in proceedings relating to vehicle accidents.[358]

The importance of the partial abolition of tort liability is evidenced by the legislature's decision to make all provisions of Pennsylvania's no fault act severable except that section.[359] Thus, invalidation of its restrictions on tort actions would be fatal to the statute. Although the Illinois no fault statute was overturned by such a challenge in *Grace v. Howlett*,[360] the Pennsylvania law has survived a similar attack.

B. Singer v. Shepard

The only major constitutional challenge to Pennsylvania's no fault act occurred in *Singer v. Shepard*.[361] *Singer* aimed directly at the act's partial abolition of tort liability.[362] The plaintiff sought a declaration that limitations on tort actions violated the state and federal constitutions.

1. Limitation on Tort Recoveries—The plaintiff in *Singer* claimed that the denial of damages for noneconomic detriment to those whose injuries do not qualify as serious[363] or whose medical expenses do not exceed the $750 threshold, violated article III, section 18 of the Pennsylvania Constitution by prescribing an impermissible limitation on the amount of recovery.[364] Article III, sec-

---

356. *Id.* § 1009.301(a)(5).
357. *Id.*
358. *Id.* § 1009.301(b).
359. *Id.* § 1009.503.
360. *See supra* notes 62-77 and accompanying text.
361. 464 Pa. 387, 346 A.2d 897 (1975).
362. PA. STAT. ANN. tit. 40, § 1009.301(a) (Purdon Supp. 1982-1983).
363. *See id.* § 1009.301(a).
364. *See* 346 A.2d at 902.

tion 18 provides:

> The General Assembly may enact laws requiring the payment by employers, or employers and employes [sic] jointly, of reasonable compensation for injuries to employes [sic] arising in the course of their employment, . . . and the maximum and minimum limits thereof . . . but in no other cases shall the General Assembly limit the amount to be recovered for injuries resulting in death, or for injuries to persons or property, and in case of death from such injuries, the right of action shall survive, and the General Assembly shall prescribe for whose benefit such actions shall be prosecuted.[365]

Rejecting that contention, the Pennsylvania Supreme Court reasoned that although the no fault act creates two classes—persons who are entitled only to economic loss recovery and persons who are entitled also to noneconomic recovery—each class may recover in its category of compensable losses without limitation.[366] It therefore found no constitutional violation in the act's limitation of monetary recovery.[367]

2. *Abolition of Right of Action*—The plaintiff also argued that elimination of the right to a remedy for some accident victims ran counter to article I, section 2 of the Pennsylvania Constitution. That section reads: "All courts shall be open, and every man for an injury done him . . . shall have remedy by due course of law, and right and justice administered without sale, denial or delay."[368] This is the same argument that appears in the constitutional cases in other states.[369]

The Pennsylvania Supreme Court treated the constitutional "access to courts" as a procedural guarantee. It first noted that the legislature may modify or abolish common law rights of action.[370]

---

365. PENN. CONST. art. III, § 18. In *Singer*, the Pennsylvania Supreme Court highlighted the history of this section. Its predecessor, art. III, § 21 of the 1874 Constitution invalidated an act of 1868 which placed a maximum on the amount a plaintiff could recover in a negligence action against a common carrier. An amendment was adopted in 1915 to provide an exception for workmens' compensation. *See* 346 A.2d at 900-02. By way of comparison, it may be noted that the present constitution of New York forbids a limitation on recovery in a wrongful death action. *See* N.Y. CONST. art. I, § 16.

366. 346 A.2d at 901.

367. *Id.* at 902. The court decided that the act did not limit the amount of recovery because "[w]hat Section 301(a) abolishes is the *right* of those injured parties to recover in tort." *Id.* Nothing in article III, section 18 prevented the abolition of a cause of action. *Id.*

368. PENN. CONST. art. I, § 2.

369. *See, e.g.*, Lasky v. State Farm Ins. Co., 296 So. 2d 9 (Fla. 1974). *Lasky* is discussed *supra* at note 181 and accompanying text.

370. 346 A.2d at 903 (citing Jackman v. Rosenbaum Co., 263 Pa. 158, 106 A. 238 (1919), *aff'd*, 260 U.S. 22 (1922)). *See* Sherwood v. Elgart, 383 Pa. 110, 117 A.2d 899 (1955).

The advocacy of immutable common law is essentially a due process argument, and the *Singer* court believed that the issue had been foreclosed by the United States Supreme Court's reasoning in *Munn v. Illinois*:[371]

> A person has no property, no vested interest, in any rule of the common law. That is only one of the forms of municipal law, and is no more sacred than any other. Rights of property which have been created by the common law cannot be taken away without due process; but the law itself, as a rule of conduct, may be changed at the will, or even at the whim, of the legislature, unless prevented by constitutional limitations. Indeed, the great office of statutes is to remedy defects in the common law as they are developed, and to adapt it to changes of time and circumstances.[372]

The *Singer* court therefore concluded that when the legislature eliminated a common law right of action, there was no legally cognizable injury upon which article I, section 2 of the Pennsylvania Constitution could operate, and consequently no violation of the "access to courts" guarantee.[373]

3. Equal Protection—The *Singer* court also ruled that the Pennsylvania no fault law did not violate equal protection guarantees.[374] The statute created at least two classes. The first class retained the right to full common law tort damages and the second was confined to no fault recovery. The court recognized that the classification produced "unequal distribution of benefits or imposition of burdens."[375]

The court noted, however, that unequal treatment is not per se unconstitutional.[376] Instead, it indicated that the fourteenth amendment is violated when there is invidious discrimination against a suspect class or if a fundamental right is impermissibly impaired.[377] The plaintiff conceded that neither a suspect class nor a fundamental right was at issue. Accordingly, the Pennsylvania

---

371. 94 U.S. 113 (1876).

372. 346 A.2d at 903 (quoting Munn v. Illinois, 94 U.S. 113, 134 (1876)).

373. *Id.*

374. *Id.* at 906.

375. *Id.* at 904 (quoting Commonwealth v. Webster, 462 Pa. 124, 337 A.2d 914, 917 (1975).

376. *Id.*

377. *Id.* To support its position, the court cited, among other cases, Loving v. Virginia, 388 U.S. 1, (1967) (holding unconstitutional state prohibition against inter-race marriages for the purpose of preserving the purity of the white race); Dunn v. Blumstein, 405 U.S. 331 (1972) (holding that a one year required state residency period for voter qualification impairs fundamental right to travel).

Supreme Court applied a rationality standard in reviewing the no fault act's classifications of accident victims.

The most obvious division separating two groups of claimants—the $750 threshold—was upheld in reasoning similar to that applied by the Massachusetts court in *Pinnick v. Cleary*.[378] The choice of a particular monetary figure was a question for the legislature. The court held that the monetary threshold was rationally related to the legitimate objectives of the no fault statute—reasonably priced insurance and prompt and adequate payment for basic losses.[379] Similarly, it concluded that the other items of differential treatment in the tort recovery rules of the Pennsylvania act also rested on rational grounds of deterrence or cost.[380]

C. *Conclusion*

In a close decision, the Pennsylvania Supreme Court upheld the crucial inseparable section of the no fault statute which abolishes recovery of general damages unless the accident victim accumulates medical expenses exceeding $750, or suffers serious injury, or, unless the one causing the injury does so intentionally or while uninsured, or as the manufacturer or repairer of a defective vehicle. The court perceived no violation of the fourteenth amendment equal protection clause of the United States Constitution and no violation of the state constitutional guarantees against limitations on tort recoveries or access to the courts.[381]

Having survived this challenge to its most critical section, the Pennsylvania no fault statute seems invulnerable to a general equal protection attack. It would also be difficult to challenge the statute on the type of due process grounds that influenced the

---

378. 360 Mass. 1, 271 N.E.2d 592 (1971), discussed *supra* at notes 94-117 and accompanying text; *see* Goodman v. Kennedy, 459 Pa. 313, 329 A.2d 224, 228-29 (1974) (abjuring judicial supervision over legislative wisdom in selecting the boundaries of classifications). *But see* Grace v. Howlett, 51 Ill. 2d 478, 283 N.E.2d 474 (1972), discussed *supra* notes 62-77 and accompanying text.

379. 346 A.2d at 905.

380. *Id.* at 906-07. The tort recovery rules are codified at PA. STAT. ANN. tit. 40 § 1009.301(a) (Purdon Supp. 1982-83). Both the cost and deterrence factors can explain the exclusion of motorcycles owners from the class of no fault insureds and their continued full liability in tort. *Id.* § 1009.301(a)(6).

381. The dissenting opinion in *Singer*, however, forcefully argued that Pennsylvania's no fault statute impermissibly limits the amount that may be recovered in an action for injuries or death. 346 A.2d at 910 (Eagen J., dissenting).

Michigan Supreme Court in *Shavers v. Kelley*,[382] because the ratemaking and consumer protection provisions of Pennsylvania's insurance law appear to be more concerned with fairness and governmental accountability than pre-*Shavers* Michigan law. However, detailed study might expose due process flaws. The narrow margin of decision in *Singer* turned on the application of article III, section 18 of the Pennsylvania Constitution to the no fault law, and on that issue the result cannot be otherwise in a future case unless the court chooses directly to overrule itself.

XII. NEW JERSEY

A. *Introduction*

The New Jersey no fault scheme was enacted as part of a comprehensive program requiring owners of automobiles registered or garaged in the state to maintain first party and liability coverage. The New Jersey Automobile Reparation Reform Act became effective in 1972.[383] The legislature directed that the act be liberally construed and that its provisions should be severable.[384] The act provides for compulsory liability coverage consisting of a limit of $15,000 for injury to or death of one person in one accident, $30,000 for all injuries or deaths in one accident, and $5,000 for damage to property in one accident.[385] The statute prohibits an insurer from refusing to renew the no fault and liability coverage required by the act without the consent of the insurance commissioner.[386] The act also contains rather standard prescriptions relating to additional coverage, discovery and other matters. Close examination of the New Jersey act, however, reveals several distinctive features as well.

1. No Fault Benefits—The act originally covered all reasonable medical expenses and continues to do so.[387] However, it was amended in 1977 to establish a common fund, financed by insurer

382. 402 Mich. 554, 267 N.W.2d 72 (1978), *aff'g in part, rev'g in part*, 65 Mich. App. 355, 237 N.W.2d 325 (1975). *See infra* notes 591-657 and accompanying text.

383. Act of June 20, 1972, ch. 70, 1972 N.J. Laws (codified at N.J. STAT. ANN. §§ 39:6A-1 to :6A-20 (West 1973)).

384. N.J. STAT. ANN. § 39:6A-16 (West 1973).

385. *Id.* § 39:6A-3.

386. *Id.* A constitutional attack upon this provision was repelled in Sheeran v. Nationwide Mut. Ins. Co., 80 N.J. 548, 404 A.2d 625 (1979).

387. N.J. STAT. ANN. § 39:6A-4(a) (West 1973).

contributions which pays any excess over $75,000.[388] The act also covers hospital expenses and defines them separately from medical expenses.[389] This distinction is important because only *medical expenses* may be tallied in aggregating the $200 threshold required for a tort action.[390] Payment for the loss of income of an "income producer"[391] and for substitute services ordinarily performed by the accident victim[392] are included. Additional no fault benefits include survivor's benefits of up to $100 a week for 52 weeks,[393] and a maximum of $1000 in funeral expenses, payable to the estate of the deceased victim.[394]

2. Tort Exemptions and Actions—

(a) Exemptions. Every person carrying the required PIP insurance is exempt from tort liability to persons entitled to PIP benefits for accidents in New Jersey if the injury is only to soft tissue[395] *and* medical expenses are less than $200.[396] These are conjunctive conditions governing tort immunity. The New Jersey statute defines medical and hospital expenses separately, with the

---

388. Act of Jan. 5, 1978, ch. 310, 1977 N.J. Laws 1215 (amending N.J. STAT. ANN. § 39:6A-4 (West 1973)). The amended provision reads:

> In the event benefits paid by an insurer pursuant to this subsection are in excess of $75,000 on account of personal injury to any one person in any one accident, such excess shall be paid by the insurer in consultation with the Unsatisfied Claim and Judgment Fund Board and shall be reimbursable to the insurer from the Unsatisfied Claim and Judgment Fund pursuant to section 2 of this act.

Id. The reference to "section 2" is to the Financial Responsibility Law. The purposes of the 1977 amendment were to relieve smaller companies of the burden of very large claims and to reduce the magnitude of loss reserves set aside for such claims which would otherwise enter into ratemaking calculations.

389. Medical expenses cover the following treatment and services: medical, surgical, dental, nursing, hospital, rehabilitation, diagnostic, medication and other necessary expenses for treatment prescribed by licensed health practitioners. *Id.* § 39:6A-2(e) (West 1973). Hospital expenses, on the other hand, appear to relate to the expenses involved in the services supplied by a hospital, and cover a semi-private room, meals, operating room, medicine, anesthetics, physiotherapy, artificial limbs and other services. *Id.* § 39:6A-2(f).

390. *Id.* § 39:6A-8. *See also infra* notes 395-397 and accompanying text.

391. N.J. STAT. ANN. § 39:6A-4(b) (West 1973). Those payments are subject to a weekly maximum of $100 and an aggregate of $5,200 (52 weeks) for any one person in one accident. *Id.*

392. *Id.* § 39:6A-4(c). The essential service benefits are subject to limits of $12 per day and a total of $4,380 (365 days) to any one person in one accident. *Id.*

393. *Id.* § 39:6A-4(d). The surviving spouse, children or the victim's estate are entitled to the benefits which would have been paid to the income producer; if the person performing essential services dies, essential services benefits are payable to the one incurring such expenses. *Id.*

394. *Id.* § 39:6A-4(e).

395. Soft tissue injury includes sprains, strains, lacerations, contusions, hematomas, tears confined to muscles, tendons, nerves, veins, arteries and similar injuries. *Id.* § 39:6A-8

396. *Id.*

medical expenses confined to treatment and services rendered and prescribed by licensed professionals. Hospital expenses include the costs of institutional, in-patient care, as well as tests, X-rays and other diagnostic charges[397] Excluding highly inflated hospital costs may give some substance to an otherwise nominal threshold of $200.

(b) Tort Actions. There are no limitations on tort actions if the injury results in death, permanent disability, disfigurement, permanent loss of bodily functions or loss of a body member.[398] Moreover, the New Jersey statute speaks in terms of exemption from "tort liability for damages" without defining damages.[399] Presumably, "damages" not only bears the usual meaning of general damages for pain, suffering and inconvenience, but also includes special damages for economic detriment in excess of first party benefits for wage loss and substitute services.

B. Rybeck v. Rybeck

The no fault statute's limitations on traditional tort actions gave rise to the first constitutional challenge to the New Jersey no fault act in *Rybeck v. Rybeck*.[400] *Rybeck* was a consolidated action in which the plaintiffs, on a motion for summary judgment, sought a declaration that the New Jersey act contravened federal and state constitutional guarantees.[401] The plaintiffs alleged that the act denied due process of law, equal protection, access to the courts and trial by jury.[402] Those issues parallel contentions raised in no fault challenges in other states. Some novel issues, however, were also introduced by *Rybeck*.

1. Due Process—The plaintiffs in *Rybeck* lodged a broad due process objection to the act. The *Rybeck* court responded by according the usual deference to legislative means and objectives and applying a rationality standard that the no fault act easily met.[403]

---

397. *Id.*

398. *Id.*

399. *Id.*

400. 141 N.J. Super. 481, 358 A.2d 828 (1976), *appeal dismissed*, 150 N.J. Super. 151, 375 A.2d 269 (1979).

401. 358 A.2d at 830-31. In the first action, Mrs. Rybeck, a passenger, sued her husband and the driver of the second car for injuries sustained in the automobile accident. She was a named insured in the policy covering her husband's car and, therefore, looked to the insurer for no fault benefits and third party recovery. In the second action, both Rybecks sued their insurer for a judgment declaring the no fault act invalid. *Id.*

402. *Id.* at 831.

403. *See id.* at 834.

The court found that the act did not violate due process because it was "legislation reasonably calculated to reform an area of the law that was not operating satisfactorily."[404]

2. *Equal Protection*—Similarly, the court rejected an equal protection challenge to the New Jersey no fault scheme.[405] Finding no suspect class or fundamental right implicated, the court applied a rational basis test. The no fault reform easily survived this test. The *Rybeck* court concluded that the act addressed rational objectives to which the classifications established by the legislature were logically relevant.[406] The equal protection challenges to the New Jersey no fault act raised in *Rybeck* were more specific than the due process arguments.

(a) Classifications Based on Threshold Requirement. For example, the plaintiffs argued that it was constitutionally impermissible to divide accident victims into those who were permitted to sue in tort and those who were relegated exclusively to no fault benefits.[407] The *Rybeck* court responded that to accomplish the legitimate goal of eliminating the cost and court congestion attributable to minor claims, the legislature was compelled to draw a line somewhere. The court found the prescribed monetary and injury threshold to be a reasonable legislative choice. The court commented: "Whether constitutionally necessary or not, a reasonable and adequate alternative was provided for the minor claimant."[408]

(b) Discrimination Against the Poor. The plaintiffs in *Rybeck* argued that the act discriminated against the poor because the poor receive medical care at lower cost, and therefore had greater difficulty meeting the $200 threshold.[409] The court rejected this argument for lack of proof. It was also alleged that the poor were disadvantaged by the cost of producing medical testimony for a preliminary hearing to determine whether or not the monetary

---

404. *Id.* The court observed: "The Constitution does not forbid the enactment of ill fated legislation. . . . It does not matter that there may have been other methods of reform the legislature might have chosen." *Id.*

405. *Id.* The equal protection clause of the New Jersey Constitution is expressed quite differently from the fourteenth amendment to the United States Constitution. The New Jersey clause provides: "No person shall be denied the enjoyment of any civil or military right, nor be segregated in the militia or in the public schools, because of religious principles, race, color, ancestry or national origin." N.J. Const., art. I, ¶5

406. 358 A.2d at 835. *See* New Jersey Chapter, Am. Insurance Planners v. New Jersey State Bd. of Professional Planners, 48 N.J. 581, 227 A.2d 313 (1967).

407. 358 A.2d at 835-36.

408. *Id.* at 837.

409. *Id.* at 837-38.

threshold was satisfied.[410] The court viewed this cost as incidental to a lawsuit and added: "Access to the courts, except in specific instances not here involved, is not constitutionally required to be subsidized."[411]

(c) Treatment of Non-Wage Earners. Plaintiffs also charged that the act discriminated against non-wage earners,[412] by providing "income continuation benefits" only to an "income producer."[413] The New Jersey statute defined "income producer" as "a person, who at the time of the accident causing personal injury or death, was in an occupational status, earning or producing income."[414] Thus, wage loss is payable under the statute only to currently employed individuals and not to those temporarily laid-off or occasionally employed. Because plaintiffs were not aggrieved by the application of this provision, however, the *Rybeck* court did not rule on the validity of excluding the unemployed from wage loss benefits.[415]

(d) Sex Discrimination. The court also rejected the plaintiff's claim that the act unconstitutionally discriminated against women. The act provides that every automobile liability policy include no fault benefits applicable to the named insured, resident family members, occupants of and pedestrians injured by the insured automobile.[416] Plaintiffs objected to this section on the ground that it deprived Mrs. Rybeck, as a married woman, of the right to purchase her own insurance or to be beyond the reach of the no fault statute. The court responded that Mrs. Rybeck was entitled to first party benefits and was restricted in her right to bring a traditional tort action not because she was a married woman, but because she was a covered passenger in the insured vehicle.[417] Mrs. Rybeck's rights and disabilities, the court concluded, were not rooted in the fact that she was the driver's spouse.[418] The court thus refuted a tenuous attempt to fashion a sex discrimination objection to the act. Mr. Rybeck would have suffered the same "discrimination" if his wife were the principal insured.

---

410. *Id.* at 838.
411. *Id.* at 838.
412. *Id.*
413. N.J. STAT. ANN. § 39:6A-4(b) (West 1973).
414. *Id.* § 39:6A-2(d).
415. 358 A.2d at 838.
416. N.J. STAT. ANN. § 39:6A-4 (West 1973).
417. 358 A.2d at 839.
418. *Id.*

3. Access to Courts—The plaintiffs further complained that the act prohibited the introduction of evidence of first party benefits paid or payable in a tort action for personal injury.[419] The argument was couched in terms of derogation of a common law negligence action. The court disposed of this contention by citing article XI, section 1, paragraph 3 of the New Jersey Constitution which states that laws in force at the time of the adoption of the current state constitution remain effective until modified or repealed.[420] The New Jersey Constitution, unlike the basic law of some other states, does not guarantee a remedy for every injury; therefore,the usual legislative authority to alter common law rights and remedies can be invoked without the need to engage in elaborate constitutional construction.[421]

4. Legislative Invasion of the Court's Authority—A provision of the New Jersey act attempted to regulate proof of loss in civil actions for bodily injury by barring evidence of the sums collectible in first party benefits.[422] Plaintiffs objected that this provision preventing proof of covered losses, legislated a rule of evidence, thus infringing on the New Jersey Supreme Court's rule-making authority and violating the separation of powers doctrine.[423] By its analysis of the statute the *Rybeck* court avoided a direct confrontation with the legislature over the court's rule-making power.

Under article VI, section 11, paragraph 3 of the state constitution, the judiciary has authority to regulate matters of procedure, but rules of evidence are not expressly included in that power.[424] The court interpreted the "curious language"[425] of the no fault act as encompassing two objectives. First, it prevented double recovery, and second, it barred testimony on out-of-pocket expenses reimbursed by the insurer. The second objective, however, could lead to confusion where a plaintiff's losses exceeded PIP benefits and the jury could be informed only of the excess expenses.[426] The

---

419. N.J. STAT. ANN. § 39:6A-12 (West 1973).

420. N.J. CONST. art. XI, § 1, ¶3.

421. The court observed: "[O]ne having a claim abrogated or limited by valid legislation is not deprived of access to the courts. . . ." 358 A.2d at 842.

422. N.J. STAT. ANN. § 39:6A-12 (West 1973).

423. 358 A.2d at 843.

424. N.J. CONST. art. VI, § 11, ¶3.

425. *Id.* at 842. The act provides: "Evidence of the amounts collectable or paid . . . to an injured person is inadmissable in a civil action for recovery of damages for bodily injury." N.J. STAT. ANN. § 39:6A-12 (West 1973).

426. The Court illustrated the potential confusion by a hypothetical set of facts:

court attempted an accommodation by interpreting the statute's section as barring evidence of covered out-of-pocket expenses as direct proof of the extent of damages, but not if covered losses were relevant to other issues.[427] In other words, where the application of the statute was substantive—to avoid double recovery—it rested within the province of the legislature; but where covered losses were relevant to other issues in controversy, the court would admit such evidence.[428] This is a hazy distinction that may require clarification in future litigation.

C. *Sheeran v. Nationwide Mutual Insurance Co.*

Issues of a different order were submitted for constitutional measure in *Sheeran v. Nationwide Mutual Insurance Co.*[429] That case concerned the interpretation and application of the act's provision prohibiting licensed insurers from refusing to renew no fault coverage.[430] Nationwide decided in 1977 to withdraw from the casualty and fire insurance market in New Jersey. It planned to renew contracts only where renewal was guaranteed, and to decline new business. However, Nationwide did not wish to surrender its license to do business in New Jersey. The Commissioner of Insurance instituted an action for a declaratory judgment that Nationwide could not continue to be licensed if it refused to renew its insurance contracts.[431] A regulation promulgated by the Commissioner listed the permissible gounds for nonrenewal and Nationwide's reason—business losses—was not among the recognized justifications.[432] In the declaratory judgment action, Nationwide not only attacked the Commissioner's application of the statute, but also contended that the mandatory nonrenewal section involved an

---

> Thus, consider the case where an injured $140 a week worker, out of work for eight weeks, received $100 a week for income continuation benefits. If the trial jury were to hear that he lost wages of $40 a week, it would be something of a puzzle to them in this day and age. The worker, say, also went to a physician 20 times and was charged $20 a visit. His PIP carrier paid only $15 a visit for 15 visits, feeling, perhaps rightfully, that the rest was unreasonable. The difference of $175 was not worth plaintiff's while to litigate with the PIP carrier. If the trial jury should hear that he owes $5 for each of his first 15 visits and then, as he recovered from his injury, he incurred a cost of $20 for each of his last five visits, the testimony will not make sense to them.

358 A.2d at 843.

427. *Id.* at 844.

428. *Id.*

429. 80 N.J. 548, 404 A.2d 625 (1979).

430. N.J. STAT. ANN. § 39:6A-3 (West 1973).

431. 404 A.2d at 625.

432. *See* 404 A.2d at 628.

invalid delegation of legislative power, denied due process and imposed an "unconstitutional condition" on licensure.

1. Application of the Mandatory Renewal Section—Nationwide argued that the no fault act's prohibition of nonrenewal by licensed insurers was intended to bar discriminatory nonrenewal of *individual* insurance contracts, and that it had no application to a general nonrenewal plan.[433] The *Sheeran* court countered that argument with a reference to the legislative history of the no fault act and determined that the statutory scheme would be frustrated if insurers could renounce renewals of existing contracts but still retain the power to insure selected applicants in the future.[434] Resorting to the "plain meaning" rule of construction, the court concluded that the act required Nationwide to surrender its license if it refused to renew existing insurance contracts.[435]

2. Improper Delegation—The *Sheeran* court responded to the argument that the act involved an invalid delegation of legislative power by noting first that the act empowered the Insurance Commissioners to promulgate "*reasonable* rules and regulations."[436] Appropriate standards therefore could be inferred from the act to keep the Commissioner's administrative discretion within proper bounds. Considering the act as a whole, the court found that the Commissioner's authority to grant or withhold consent for nonrenewals was "sufficiently circumscribed" to meet constitutional requirements.[437]

3. Due Process—In its second constitutional challenge to the act's mandatory renewal section,[438] Nationwide argued that the section's application forced it to commit resources to lines of insurance it did not wish to undertake, and thus deprived it of property without due process. The court bluntly rejected this contention.[439] Nationwide would not be compelled to write automobile insurance in New Jersey, the court indicated, if it relinquished its license. If, on the other hand, Nationwide wished to be licensed, it must accept its fair share of the market. Rather than selective renewal, the

---

433. *Id.*

434. *Id.* at 629.

435. *Id.* at 631 (citing N.J. STAT. ANN. § 39:6A-3 (West 1973).

436. *Id.* at 630 (citing N.J. STAT. ANN. §39:A-19 (West 1973)).

437. *Id.* at 630. The Commissioner "must act reasonably in light of the policies underlying the act." *Id. See* Avant v. Clifford, 67 N.J. 496, 341 A.2d 629 (1975).

438. N.J. STAT. ANN. § 39:6A-3 (West 1973).

439. 404 A.2d at 631.

court observed that Nationwide's appropriate recourse would be to seek approval of higher rates if it was not able to return a reasonable profit from existing premiums.[440]

4. Unconstitutional Condition—Responding to Nationwide's final contention that the act imposed an unconstitutional condition on licensure, the *Sheeran* court concluded that requiring a licensed carrier to renew its contracts is not an unreasonable condition on the privilege of engaging in business in the state.[441] The court found a rational connection between imposing the obligation to renew insurance contracts, and the legitimate underlying statutory purpose of protecting the public.[442]

5. Summary—The court's general approach in *Sheeran* was to uphold the Commissioner in enforcing the responsibility of carriers to act in the public interest. To rule otherwise might have opened the opportunity for highly selective practices that would have defeated the goal that each carrier must accept its fair share of the spectrum of risks where the law mandates insurance coverage for all registrants of motor vehicles.

D. Conclusion

Both *Rybeck* and *Sheeran* are significant. *Rybeck* underscores the New Jersey Supreme Court's affirmative approach to automobile accident reparations reform. *Sheeran* confirms and upholds the state's regulatory authority over the insurance industry.

Rybeck, however, provided an inadequate resolution of possible constitutional challenges to New Jersey's no fault act. The due process attack was not sharply focused. Equal protection issues, although specific, appeared muted by an incomplete record or by an absence of standing. The alleged discrimination against the poor, due to lower medical charges in depressed areas, for example, was not supported by an offer of statistical evidence.[443] Furthermore, *Rybeck* did not question the New Jersey statute's demarcation between medical and hospital expenses. There is a practical difficulty in assigning some charges to one side of the line or to the other, and the attempted differentiation only invites argument. There may be no supportable rationale for excluding steeply accelerating

---

440. *Id.* at 631.

441. *Id.*

442. *Id.*

443. By contrast, New York has recognized regional differences in physician fees and hospital costs. If similar disparities exist in New Jersey, a credible economic discrimination challenge possibly could be framed in a future case.

hospital service expenses from the threshold for tort actions.[444]

Similarly, the arguable inequity of confining wage loss first party benefits to persons actually producing income at the time of their injury fell by the wayside in *Rybeck* because plaintiffs lacked standing.[445] Clearly, this question might be revived by a person seasonally employed, recently unemployed, or to be employed in the future.[446] Other no fault statutes permit wage loss benefits in these circumstances.[447]

XIII. New York

A. *Introduction*

New York enacted its no fault law, the Comprehensive Automobile Insurance Reparations Act,[448] in 1973. In practical effect, New York became a compulsory automobile liability insurance state in 1956. The earlier Motor Vehicle Financial Security Act of 1956[449] imposed proof of financial security as a condition precedent to registration of a motor vehicle.[450] The 1956 Act supplemented the Motor Vehicle Safety Responsibility Act of 1941,[451] which permitted proof of financial responsibility by filing a certificate of insurance.[452] Thus, the New York no fault act augmented the already extensive body of legislation and administrative regulation pertaining to the ownership and operation of motor vehicles in the state.

As amended in 1979[453] the required limits of an owner's motor

---

444. The exclusion of hospital expenses to be aggregated for the $200 threshold may be predicated on the obvious fact that the cost of one day's hospitalization could easily surmount the threshold. The fault really lies in assuming that the monetary threshold, whether $200 or $1000, serves any purpose other than to challenge the ingenuity of client and counsel, inviting "padding" of medical expenses. The better approach is to rely on a verbal threshold confined to proof of serious injury.

445. 358 A.2d at 838.

446. A lower court decision has addressed this issue. In Miller v. Ohio Casualty Group, 177 N.J. Super. 328, 326 A.2d 547 (1980), the plaintiff was injured in an automobile accident the day before she planned to commence employment. The court held plaintiff was entitled to income continuation benefits under her no fault coverage equal to the difference between what she would have earned and what she was receiving in temporary disability benefits.

447. *See infra* mpte 448.

448. Act of Feb. 13, 1973, ch. 13, 1973 N.Y. Laws 56 (codified at N.Y. Ins. Law §§ 670-678 (McKinney Supp. 1981-1982)).

449. N.Y. Veh. & Traf. Law §§ 310-321 (McKinney 1970 & Supp. 1981-1982).

450. *Id.* § 312(1).

451. *Id.* §§ 330-368.

452. *Id.* § 343(a).

453. 1979 N.Y. Laws 6 (codified at N.Y. Veh. & Traf. Law § 345(b)(3) (McKinney 1970 & Supp. 1981-1982).

vehicle liability policy are: (1) $10,000 for injury to one person in an accident, (2) $20,000 for injury to two or more persons in one accident, (3) $50,000 for death of one person in one accident, and (4) $100,000 for death of two or more persons in one accident.[454] Other important provisions governing automobile liability insurance are set forth in provisions of the Insurance Law.[455]

1. No Fault Benefits—The Comprehensive Automobile Insurance Reparations Act provides payments for "basic economic loss" resulting from personal injury.[456] "Basic economic loss" is a broad category that includes most medical and some non-medical costs.[457] The act also covers lost earnings,[458] death benefits,[459] and other reasonable and necessary expenses.[460] The total payable as first party benefits is $50,000.[461] Various offsets such as workers' compensation and federal and state disability benefits are prescribed to avoid duplicative recoveries.[462] No fault insurance generally "follows the car," unless the victim is a passenger of a bus or school bus. In those circumstances, the injured party must look first to any no fault coverage on a vehicle of his or her

---

454. N.Y. VEH. & TRAF. LAW § 345(b)(3) (McKinney 1970 & Supp. 1981-1982).

455. N.Y. INS. LAW §§ 167-168 (McKinney 1966 & Supp. 1981-1982). During the same period when the New York legislature enacted those financial security and liability insurance statutes, it introduced a complex system of uninsured motorist law. The Motor Vehicle Accident Indemnification Corporation (MVAIC) Act, 1958 N.Y. Laws 1638 (codified at N.Y. INS. LAW §§ 600-626 (McKinney 1966 & Supp. 1981-1982)) must be supplemented by reference to section 167(2-a) of the New York Insurance Code to ascertain the full scope and context of New York's uninsured motorist law. Of particular relevance to no fault is the recent amendment of MVAIC that extended first party benefits to victims who are "qualified" persons. A qualified person is now defined in the statute as a New York resident who is neither an insured nor owner of an uninsured motor vehicle. N.Y. INS. LAW §§ 601(b) (McKinney Supp. 1981-1982). The complexities of MVAIC are not analyzed here. However, there appears to be little indication that the implications of section 621-a and its effective coordination with New York's no fault statute have been realized.

456. N.Y. INS. LAW § 671(1)(a) (McKinney Supp. 1981-1982).

457. Basic economic loss covers necessary expenses for medical, hospital, surgical and similar services, psychiatric and physical therapy, rehabilitation, non-medical care based on a recognized religious practice, and any other professional health care. *Id.* § 671(1)(a). However, charges must conform to the schedule of fees established by the state Workers' Compensation Board. *Id.* § 678.

458. *Id.* § 671(1)(b),(2)(a). The maximum reimbursement is $1000 per month for three years. Lost earnings are subject to a 20% reduction in calculating first party benefits. *Id.*

459. *Id.* § 672(1)(c). The estate of a covered person is entitled to $2000 for the death of that person. *Id.* Covered persons are the named insured, household members, and persons not occupants of another motor vehicle or motorcycle. *Id.* §§ 671(10); 672(1)(a), (b). Reimbursement is limited to $25 a day for one year from the date of the injury. *Id.* § 671(1)(c).

460. *Id.* § 671(1)(c).

461. *Id.* § 671(1).

462. *Id.* § 671(2)(b).

household.[463]

2. Tort Actions and Exemptions—The New York statute provides that in an action brought by a "covered person" against a "covered person" for personal injury resulting from a motor vehicle accident within the state, there is no right of recovery for basic economic loss, or for noneconomic detriment, unless a serious injury has occurred.[464] "Covered person" is defined as a pedestrian injured by a motor vehicle, or an owner, operator or occupant of a properly insured motor vehicle.[465] The owner, operator or occupant of a motorcycle who maintains the requisite financial security[466] enjoys the same immunity as a covered person from a tort action.[467]

A noncovered person forfeits tort immunity and may be sued for the full loss resulting from personal injury.[468] The noncovered person is one who has failed to comply with the compulsory insurance provisions or who may be excluded by the insured from coverage on the usual grounds, such as intentionally causing his own injury.[469] Where a covered person recovers damages in an action against a noncovered person, the insurer who has paid first party benefits has a lien against the recovery to the extent of benefits paid or payable to the covered person.[470]

The statute does not address the question of tort immunity for the covered person when he or she is sued by a noncovered person. If in that case the covered defendant is open to full tort liability,

---

463. *Id.* § 672(1)(a). The purpose of this provision is to mitigate the high cost of insurance, especially for school districts.

464. *Id.* § 673(1). A serious injury is defined as personal injury resulting in death, dismemberment, significant disfigurement, fracture, permanent loss or significant limitation or use of a body organ, member, functon or system, or non-permanent injury which prevents performance of usual activities for at least 90 days during the 180 days following the injury. *Id.* § 671(4). The cognizance of nonpermanent injuries of a prescribed duration and disabling consequence is an unusual provision of the New York Act.

465. *Id.* § 671(10).

466. N.Y. VEH. & TRAF. LAW §§ 310-321 (McKinney 1970 & Supp. 1981-1982).

467. N.Y. INS. LAW § 673(1) (McKinney Supp. 1981-1982). The statute as originally written and construed included motorcyclists as covered persons. *See* Perkins v. Merchants Mut. Ins. Co., 50 A.D.2d 1070, 377 N.Y.S.2d 319 (1975), *aff'd*, 41 N.Y.2d 394, 361 N.E.2d 997, 353 N.Y.S.2d 347 (1977). However, a 1977 amendment had the effect of denying first party benefits to motorcycle occupants, while at the same time holding motorcycle owners liable for payment of first party benefits. N.Y. INS. LAW § 672(1)(6) (McKinney Supp. 1981-1982).

468. N.Y. INS. LAW § 673(2) (McKinney Supp. 1981-1982).

469. *Id.* §§ 671(10), 672(2).

470. *Id.* § 673(2).

an irrational and inequitable result obtains.[471] Where both parties are covered persons, a claimant may sue in tort for economic loss in excess of the $50,000 basic economic loss, and for noneconomic loss (pain and suffering damages) only if the serious injury threshold[472] is met.

B. Montgomery v. Daniels

The major constitutional analysis of the New York no fault law is contained in *Montgomery v. Daniels*.[473] At the time that case arose, the Comprehensive Automobile Insurance Reparations Act afforded approximately the same range of benefits as provided in its current language. Subsequent amendments and the extensive body of administrative law interpreting the statute have not altered the central concepts of first party benefits for basic economic loss, or of restrictions on actions for pain and suffering damages. There is one difference, however, which must be noted. Originally, the New York plan incorporated a $500 monetary as well as a serious injury threshold.[474] The legislature rewrote the "serious injury" definition in 1977, excluding the $500 alternative, effective December 1, 1977.[475]

In *Montgomery*, plaintiffs challenged the tort action thresholds and the act's classifications as violations of the federal and state constitutions. They also contested abridgement of the right to jury trial on state constitutional grounds.

1. Due Process—The plaintiffs' injuries in this case did not qualify as "serious" under the statute and their medical expenses failed to exceed $500; hence, under the existing no fault law, they could not bring a negligence action for damages. They contended that the act deprived them of a right protected by due process.[476]

---

471. The Court of Appeals, however, has found no constitutional infirmity in the statute's classifications of covered and noncovered persons. *See* Montgomery v. Daniels, 38 N.Y.2d 41, 340 N.E.2d 444, 378 N.Y.S.2d 1 (1975).

472. N.Y. INS. LAW § 673(1) (McKinney Supp. 1981-1982). *See supra* note 464.

473. 38 N.Y.2d 41, 340 N.E.2d 444, 378 N.Y.S.2d 1 (1975).

474. The New York plan originally provided that a serious injury was one for which "the reasonable and customary charges for medical, hospital, surgical, nursing, dental, ambulance, X-ray, prescription drug and prosthetic services necessarily performed as a result of the injury would exceed five hundred dollars." Comprehensive Automobile Insurance Repatation Act, ch. 13, 1973 N.Y. Laws 56, 57 (current version at N.Y. INS. LAW § 671(4) (McKinney Supp. 1981-1982).

475. Act of Aug. 11, 1977, 1977 N.Y. Laws, ch. 892, § 6, 7 (codified at N.Y. INS. LAW § 671(4) (McKinney Supp. 1981-1982)).

476. 340 N.E.2d at 450, 378 N.Y.S.2d at 11.

The New York Court of Appeals held that the challenged statute did not "deprive the victim of a right or interest protected by the due process clause of either our State or the Federal Constitution."[477] In reaching this result, the court applied a rationality test, examining the objectives of the legislation and the relationship of the means to those purposes. It readily identified as a permissible objective the exercise of the police power in regulating the use of motor vehicles.[478] Federal and state studies had established the need for reform of the prior fault system of automobile insurance. The act, in the court's view, worked "less deprivation" than the Connecticut guest statute upheld by the United States Supreme Court in *Silver v. Silver*[479] and was aimed at "the promotion of the health, comfort, safety and welfare of society."[480] The *Montgomery* court drew added support for this conclusion from the presumption of constitutionality attached to regulatory and economic legislation, a presumption very deeply imbedded in New York case law.[481]

In considering whether the means employed by the legislature were reasonably related to its objectives, the court found the elimination of tort actions for noneconomic damages in minor injury cases was reasonably related to the no fault law's objectives.[482] The partial elimination of tort actions allowed the guarantee of prompt compensation for economic loss, and the reduction of court congestion under the no fault act.[483] To survive rationality review, the statute did not need to be the best response, or afford a certainly

---

477. *Id.* at 451, 378 N.Y.S.2d at 11.

478. *Id.* at 452, 378 N.Y.S.2d at 12.

479. 280 U.S. 117 (1929). *See* 340 N.E.2d at 452, 378 N.Y.S.2d at 12. In *Silver*, the United States Supreme Court upheld Connecticut's guest statute against a challenge that the legislature could not constitutionally abrogate a common law right of action.

480. 340 N.E.2d at 451, 378 N.Y.S.2d at 11 (quoting Nettleton v. Diamond, 27 N.Y.2d 182, 193, 264 N.E.2d 118, 123, 315 N.Y.S.2d 625, 633 (1970)). In *Nettleton*, the New York Court of Appeals upheld the constitutionality of a law forbidding the sale of the skin of certain species of wild animals, as a valid exercise of the state's police power for the purpose of wildlife conservation. *Accord*, Health Ins. Ass'n v. Harnett, 44 N.Y.2d 302, 310, 376 N.E.2d 1280, 1283-84, 405 N.Y.S.2d 634, 639 (1978) (rejecting a due process challenge to a statute requiring health insurance contracts to include coverage for maternity care).

481. 340 N.E.2d at 452, 378 N.Y.S.2d at 12. *See* Health Ins. Ass'n v. Harnett, 44 N.Y.2d 302, 376 N.E.2d 1280, 405 N.Y.S.2d 634 (1978); People v. Pagnotta, 35 N.Y.2d 333, 337, 253 N.E.2d 202, 205, 305 N.Y.S.2d 484, 488 (1969); *see also* Matter of Van Berkel v. Power, 16 N.Y.2d 37, 40, 209 N.E.2d 539, 541, 261 N.Y.S.2d 876, 878-79 (1955) (expressing the view that the "presumption of validity [is] so strong as to demand of those who attack [legislation] a demonstration of invalidity beyond a reasonable doubt").

482. 340 N.E.2d at 452, 378 N.Y.S.2d at 12-13.

483. *Id.* at 452-53, 378 N.Y.S.2d at 13.

successful solution to the problems perceived by the legislature.[484]

A second aspect of the plaintiffs' due process objection rested on the alleged failure to provide an adequate alternative to the traditional tort action. The *Montgomery* court questioned the applicability of an "adequate substitute" test, but observed that if a substitute remedy were essential to due process, the legislature had in this instance provided a more than adequate alternative.[485] The court also rejected a final due process challenge to the New York act, holding that the term "significant disfigurement" in the act's definition of serious injury was not unconstitutionally vague. Comparable, although not identical, language in New York's workmen's compensation law had been successfully applied without encountering any vagueness problems.[486]

2. Equal Protection—In confronting the equal protection challenge to the New York act, the *Montgomery* court categorically rejected a strict scrutiny standard of review.[487] Plaintiffs had asserted fundamental rights of security of the person and access to the courts. Although those rights might have been catalogued more accurately as due process claims, the court addressed them in its equal protection analysis and found no substantive underpinnings in *Griswold v. Connecticut*[488] or *Eisenstadt v. Baird*[489] that would equate the personal security claim with a fundamental right. Neither did it find any Supreme Court precedent recognizing access to courts as an independent constitutional right.[490]

Having rejected the most demanding standard of review, the New York Court of Appeals considered the applicability of an intermediate test. It found no occasion to apply the "sliding scale" analysis suggested by Justice Marshall in his dissenting opinion in *San Antonio School District v. Rodriquez*.[491] Instead it applied the traditional rationality standard, which it found appropriate to re-

---

484. *Id.* at 453, 378 N.Y.S.2d at 13.

485. *Id.* at 453, 378 N.Y.S.2d at 14. The court referred to article I, section 14 of the New York Constitution and to Munn v. Illinois, 94 U.S. 113, 134 (1876), affirming that the legislature may, within due process limits, alter the common law.

486. 340 N.E.2d at 454, 378 N.Y.S.2d at 15.

487. *Id.* at 455, 378 N.Y.S.2d at 16.

488. 381 U.S. 479 (1965).

489. 405 U.S. 438 (1972).

490. 340 N.E.2d at 455-56, 378 N.Y.S.2d at 17 (citing Ortwein v. Schwab, 410 U.S. 656 (1973); United States v. Kras, 409 U.S. 434 (1973); Boddie v. Connecticut, 401 U.S. 371 (1971)).

491. 340 N.E.2d at 456, 378 N.Y.S.2d at 18 (citing San Antonio School Dist. v. Rodriquez, 411 U.S. 1, 70-135 (1973)). *See supra* note 49 and accompanying text.

view social and economic legislation.[492]

(a) Covered and Noncovered Persons. The plaintiffs challenged the section of the act that barred tort actions between covered persons unless serious injury or loss exceeding basic economic loss was present.[493] That section permitted tort actions, however, where a noncovered person was a party.[494] The court perceived a rational basis for the distinction and for the statutory designations of noncovered persons.[495]

(b) Serious Injury Threshold. The plaintiffs objected to the statute's two-pronged definition of serious injury: (1) death, dismemberment, significant disfigurement, compound or comminuted fracture, or loss of body organ or function, or (2) medical and related charges in excess of $500.[496] The court upheld both thresholds as classifications reasonably related to the objective of eliminating tort actions to recover for minor injuries.[497] Changes in New York's no fault law, subsequent to *Montgomery* would have eliminated some of the objections raised by the plaintiffs to the serious injury classification.[498] For example, the contention that an absolute $500 threshold unequally affected injured persons because medical charges vary substantially in different parts of the state has been mooted by the elimination of the monetary threshold from the statute.[499]

---

492. 340 N.E.2d at 456, 378 N.Y.S.2d at 18 (citing Dandrige v. Williams, 397 U.S. 471, 485 (1970)).

493. N.Y. INS. LAW § 673 (McKinney Supp. 1981-1982).

494. *Id.*

495. 340 N.E.2d at 457-58, 378 N.Y.S.2d at 18-19. The legislature could rationally exclude motorcyclists from the act's tort limitation because the cost of no fault insurance for motorcyclists "would have been prohibitively high." *Id.* at 457, 378 N.Y.S.2d at 18. Uninsured motorists could be excluded because they made no financial contribution to the no fault compensation system. *Id.* at 457, 378 N.Y.S.2d at 19. Nonresidents could be rationally excluded from the act's tort limitations because the legislature had no power to require nonresidents to comply with the act. *Id.* at 457-58, 378 N.Y.S.2d at 19.

496. Comprehensive Automobile Insurance Reparation Act, ch. 13, 1973 N.Y. Laws 56, 57 (current version at N.Y. INS. LAW § 671(4) (McKinney Supp. 1981-1982)).

497. 340 N.E.2d at 459, 378 N.Y.S.2d at 210.

498. The current provisions of the act do not contain the monetary threshold, and include "fracture" without limitation as to kind. *See* N.Y. INS. LAW § 671(4) (McKinney Supp. 1981-1982).

499. Interestingly, geographic variations in charges for the same types of medical service have been acknowledged and must be dealt with in determining the reasonableness of medical expenses to be covered by no fault benefits. To achieve some standardization of medical fees, the Superintendent of Insurance has promulgated regulations establishing no fault fee schedules for medical services, incorporating a "regionalization factor" in recognition of the differing costs of medical practice in various parts of the state. N.Y. ADMIN. CODE tit. 11, § 68.83 (1981). The regional factor was first incorporated in the fee schedule by the

3. Right to Jury Trial—Article I, section 2, of the New York Constitution provides that: "Trial by jury in all cases in which it has heretofore been guaranteed by constitutional provision shall remain inviolate forever."[500] The *Montgomery* court perceived no violation of this right under the provisions of the no fault law.[501] The court reasoned that the legislature had partially eliminated tort actions and where the right of action no longer existed, "there remained nothing to which the right to trial by jury may attach."[502]

C. *Recent Cases*

No fault litigation following *Montgomery* provides no indication that the New York Court of Appeals would retreat from its firm endorsement of the constitutionality of the state's no fault law. Subsequent challenges to the law have not come from injured parties whose actions for minor pain and suffering damages had been barred by the law's serious injury threshold. For example, in *Rachlin v. Lewis*,[503] attorneys successfully objected to the Superintendent of Insurance's schedule of attorney's fees established under the act.[504] The *Rachlin* court held that the Insurance Department had no authority under Insurance Law[505] to bar an attorney from receiving fees from his client in excess of what the insurance carrier paid, or to place a maximum on disbursements an arbitrator may award.[506] The wording of the regulation has been amended to conform to the court's ruling.[507] In *Country-Wide Insurance Co. v. Harnett*,[508] a three-judge federal district court rejected the insurance company's contentions that the provisions for binding arbitration at the demand of the claimant,[509] and for automatic renewal of most automobile insurance policies,[510] violated equal protection, due process and the contract clause of the federal

---

third amendment to the Regulation, Sept. 29, 1980.

500. N.Y. CONST. art. I, § 2.
501. 340 N.E.2d at 459, 378 N.Y.S.2d at 21.
502. *Id.* at 460, 378 N.Y.S.2d at 23.
503. 96 Misc.2d 701, 409 N.Y.S.2d 594 (N.Y. Sup. Ct. 1978).
504. N.Y. ADMIN. CODE tit. 11, § 65.16(c)(8) (1981)
505. N.Y. INS. LAW § 675(1) (McKinney Supp. 1981-1982).
506. 409 N.Y.S.2d at 597.
507. N.Y. ADMIN. CODE tit. 11, § 65.16(c)(7)(vii)(ix) (1981).
508. 426 F. Supp. 1030 (S.D.N.Y. 1977), *aff'd*, 431 U.S. 934 (1978).
509. *N.Y. Ins. Law* § 675(2) (McKinney Supp. 1981-1982).
510. *Id.* § 167-a(4), (6).

Constitution.[511]

Finally, the statute's authorization of the Superintendent of Insurance to establish permissible charges for health services furnished to no fault patients[512] was contested by a surgeon and a severely injured accident victim in *Goldberg v. Carey*.[513] The plaintiffs pleaded multiple causes of action including deprivation of property without due process, denial of equal protection, and impairment of the right to contract. Although the federal court abstained, awaiting state court determination of state issues, no state rulings have been forthcoming. The state courts have not ruled on these claims yet. Nevertheless, the Insurance Department's power to promulgate fee schedules appears to be constitutionally secure.[514]

New York's Comprehensive Automobile Insurance Reparations Act was amended in 1977 to direct the Superintendent to consult with the Chairman of the Workers' Compensation Board and coordinate their regulations respecting health provider service charges.[515] Fee schedules are to be established for all "necessary expenses" as defined in the act, including hospital, ambulance, and prescription drug charges, fees for psychiatric and occupational therapy and rehabilitation, as well as medical, surgical, dental and nursing care.[516] Since January 1, 1978, no fault medical expenses have been subject to the Workers' Compensation Board schedules, and the schedules have been expanded to cover an increasing number of services.[517]

Possible challenges of specific fees might be raised on the basis of statistical evidence, but it is doubtful that the entire scheme could be invalidated in view of its acceptance in the workers' compensation field. There appears to be strong support by the insurance industry for fixed fees in the treatment both of employment and automobile accident injuries.

---

511. 426 F. Supp. 1030. The court dismissed the complaint. *Id.* at 1035.

512. N.Y. INS. LAW § 678 (McKinney Supp. 1981-1982).

513. 462 F. Supp. 872 (E.D.N.Y. 1978), *aff'd*, 601 F.2d 653 (2d Cir. 1979). Although *Goldberg* was filed in federal district court, that court abstained from exercising its jurisdiction until a state court ruled on the important issues of state constitutional law and statutory construction presented by the case.

514. Under section 678, the fee schedule promulgated by the Superintendent of Insurance may not exceed the fees established by the Workers' Compensation Board. At present, the two schedules are identical. N.Y. INS. LAW § 678 (McKinney Supp. 1981-1982).

515. Act of Aug. 11, 1977, N.Y. Laws ch. 892, § 15 (codified at N.Y. INS. LAW § 671(1)(a) (McKinney Supp. 1981-1982).

516. N.Y. INS. LAW § 678(2) (McKinney Supp. 1981-1982).

517. *See* N.Y. ADMIN. CODE tit. 11, app. 17-C (1981).

XIV. MICHIGAN

A. *Introduction*

1. *The No Fault Statute*—Michigan enacted the Motor Vehicle Personal and Property Protection Act in 1972.[518] It requires security for payment of first party benefits,[519] and prescribes the usual PIP coverage[520] and survivor's benefits.[521] Tort liability for noneconomic loss remains only if the injured person suffers "death, serious impairment of body function, or permanent serious disfigurement."[522] The somewhat unusual property provisions are not crucial to the no fault scheme and are discussed elsewhere.[523]

2. *Overview*—Michigan's no fault law, like Illinois', was challenged prior to its effective date, October 1, 1973, in the unusual form of a request by the governor and state senate for an advisory opinion.[524] The Michigan constitution permits the highest court of the state to respond to a legislative or executive request for judicial guidance on important questions of law.[525]

The Michigan Supreme Court rendered its advisory opinion

---

518. Ch. 13, 1972 Mich. Pub. Acts 902 (codified at MICH. COMP. LAWS §§ 500.3101 to .3179 (1983).

519. MICH. COMP. LAWS § 500.3101 (1983).

520. The Act provides for allowable expenses for all reasonable charges for medical, hospital, rehabilitation services. *Id.* § 500.3107. No monetary limit is attached to these allowable expenses except for a $1000 maximum for funeral and burial expenses. *Id.* Work loss for three years from the date of the accident is reduced by 15%. *Id.* The original maximum of $1000 for a 30-day period has been annually adjusted to reflect increases in the cost of living; the maximum work loss payment effective from Oct. 1, 1981 through Sept. 30, 1982 is $2049.00 per month. Michigan Dep't. of Licensing and Regulation, Insurance Bureau, Bulletin 81-17, Aug. 10, 1981. Work loss also includes payments for substitute services replacing services the accident victim would have performed. MICH. COMP. LAWS § 500.3107 (1983). These services may be compensated at a rate not exceeding $20 a day for three years. *Id.* If the injured person is temporarily unemployed at the time of the accident, work loss is based on income earned in the last month of full time employment. *Id.* § 500.3107(a).

521. Survivor's loss benefits paralleling work loss benefits are payable to dependents if the accident victim dies within three years of the date of the accident. *Id.* § 500.3108. They are subject to the same maximum adjustment for cost of living increases as work loss benefits. *Id.* Recovery is allowed for damages for allowable expenses, work loss and survivor's loss, in excess of the limitations contained in sections 3107 to 3110. *Id.* Section 3135(2)(c) provides: "The party liable for damages is entitled to an exemption reducing his [or her] liability by the amount of taxes that would have been payable on account of income the injured person would have received if he [or she] had not been injured." *Id.*

522. *Id.* § 500.3135(1) (1983).

523. *See infra* notes 651 and 683 and accompanying text.

524. *In re* Requests of the Governor and the Senate on the Constitutionality of Act No. 294 of the Public Acts of 1972, 389 Mich. 441, 208 N.W.2d 469 (1973).

525. MICH. CONST. art. III, § 8.

on June 18, 1973, emphasizing that "an advisory opinion does not constitute a decision of the Court and is not precedentially binding in the same sense as a decision of the court after a hearing on the merits."[526] The court then advised the governor and legislature that the new statute did not violate the Michigan state constitution or state rules applicable to statutory construction. The court declined, however, to render an advisory opinion concerning the equal protection issues raised by the governor and senate.[527]

A closer look at the advisory opinion issued by the Michigan Supreme Court in its first encounter with the no fault law discloses that the court concentrated on statutory construction rather than constitutionality. Nevertheless, the opinion illuminates the court's perception of its function. The authority to render an advisory opinion, which appears to endow the court with "veto" power over legislation at the "request" of one of the political branches was cautiously exercised. The court gave due deference to the legislature, sought its intent and accepted its language.

The advisory opinion addressed three issues: first, whether the act embraced more than one object in violation of a Michigan constitutional limitation that: "No law shall embrace more than one object, which shall be expressed in its title;"[528] second, whether the act[529] modified, amended or amended by reference any other Michigan statutory provisions with respect to the substantive law of torts in violation of a Michigan constitutional limitation that: "No law shall be revised, altered or amended by reference to its title only. The section or sections of the Act altered or amended shall be reenacted and published at length";[530] finally, whether the phrases "serious impairment of body function" and "permanent serious disfigurement" as used in the act were "sufficient for legal interpretation."[531]

The majority of the court answered the first two questions in the negative.[532] The third question focused on the act's serious injury threshold: "A person remains subject to tort liability for noneconomic loss caused by his or her ownership, maintenance, or

526. 208 N.W.2d at 471 n.1.

527. *Id.* at 482 (Levin, J., concurring).

528. MICH. CONST. art. 4, § 24. *See* 208 N.W.2d at 472.

529. *See* MICH. COMP. LAWS § 500.3135 (1983).

530. MICH. CONST. art. 4, § 25. *See* 208 N.W.2d at 472.

531. *See* MICH. COMP. LAWS § 500.3135(1) (1983).

532. *Id.* For a dissenting view on the second question, see 208 N.W.2d at 494 (Williams, J., dissenting).

use of a motor vehicle only if the injured person has suffered death, serious impairment of body function, or permanent serious disfigurement."[533] The Michigan Supreme Court upheld the adequacy of these criteria for residual tort liability on the basis of general rules of statutory construction, and historically accepted competence of the trier of fact to interpret comparable inexact terminology. In matters of construction a court must honor the intent of the legislature and "read the language in the light of the general purpose sought to be accomplished."[534] The court referred to Michigan law directing that "[a]ll words and phrases shall be construed and understood according to the common and approved usage of the language."[535] The word "serious" had been construed exhaustively over many years.[536] The court concluded that a jury could as capably interpret "serious impairment of body functions" and "serious disfigurement" as it could the traditional jury instruction language of proximate cause, reasonable care, and reasonable, fair and adequate damages.[537]

The three questions to which the court responded in its advisory opinion were not the sole issues reflecting on the validity of the no fault law. Indeed, the most significant question posed was whether the substantial curtailment of tort liability expressed in the statute[538] violated federal and state guarantees of due process and equal protection. Explaining the court's silence on these points, Justice Levin, in his concurring opinion, indicated that the Michigan Constitution permitted the court to pronounce advisory opinions only on the constitutionality of important questions of law raised with specificity by a co-equal branch of state government.[539] This limitation precluded consideration of questions of fact, and of judicially conjured factual contexts. Justice Levin stated: "It is not properly our function to hypothesize particularized claims or to set up, speculatively, strawmen classes of persons who might claim to be disadvantaged in various ways by the classifications and provisions of the act."[540] Neither the governor nor the legislature had particularized any claims of unconstitutionality. Al-

533. MICH. COM]. LAWS § 500.3135(1) (1983).
534. 208 N.W.2d at 480.
535. *Id.* (citing MICH. COMP. LAWS § 8.3a (1981).
536. *Id.* at 480 n.10.
537. *Id.* at 481-82.
538. MICH. COMP. LAWS § 500-3135 (1983).
539. 208 N.W.2d at 482 (Levin, J., concurring).
540. *Id.* at 483-84.

though bar associations had submitted specific issues on behalf of identifiable classes of persons, the court lacked authority, according to Justice Levin, to address issues not raised by a co-equal branch of state government.[541]

In summary, the advisory opinion issued before the effective date of the Michigan no fault law answered technical questions relating to form and enactment procedure, and what might be termed a facial "void for vagueness" attack on the serious injury phraseology of the statute. The Michigan Supreme Court declined to commit itself through the medium of an advisory opinion on the broader issues of whether the no fault system enacted by the legislature complied with due process and equal protection.

B. Shavers v. Kelley—Trial Court (1974)

The major constitutional challenge of the Michigan no fault law commenced as an action for a declaratory judgment in the Wayne County Circuit Court. That litigation turned into the landmark case of *Shavers v. Kelley*.[542] The plaintiffs were owners of motor vehicles, some having purchased no fault insurance, and the named defendants were three state officials and several insurance companies. The plaintiffs appear to have raised almost every conceivable constitutional objection to the no fault statute. The multiplicity of issues, which the trial court entertained, arose in part from its expansive view of standing—a view not shared by the appellate courts.[543]

1. Statutory Provisions Held Constitutional by the Trial Court—The opinion of Judge Gilmore who tried the *Shavers* case is not reported but it is summarized in the Michigan Supreme Court and Court of Appeals decisions.[544] Among others, the following provisions of the Michigan act were held constitutional by the trial court:[545] (1) requiring purchase of no fault insurance as a condition precedent to registration or operation of a motor vehicle;[546]

---

541. *Id.* at 484. The Michigan Trial Lawyers Association had "propounded 17 questions . . . on behalf of identifiable classes of persons" including motorcyclists, unemployed persons and others. *Id.*

542. 402 Mich. 554, 267 N.W.2d 72 (1978), *aff'g in part, rev'g in part*, Shavers v. Attorney Gen., 65 Mich. App. 355, 237 N.W.2d 325 (1975), *cert. denied*, 442 U.S. 934 (1979). The trial court opinion is not reported.

543. *See infra* notes 560, 569-571, 600-605 and accompanying text.

544. *See* Shavers v. Kelley, 402 Mich. 544, 267 N.W.2d 72, 79-80 (1978); Shavers g. Att. Gen., 65 Mich. App. 355, 237 N.W.2d 325, 329-30 (1975).

545. 267 N.W.2d at 79, 237 N.W.2d at 329.

546. MICH. COMP. LAWS § 500-3101(1) (1983).

(2) imposing penalties for failure to obtain insurance;[547] (3) the personal injury protection scheme; (4) limiting funeral and burial expenses to $1,000;[548] (5) requiring non-resident motorists, when in Michigan more than an aggregate of 30 days in a calendar year, to obtain no fault insurance;[549] and (6) empowering insurance companies to seek reimbursement for no fault benefits from an insured's tort recovery.[550] The trial court construed the act as requiring that only parallel benefits recovered in tort be offset against personal injury protection benefits.[551]

2. *Statutory Provisions Held Unconstitutional by the Trial Court*—The trial court determined that the following sections of the law violated the equal protection clauses of the federal and Michigan constitutions and that these sections were severable:[552] (1) provisions excluding vehicles having two wheels or less,[553] (2) the difference in the payment procedure applicable to replacement services, as opposed to other no fault benefits,[554] (3) the subtraction of governmental statutory benefits from personal injury benefits[555] and (4) *property* protection insurance mandated by the act.[556] The trial court also decided that the authority granted to the Commissioner of Insurance to permit inclusion in insurance contracts of deductibles exceeding $300 per accident[557] violated the Michigan constitution[558] by delegating legislative power without prescribing standards for its exercise. Similarly, provisions of the act stating that nonresident owners and occupants of a motor vehicle not registered in Michigan could not sue in tort under Michigan law, unless those persons' insurance entitled them to personal injury protection benefits, were also declared unconstitutional.[559]

3. *Standing*—In response to defendants' contention that

---

547. *Id.* § 500.3102(2).
548. *Id.* § 500.3107(1).
549. *Id.* § 500.3102(1).
550. *Id.* § 500.3116.
551. 267 N.W.2d at 79, 237 N.W.2d at 329.
552. 267 N.W.2d at 79-80, 237 N.W.2d at 329-30.
553. MICH. COMP. LAWS § 500.3101(2) (1983).
554. *Id.* § 500.3107(b). The act specified prior payment by the insured and subsequent reimbursement for replacement services. *Id.*
555. *Id.* § 500.3109(1).
556. *Id.* §§ 500.3101(1), (3), .3121, .3123, .3127, .3135, .3145(2), .3148, .3163(1), (3). *See* 237 N.W.2d at 333. Because the act could not abolish tort liability for property damage, its required residual liability insurance coverage would have to be read to include liability for property damage. MICH. COMP. LAWS § 500.3131 (1983).
557. MICH. COMP. LAWS § 500.3109(3) (1983).
558. MICH. CONST. art. III, § 2.
559. MICH. COMP. LAWS § 500.3113(c) (1983).

plaintiffs proffered some issues not properly before the court, the trial court invoked the taxpayer suit—real party in interest rule:

> [A]n action to prevent the illegal expenditure of state funds or to test the constitutionality of a statute relating thereto may be brought in the name of a domestic non-profit corporation organized for civic, protective, or improvement purposes, or in the names of at least 5 residents of this state who own property assessed for direct taxation by the county wherein they reside.[560]

In the trial court's view, this court rule supplied a foundation for plaintiffs to attack all possible flaws in the statute.

C. Shavers v. Kelley—Court of Appeals (1975)

1. *Trial Court Rulings*—On appeal, the Michigan Court of Appeals reversed the trial court on the issue of standing by refusing to apply the "taxpayer suit rule" to the plaintiffs.[561] It also invalidated the trial court's ruling on certain reimbursement provisions of the act, and the exclusion of motorcycles.[562] It affirmed the lower court's decision as to the constitutionality of the act's personal injury protection scheme, and the unconstitutionality of the act's property protection plan.[563] Beyond these individual determinations, the court of appeal also considered the question of the appropriate standard of constitutional review; it decided that a rational basis test incorporated the proper standard.[564]

In overturning the trial court's use of the taxpayer suit—real party in interest rule to authorize standing, the court of appeals stated:

> The court rule allows taxpayers aggrieved by the outlay of state funds to hurdle the traditional standing obstacle in taxpayers suits. We do not read it as permitting a group to challenge any legislation merely because of an incidental expenditure of state funds; almost all legislation involves some public spending.[565]

Instead, the court of appeals looked to the General Court Rule pertaining to declaratory judgments to determine plaintiff's stand-

---

560. *See* 267 N.W.2d at 81; 237 N.W.2d at 330. *See also* MICH. COMP. LAWS § 600.2041(3) (1981) (identical provision).

561. Shavers v. Attorney Gen., 65 Mich. App. 355, 362, 237 N.W.2d 325, 330 (1975).

562. 237 N.W.2d at 341.

563. *Id.* at 332-35.

564. *Id.* at 333-35.

565. *Id.* at 330 (rebutting the trial court's application of MICH. COMP. LAWS § 600.2041(3) (1981)).

ing.[566]. Under that rule, the court found that the presence of "a case of actual controversy"[567] was necessary to establish standing. Because hypothetical issues were insufficient, the court of appeals held that a party must establish that a statute is applicable to him before he may challenge its constitutionality by seeking a declaratory judgment.[568]

Applying its interpretation of standing and the requirements for declaratory judgment to limit the justiciable issues, the court of appeals also reversed the trial court's rulings pertaining to replacement services reimbursement, authority of the Insurance Commissioner to approve certain deductibles, subtraction of government benefits, tort action rights of nonresidents and reimbursement of an insurer for no fault benefits from an insured's tort recovery.[569] The plaintiffs had not demonstrated a need for determination of their rights in these matters.[570] Finally, the court of appeals also reversed the trial court's ruling that the exclusion of motorcycles violated equal protection.[571]

2. *Trial Court Rulings Affirmed*—The Michigan Court of Appeals affirmed the trial court's general holding that the no fault statute was constitutional.[572] Because the court of appeals reversed most of the trial court's specific holdings of unconstitutionality on the procedural ground of lack of standing, the affirmation of these statutory provisions would appear to be by implication rather than specific supportive reasoning. Nonetheless, in its explication of the proper standard for review, the court of appeals clearly upheld the personal injury protection plan.[573] The appellate court agreed both procedurally and substantively with the lower court that the property protection scheme must be struck from the statute. That determination rested on equal protection and due process grounds.[574]

3. *Standard of Constitutional Review*—Seriatim review of the trial court's rulings on individual sections of the Michigan no fault act did not represent the sole function of the court of appeals'

---

566. MICH. STAT. ANN. G.C.R. 521 (1963) (Callaghan 1976).

567. 237 N.W.2d at 330. *See* Kuhn v. City of East Detroit, 50 Mich. App. 502, 213 N.W.2d 599, 601 (1973); Welfare Employees Union v. Civil Service Comm., 28 Mich. App. 343, 184 N.W.2d 247, 250 (1970).

568. 237 N.W.2d at 331.

569. *Id.* at 331-32.

570. *Id.* at 331.

571. *Id.* at 332-33. *See* MICH. COMP. LAWS § 500.3101(2) (1983).

572. 237 N.W.2d at 335.

573. *Id.* at 332.

574. *Id.* at 333-35.

opinion. It also addressed the issue of the proper standard of constitutional review. The plaintiffs had alleged that the trial court erred in applying a rationality rather than a strict scrutiny test. The justification for strict scrutiny, according to plaintiffs, rested on the no fault law's purported infringement of a fundamental right—the right to travel.[575] The argument did not convince the court:

> We are unwilling to extend the right to travel to include mere locomotion and thus severely limit legislative ability to respond to the many hazards present-day automobile operation presents. *Rather than "strict scrutiny", the more permissive traditional tests of due process and equal protection should apply to this legislation.*[576]

The rationality test adopted by the Michigan Court of Appeals permits deferential regard for legislative choices in defining classifications, and a consideration of administrative convenience and economic impact of a challenged governmental regulation or program.[577]

Applying the rationality test, the court of appeals identified the general legislative goal of the Michigan no fault law: "[T]o provide timely and adequate relief, at the lowest cost to the system and the individual, for economic losses arising out of motor vehicle mishaps."[578] The means embodied in the personal injury protection provisions of the act bore a rational relationship to this permissible legislative goal.[579] The court of appeals therefore held that the act's replacement of tort liability with compulsory first-party insurance did not violate due process.[580]

Similarly, the court found no violation of equal protection in treating automobile accident injuries differently from other injuries.[581] It perceived no invidious discrimination in devising a benefit and liability system applicable only to motor vehicle accident

---

575. *Id.* at 331. *See, e.g.*, Dunn v. Blumstein, 405 U.S. 330 (1972) (striking down state residency requirements for voting as an unconstitutional restriction on the right to travel).

576. 237 N.W.2d at 331-32 (emphasis added).

577. Recent decisions of the United States Supreme Court have reflected a disinclination to override legislative judgment in social insurance schemes at the federal and state level. *See, e.g.*, Weinberger v. Salfi, 422 U.S. 749, 774 (1975); Jefferson v. Hackney, 406 U.S. 535, 546-51 (1972), *reh'g denied*, 408 U.S. 848 (1972); Dandridge v. Williams, 397 U.S. 471, 484 (1970).

578. 237 N.W.2d at 334.

579. *Id.* at 332.

580. *Id.*

581. *Id.* at 333.

victims and tortfeasors.[582] On the other hand, the court reversed the trial court's decision that the exclusion of motorcycles from the class of motor vehicles subject to the act violated equal protection.[583] The legislature's decision that inclusion of motorcycles would adversely affect the cost of coverage reflected an assessment of economic impact, a legitimate consideration in designing a benefits program.[584]

The appeals court found, however, that the property protection plan failed even a minimum scrutiny test.[585] As enacted, the scheme applied only to third parties, providing coverage up to $1,000,000 for damage to properly parked cars and their contents.[586] The court could discern no rational basis for excluding moving vehicles while covering other types of property.[587] The owner of a vehicle damaged while in operation would not be compensated by no fault coverage and would have to purchase collision insurance for protection. This plan appeared inconsistent with the act's general purpose of providing compensation for economic loss resulting from motor vehicle accidents. The property protection provisions could not be reconciled with the requirements of equal protection,[588] and concommitantly, fell short of meeting due process criteria; the means appeared arbitrary and without a reasonable relationship to the general purpose of the act.[589] The court of appeals held the invalid property protection provisions severable. Thus, the result of the opinion was to affirm solely Michigan's personal injury no fault scheme.[590]

In summary, the court of appeals agreed generally with the trial court that the personal injury no fault plan was valid and the property damage plan invalid. It adopted a narrower view of standing to challenge the no fault law's provisions, and focused on clari-

---

582. *Id.*

583. *Id.* at 332-33.

584. *Id.* at 320. *See* Geduldig v. Aiello, 417 U.S. 484, 495-96 (1974) (upholding the exclusion of pregnancy related disabilities from the coverage provided in California disability insurance); Gauthier v. Campbell, Wyant & Cannon Foundry Co., 360 Mich. 510, 104 N.W.2d 182, 188 (1960) (upholding limitations on workmens' compensation benefits for death from silicosis, giving weighty consideration to the economic effect on the industry of unlimited benefits).

585. 237 N.W.2d at 333.

586. MICH. COMP. LAWS §§ 500.3121, .3123(1)(a) (1983).

587. 237 N.W.2d at 333-34.

588. *Id.*

589. *Id.*

590. *See id.* at 335-36 (Burns, J., concurring) (expressing uncertainty relative to the wisdom and practical effects of the no fault law).

fying the proper standard of constitutional review. Neither the court of appeals nor the trial court, however, touched the due process issues that would deeply engage the Supreme Court of Michigan on appeal.

D. Shavers v. Kelley—Supreme Court of Michigan (1978)

The Supreme Court of Michigan agreed with the court of appeals on the issues of standing, the authority of the legislature to enact a no fault protective scheme and, generally, the standard of review to be applied by the judicial branch.[591] It diverged sharply from the intermediate court, however, in its perception of the new system's fulfillment of due process requirements. The Michigan Supreme Court held the no fault statute "constitutional in its general thrust but unconstitutionally deficient in its mechanisms for assuring that compulsory no-fault insurance is available to Michigan motorists at fair and equitable rates."[592] The "general thrust" of constitutionality supported both the personal injury and property damage schemes[593] but, given existing ratemaking mechanisms and availability of insurance, the legislature could not mandate no fault insurance as a condition precedent to registration and operation of a motor vehicle in Michigan.

The Michigan Supreme Court identified three deficiencies in the no fault plan. First, the statutory protection against "excessive, inadequate or unfairly discriminatory" rates was unsupported by clarifying rules from the Commissioner of Insurance, without legislatively sufficient definition, and without any history of prior court interpretation.[594] Second, there were inadequate statutory provisions for a motorist attacking the validity of an individual rating decision.[595] Finally, the act failed to afford individuals adequate means or opportunity to challenge insurance refusal, discriminatory cancellation, or assignment to the "Automobile Placement Facility" with its presumptively higher rates.[596] Mindful of the fact that its opinion was being issued five years after the effective date of the act, the court sustained the law for eighteen months to permit the legislature and Commissioner of Insurance to remedy the

591. *See* Shavers v. Kelley, 402 Mich. 554, 267 N.W.2d 72 (1978), *cert. denied*, 442 U.S. 934 (1979).

592. 267 N.W.2d at 78.

593. *Id.*

594. *Id.* at 78.

595. *Id.*

596. *Id.*

constitutional deficiency.[597] The court suggested forms of corrective action[598] and retained jurisdiction to assess the constitutional status of the act at the conclusion of the eighteen month grace period.[599]

1. Standing—The Michigan Supreme Court agreed with the court of appeals that plaintiffs' standing must be predicated on the state's declaratory judgment rule,[600] and not on the basis of a taxpayer suit. To hold otherwise, the court stated, "would virtually abolish the law of standing, a result not clearly contemplated" by the taxpayer suit and declaratory judgment rules.[601] Before a declaratory judgment could be granted, it was essential that the plaintiff plead and prove facts demonstrating adverse interests and a justiciable controversy.[602] Under this test, the court judged that both the plaintiffs, who owned automobiles and motorcycles, and the cross-plaintiff insurance companies required to issue insurance contracts that complied with the no fault act, possessed standing to raise the following issues: (1) the constitutionality of *compulsory* personal injury, property damage and residual liability insurance; (2) whether personal injury and property damage no fault insurance violated due process and equal protection guarantees;[603] (3) whether the exclusion of two-wheel vehicles violated due process and equal protection clauses; (4) whether the work loss and replacement services reimbursement provisions violated equal protection clauses; and (5) whether the provisions relating to non-resident motorists violated due process and equal protection clauses.[604] The court stated that the plaintiffs had not established standing to challenge the subtraction of tort claim recoveries or governmental

---

597. The Michigan Supreme Court continued the act's effectiveness for eighteen months from June 8, 1978. *Id.*

598. *Id.* at 91.

599. *Id.* at 78.

600. *Id.* at 81-82. The declaratory judgment rule provides that "[i]n a case of actual controversy within its jurisdiction, any circuit court of this state may declare the rights and other legal relations of any interested party seeking a declaratory judgment, whether or not other relief is or could be sought or granted." MICH. STAT. ANN. GCR 521.1 (Callaghan 1982). For explication of the rule, the court cited: Merkel v. Long, 368 Mich. 1, 117 N.W.2d 130 (1962); Updegraff v. Attorney Gen., 298 Mich. 48, 298 N.W. 400 (1941); City of Flint v. Consumers Power Co., 290 Mich. 305, 287 N.W. 475 (1939); Welfare Employees Union v. Civil Service Comm'n, 28 Mich. App. 343, 184 N.W.2d 247 (1970).

601. 267 N.W.2d at 81 n.4. The court also noted that the Michigan taxpayer's suit rule included the phrase "to test the constitutionality of a statute relating thereto" which was not present in the Illinois and California taxpayer's suit statutes. *Id.*

602. *Id.* at 82.

603. *See* U.S. CONST. amend. XIV; MICH. CONST. art. I, § 2.

604. 267 N.W.2d at 83-84.

benefits from personal injury benefits, or the delegation of power to the Commissioner of Insurance to approve deductibles.[605]

2. Constitutionality of the Michigan Act's Regulatory Scheme—The most important questions identified by the Michigan Supreme Court were two due process "sub-issues" not addressed by the courts below, but that the plaintiffs had repeatedly raised in their briefs: whether it was constitutional to require no fault insurance as a condition precedent to registration and operation of a motor vehicle, and whether the Michigan regulatory scheme complied with due process. The court analyzed the constitutional questions in two phases: first, whether the compulsory character of no fault insurance created "an entitlement to fairness and availability of such insurance;"[606] and second, if that entitlement existed, whether the no fault act afforded due process consistent with the motorist's interest.[607]

(a) The Entitlement Doctrine. In evaluating the constitutionality of the Michigan no fault law's compulsory insurance provisions, the *Shavers* court employed a novel analysis. It invoked a version of the entitlement doctrine.[608] That doctrine addresses the right of recipients of governmental benefits to expectation that the benefits will continue. An entitlement issue may arise when the government has established a benefit to which the plaintiff has a

---

605. *Id.* at 83-84 n.13. Two of these issues were before the court in other cases as the time *Shavers* arose. *See* O'Donnell v. State Farm Mutual Auto. Ins. Co., 404 Mich. 524, 273 N.W.2d 829, *appeal dismissed*, 444 U.S. 803 (1979) (no fault benefits could be reduced by the amount of federal Social Security benefits payable to the beneficiary without violating equal protection clause); Workman v. Detroit Auto. Inter-Ins. Exch., 404 Mich. 477, 274 N.W.2d 373 (1979) (tort recovery of injured party could be set off by insurance carrier against PIP benefits paid to injured party only if the tort recovery was for losses already paid for by insurance carrier).

606. 267 N.W.2d at 85, n.14.

607. *Id.*

608. *See id.* at 86-87. For explanation of the doctrine, see Casenote, *Entitlement Doctrine—Michigan Compulsory No Fault Automobile Insurance Law Violates Due Process*, 1979 B.Y.U. L. REV. 433 (examining the doctrine with particular application to *Shavers*); Comment, *Entitlement, Enjoyment, and Due Process of Law*, 1974 DUKE L.J. 89. For application of the doctrine, see Board of Regents v. Roth, 408 U.S. 564 (1972). In *Roth*, a state university professor on a one year appointment claimed he could not be denied renewal without a hearing and a statement of the reasons for non-renewal; the Supreme Court held that no property or liberty due process interests were involved. *Id.* at 578-79. *See also* Bell v. Burson, 402 U.S. 535 (1971) (important interests attach to possession of a driver's license; the licensee must be given the opportunity for a hearing before he is deprived of his license); Pennsylvania Coal Mining Ass'n v. Insurance Dep't, 471 Pa. 437, 370 A.2d 685 (1977) (coal mining companies were denied due process when rates for their compulsory black lung disease insurance were increased without notice and hearing.)

legitimate claim he relies on and presently enjoys.[609]. If all those elements are present, a court may elect to treat the entitlement as "property" meriting due process protection. In *Shavers*, the state action was the legislation mandating no fault insurance; the benefit could be characterized as a fair and accountable insurance system.[610] Applying an entitlement analysis, the court concluded that Michigan motorists possessed a legitimate claim, or a protected property interest, in receiving this benefit because they could not register or operate a car without no fault insurance.[611] Noting that "the independent mobility provided by an automobile is a crucial, practical necessity,"[612] the court observed: "In choosing to make no-fault insurance compulsory for all motorists, the Legislature has made the registration and operation of a motor vehicle inexorably dependent on whether no-fault insurance is available at fair and equitable rates. Consequently due process protections under the Michigan and United States Constitutions . . . are operative."[613]

The *Shavers* court further found that reliance on the benefit had been fostered by the legislative prohibition of excessive, inadequate or unfairly discriminatory rates,[614] and by the no fault statute's "guarantee" that no fault coverage would be available to any person.[615] The court did not allude to present enjoyment of the benefit. Indeed, application of that requirement appears inappropriate when legislation has imposed compulsory conditions on an activity, which for many is a necessity. The crucial consequence of the invocation of the entitlement doctrine emerged in the court's holding that when the state compels individuals to purchase insurance in order to operate an automobile without incurring penalites, it confers an entitlement of fair rates that enjoys the protection of constitutional due process.[616]

(b) Existing Ratemaking Practices. The *Shavers* court described the existing practices in the state for filing and approval of rates, and found that they violated due process.[617] The general requirement in Michigan, as in other states, was that rates shall not

609. *See* Casenote, *supra* note 608, at 440.
610. 267 N.W.2d at 86-87.
611. *Id.* at 87.
612. *Id.*
613. *Id.*
614. *Id.* (citing MICH. COMP. LAWS § 500.2403(1)(d) (1983)).
615. 267 N.W.2d at 87 (citing MICH. COMP. LAWS § 500.3301(1)(a) (1983)).
616. For criticism of the majority's application of the entitlement doctrine, see 267 N.W.2d at 111-13 (Ryan, J., concurring).
617. *Id.* at 88-90.

be "excessive, inadequate or unfairly discriminatory".[618] This vague guideline had not been particularized or made effective by administrative regulations:

> The legislative due process mandate is thus reduced to mere exhortation. When we add that the statute authorizes insurers to utilize any classification scheme which "may measure any differences among risks that may have a probable effect on losses or expenses" it becomes clear that rates can be established on insubstantial bases which do not satisfy due process. Absent administrative rules or legislative definition giving substance to the statutory language, there are inadequate safeguards against arbitrary action or invidious discrimination.[619]

In addition to the suspect rate structure, the court faulted the system for failing to provide the motorist with practical and effective means of challenging rates, or of pursuing a claim of discriminatory cancellation or refusal of insurance.[620] The motorist who was unable to procure insurance through ordinary methods was guaranteed coverage through the Automobile Placement Facility, an assigned risk plan.[621] But here too, the motorist could not demand justification of the rate charged, and was afforded no statutory procedure for initially challenging his assignment to the facility along with its higher rates and more limited coverage options.[622]

Because these deficiencies in the ratemaking process resulted in a denial of due process, the court held the compulsory insurance provisions of the Michigan no fault act unconstitutional. The court's decision did not strike down those provisions immediately, however, but allowed the legislature and Commissioner of Insurance eighteen months to remedy the identified defects.

(c) Required Revisions. The Supreme Court of Michigan instructed the legislature and Commissioner of Insurance that certain revisions of the law were necessary to render it constitutional. First, it required that "substantial meaning" be infused into the statutory standard that rates not be excessive, inadequate or unfairly discriminatory.[623] Accordingly, the court indicated that a rate schedule must provide reasonable premiums without regard to

---

618. *See* MICH. COMP. LAWS §§ 500.2403(1)(d), .2027 (1983).
619. 267 N.W.2d at 88 (emphasis added).
620. *Id.* at 88-89.
621. MICH. COMP. LAWS § 500.3301(1)(a) (1983).
622. 267 N.W.2d at 89-90.
623. *Id.* at 91.

factors asserted to warrant differences in premiums among insureds, and set forth factors justifiably considered in differentiating premiums, and the appropriate differential assigned to each factor.[624] Second, this information must be publicized by insurers so that insureds may ascertain the factors and applicable differentials and calculate their premiums.[625] Finally, every motorist must have the opportunity for "a prompt and effective administrative review" of his premium and how it is determined, or of the grounds for refusal or cancellation of his insurance.[626]

3. Standard of Review—Although the Supreme Court of Michigan held the no fault law's compulsory insurance requirement unconstitutional on due process grounds, other issues raised in the litigation impelled the court to determine the appropriate standard of review both for due process and for equal protection challenges. Its explication is significant for all future adjudications of these constitutional issues:

> That test to determine whether legislation enacted pursuant to the police power comports with due process is whether the legislation bears a reasonable relation to a permissible legislative objective. . . . The test to determine whether a statute enacted pursuant to the police power comports with equal protection is, essentially the same. . . .[627]

The court rejected both the heightened rationality and the strict scrutiny tests.[628] In rejecting the heightened rationality test, the court distinguished an earlier Michigan case[629] in which it had held that a challenged statute must bear a substantial relationship to the legislative goal because the "challenged statute carves out a discrete exception to a general rule and the statutory exception is no longer experimental."[630] By contrast, the *Shavers* court distinguished the no fault law as a recent innovation and entitled to judicial forbearance pending experience in its application.[631] The strict scrutiny test appeared equally inappropriate because the fundamental right to travel was not involved. The Michigan Su-

---

624. *Id.*
625. *Id.*
626. *Id.*
627. *Id.* (citations omitted).
628. *Id.* at 93.
629. Manistee Bank Trust Co. v. McGowan, 394 Mich. 655, 232 N.W.2d 636, 642 (1975) (constitutional challenge to Michigan guest statute).
630. 232 N.W.2d at 642.
631. 267 N.W.2d at 95.

preme Court agreed with the court of appeals that the right to travel "protects . . . migration, not the individual's choice of a particular means of transportation".[632]

Having rejected the more exacting tests, the *Shavers* court assessed the no fault law under a rationality standard of review, according the legislation a presumption of constitutionality. However, it interpreted this presumption more broadly than the court of appeals, declaring that legislation would be upheld when "any state of facts either known or which could reasonably be assumed" affords support for the legislation.[633] Consequently, the court could take judicial notice of ascertainable facts in reviewing legislation. A party seeking to rebut the presumption of constitutionality could establish facial invalidity by demonstrating that the legislative judgment was arbitrary. Or a challenger could claim that the legislation lacked a rational basis and introduce judicially noticable facts to support that contention.[634] If the legislative judgment could not be reviewed either in the context of purely legal argument or judicially noticed facts, however, the court stated that evidence must be adduced at trial to supply an appropriate factual foundation for the decision.[635]

Applying this analysis, the *Shavers* court was convinced that the legislative judgments underlying the passage of the novel no fault law rested on complex actuarial and statistical facts associated with the automobile insurance industry.[636] Therefore, it viewed one of its tasks as determining whether the record revealed "any reasonable state of facts" generally supporting the challenged statute.[637] Additionally, the court recognized that it had to decide whether specific provisions of the statute violated the due process and equal protection clauses of the federal and state constitutions.

4. Rulings of the Court—

---

632. *Id.* at 94 n.37. *See* Memorial Hosp. v. Maricopa County, 415 U.S. 250 (1974) (holding unconstitutional a state statute imposing a residency requirement on eligibility for publicly financed non-emergency medical treatment).

633. 267 N.W.2d at 94 (quoting United States v. Carolene Prods. Co., 304 U.S. 144, 154 (1938)). The same proposition was expressed earlier by the Michigan Supreme Court in Carolene Prods. Co. v. Thomson, 276 Mich. 172, 178, 267 N.W. 608, 610 (1936).

634. 267 N.W.2d at 94-95.

635. *Id. See* Borden's Farm Prod. Co. v. Baldwin, 293 U.S. 194, 210 (1934) (declaring that when the court must consider facts relating to commerce and industry affected by government regulation, those facts "should be presented concretely").

636. 267 N.W.2d at 95. But perhaps inconsistently, the court decided that a factual context was unnecessary for deciding whether or not the rate regulation scheme conformed to due process since that issue was "purely procedural." *Id.* at 96 n.45.

637. *Id.* at 96 (emphasis added).

(a) Partial Abolition of Tort Liability. Applying the "reasonable state of facts" test, the court upheld the partial abolition of tort liability contained in the no fault law both on equal protection and due process grounds,[638] on the ground that the no fault law was reasonably related to the legislative goal of overcoming deficiencies in the tort system.[639]

The plaintiffs also argued that the partial abolition of tort liability violated "equal protection" by creating two impermissible statutory classifications.[640] The first classification was the distinction between motor vehicle tortfeasors and their victims and all other tortfeasors and their victims; and the second was the distinction between victims of insured motor vehicle tortfeasors and victims of uninsured motor vehicle tortfeasors.[641] The *Shavers* court concluded that the first distinction was justified by the frequency and severity of injury arising from motor vehicle accidents and the need to provide a reparations system that would promptly aid accident victims.[642] The second distinction, which permitted victims of uninsured motorists to sue for injuries below the threshold, was justified as an incentive for motorists to procure the required insurance to avoid full tort exposure.[643] In the court's view, the legislative classifications did not discriminate invidiously and did not contravene constitutional principles of equal protection.

(b) Property Damage Provisions. The court reversed the lower courts' rulings[644] and upheld the property damage provisions of the no fault law. It conceded, however, that the provisions could not be justified on the basis of correcting demonstrated "evils" in the existing system because the record revealed "relatively prompt, equitable compensation for damage to property resulting from motor vehicle accidents."[645] The Michigan Supreme Court exposed the latitude of its due process review in explaining its holding on this issue:

> The analysis by the trial court and the Court of Appeals suggests that there must be an identifiable evil which the Legislature intends to correct. We do not believe this is constitutionally necessary. The

638. *Id.*

639. *Id.* at 97. It was not incumbent on the legislature to provide an "adequate substitute" before abolishing the common law remedy. Id.

640. *Id.* at 99.

641. *Id.*

642. *Id.* at 99-100.

643. *Id.*

644. *See supra* notes 563-571 and accompanying text.

645. 267 N.W.2d at 100.

> legislature is as free to experiment with other ways of dealing with a subject in the hope of making a good system better as it is to correct a perceived evil system.[646]

The legislature had anticipated a number of ameliorative effects in abolishing tort actions for property damage and fostering greater reliance on collision insurance:[647] (1) rates would be calculated on the basis of the value and repair costs of the insured's vehicle rather than of some unknown vehicle; (2) the act would also create an incentive for safer cars; (3) it would produce savings in premiums by obviating expensive accident investigations; and (4) it would encourage group insurance with lower costs by concentrating on "the risk to the insured, not the exposure to some unknown party."[648] This final prospect might attract other insurers and increase "beneficial competition" among insurance companies.[649] The anticipatory nature of these benefits did not, in the court's view, diminish the legitimacy of the goals sought to be achieved or the reasonableness of the means adopted.[650] In upholding the property damage provisions, the court stretched the presumption of constitutionality by accepting conceivable (not provable or probable) effects of the no fault legislation as legitimate goals, and declared that regulation in the economic field is due "great deference."[651]

(c) Other Provisions. The *Shavers* court also held that exclusion of two-wheeled vehicles from coverage under the act did not violate the equal protection clause of the state and federal constitutions. Treating motorcycle owners differently from other vehicle owners was justified on a premium cost basis.[652]

---

646. *Id.*

647. *Id.* at 101.

648. *Id.*

649. *Id.*

650. *Id.*

651. *Id.* Similarly, the court swept aside the equal protection challenge to the classifications created by the property protection scheme. The act provided that owners of tangible property and legally parked vehicles would be compensated by the mandatory third party (no fault) property damage insurance (up to $1,000,000) carried by the motorist who inflicted the damage, while owners of moving or improperly parked vehicles could seek compensation for damages to their vehicles only from their own insurers under optional first party collision coverage. *Id.* at 102-03. The only rationale offered by the court for approving the differentiation is that when motor vehicles and tangible property collide the motor vehicle is usually at fault. *Id.*

652. *Id.* at 103-04. The court agreed with the Michigan Court of Appeals that the economic impact of reform legislation is a valid consideration, both for legislatures and for courts. *Id.* at 104 n.63.

The court remanded the equal protection issues raised by the act's provisions on reimbursement for work loss and other services.[653] The plaintiffs had alleged that those provisions invidiously discriminated between workers in the home and other workers by assigning a different scale of benefits.[654] The court decided that a determination of these equal protection issues could not be undertaken without an adequate factual basis, which only the trial court on remand could provide.[655] Similarly it refused to rule on the consitutionality of the act's differing treatment of resident and nonresident motorists. It found no adequate factual record upon which to adjudicate those constitutional issues, and therefore remanded to the trial court to receive relevant evidence.[656]

5. *Summary*—The Michigan Supreme Court decision rendered in the *Shavers* case ranks as the most significant judicial pronouncement to date in no fault jurisprudence. Its substance may be distilled in two counterpoised principles. First, the standard of constitutional review for economic regulatory reforms such as compulsory first party motorist benefits, is a minimum rationality test. Judicial deference is due legislative choices in identifying socially desirable goals and engineering means to achieve those goals. Second, the consumer-motorist may not be forced to purchase protective no fault insurance unless he is assured of fair and equitable ratemaking, both procedurally and substantively. To arrive at that conclusion, the court invoked the entitlement doctrine and accorded a status akin to a property right to the motorist's concern over the determination of his premium and his opportunity to obtain and maintain insurance at fair rates. This entitlement to fairness imposed due process requirements on insurers and administrators that the legislature must prescribe.[657]

---

653. *Id.* at 105. The court construed MICH. COMP. LAWS § 500.3107(b) (1983), which fixes reimbursement for services in the household at a maximum of $20 per day for three years and work loss for employees at a maximum monthly payment indexed to the cost of living for up to three years.

654. 267 N.W.2d at 104. Additionally, recovery for home employees was restricted to reimbursement for expenses "reasonably incurred," that is, the cost of the substitute services must be actually incurred before they could be reimbursed. *Id.*

655. *Id.* at 104-05.

656. *Id.* at 105. The portions of the act relating to this problem may be found at MICH. COMP. LAWS §§ 500.3102(1), .3113(c), .3135(2) (1983). The trial court had held that section 3113(c), which denied nonresidents recovery of no fault benefits and also of tort recovery below the threshold, was invalid. *See* 267 N.W.2d at 105. *See also supra* note 559 and accompanying text.

657. In its decision, the Michigan Supreme Court pursued a due process analysis considerably beyond that of the Massachusetts and Connecticut Supreme Courts which upheld

E. Legislative Reaction—1979 Amendments

Responding to the *Shavers* decision, the Michigan legislature enacted Public Acts 145[658] and 147,[659] which were approved by the governor on November 13, 1979. Public Act 145 dealt with ratemaking, underwriting, and modifying or supplementing specific no fault sections.[660]

1. Amendments Relevant to Ratemaking Practices—Section 2107[661] required automobile insurers to make filings to be effective January 1, 1981. Section 2108[662] prescribes the content of the filing, which includes manuals of classification, rules, rates, rating plan, character and extent of coverage contemplated. Additionally, that section requires the submission of statistics, experience data and other information supporting the filing.

Section 2109(1)(a) requires automobile insurers to establish rates which are not "excessive, inadequate or unfairly discriminatory."[663] That section describes a rate as excessive if it is "unreasonably high" for the coverage and "a reasonable degree of competition does not exist for the insurance to which the rate is applicable."[664] A rate is inadequate if it is "unreasonably low," and use of the rate endangers the insurer's solvency or will destroy competition.[665] A rate is deemed unfairly discriminatory compared

---

the regulatory systems of those states. *See supra* notes 85-135, 257-279 accompanying text.

658. Act of Nov. 13, 1979, Pub. L. No. 79-145, 1979 Mich. Pub. Acts 849 (codified at MICH. COMP. LAWS §§ 500.1209, .2101-.2131, .2403, .2603, .2604, .2827, .2901, .2912, .2920, .2921, .2923-.2925(c), .2930, .2930(a), .2932, .3037, .3102, .3135, .3301, .3303, .3310, .3320, .3330, .3340, .3350, .3355, .3365, .3370, .3380, .3385 (1983).

659. Act of Nov. 13, 1979, Pub. L. No. 79-147, 1979 Mich. Pub. Acts 894 (codified at MICH. COMP. LAWS §§ 500.2604, .3037, .3135 (1983)).

660. MICH. COMP. LAWS §§ 500.3101-.3179 (1983). As it pertains to automobile insurance, Public Act 145 changed the effective dates for sections .3037 and .3135 of the act. *Id.* §§ 500.3037, .3135. Section .3037 prescribes the options in collision coverage for no fault insureds: limited coverage—full payment without a deductible when the operator is not substantially at fault and broad form and standard collision coverage with deductibles. *Id.* § 500.3037(1). Section .3135 defines residual tort liability for noneconomic loss and damages in excess of allowable expenses, work loss and survivor's loss. *Id.* § 500.3135. Furthermore, a mini-property damage action is excepted from the abolition of tort liability. Tort liability is abolished except as to: "Damages up to $400.00 to motor vehicles, to the extent that the damages are not covered by insurance. An action for damages pursuant to this subdivision shall be conducted in compliance with subsection (3)." *Id.* § 500.3135(2)(d).

661. *Id.* § 500.2107(1).

662. *Id.* § 500.2108(1)-(3).

663. *Id.* § 500.2109(1)(a).

664. *Id.* The amendment provided that in determining the degree of competition one should consider number of insurers writing the same kind of insurance, availability, underwriting return, and market entrance barriers encountered by new insurers. *Id.* § 500.2109(2).

665. *Id.* § 500.2109(1)(b).

with another rate if the differential is not justified by losses, expenses or uncertainty of loss.[666]

Section 2110 enumerates factors to be considered in developing and evaluating rates under section 2109—loss experience, both within and without the state, relating to catastrophic hazards, underwriting profit, dividends, savings and "all other relevant factors within and outside this state."[667] Section 2111 delimits and restricts the factors which may be the basis for classifications and territorial base rates.[668] Classifications must be based only on any or all of the following factors: (1) driver's age, experience, or years licensed; (2) driver primacy; (3) average miles weekly or annually; (4) vehicle characteristics; (5) commuting mileage; (6) number of cars or operators in the household; (7) amount of insurance.[669] In establishing personal protection insurance, the factors to be considered are: (1) earned income, (2) number of dependents of insured, and (3) coordination of benefits.[670] Additional factors with respect to collision and comprehensive insurance are the anticipated cost of repairs, the vehicle make and model, and the vehicle design as that relates to damageability.[671] Section 2111 prohibits the use of sex or marital status in determining rates,[672] and provides that an insurer may not utilize more than twenty different territorial base rates for automobile insurance coverage.[673] The scope and particularity of regulations imposed by section 2111 significantly affect traditional ratemaking practices.

Section 2112 requires insurers to inform their policyholders annually concerning rating classifications affecting their premiums, procedures for obtaining more detailed information, rights of in-

---

666. *Id.* § 500.2109(1)(c).

667. *Id.* § 500.2110(1).

668. *Id.* § 500.2111. This section contributes substantially to compensating for the due process deficiencies of the no fault law by prescribing exclusive rating factors.

669. *Id.* § 500.2111(2)(a).

670. *Id.* § 500.2111(2)(b).

671. *Id.* § 500.2111(2)(c). This section also requires insurers to establish a secondary or merit rating plan incorporating premium surcharges based upon "substantially at-fault" accidents, or civil or criminal liability resulting from violation of law. *Id.* at § 500.2111(3). The act defines "substantially at fault" as denoting the person who was "more than 50% of the cause of an accident." *Id.* § 500.2104(4).

672. *Id.* § 500.2111(4).

673. *Id.* § 500.2111(5)(a). Furthermore, no territorial base rate for an automobile insurance package policy shall be less than 45% of the highest territorial base rate for the same policy or less than 90% of the territorial base rate for the same policy in any adjacent territory. *Id.* § 500.2111(5)(b), (c). Exemptions and variances are also provided for. *Id.* § 500.2111(6)-(10).

sureds, and underwriting rules pertaining to eligibility points.[674]

2. Amendments Affecting Underwriting Practices—Section 2118 requires an insurer to write insurance for all eligible persons.[675] Section 2119 directs that insurers file their underwriting rules with the Commissioner,[676] and apply the rules uniformly throughout the state.[677] Sections 2122 and 2123 vest insurers with certain duties relating to declination or termination of insurance;[678] the insurer must inform the applicant of specific reasons for refusal to insure or for termination.[679]

3. Amendments of the No Fault Sections—Some of the post-*Shavers* modifications incorporated in Public Act 145 specifically affected the no fault provisions of the Michigan automobile insurance scheme. Section 3037, for example, makes available limited and broad collision coverage.[680] Section 3102 requires nonresident owners of motor vehicles not registered in Michigan to maintain security for payment of benefits if the vehicle is operated within the state for an aggregate of more than thirty days in a calendar year.[681]

Section 3135, the crucial tort action-tort exemption section,[682] was amended to permit a tort action for certain motor vehicle damage not covered by insurance.[683] The original portions of the

---

674. *Id.* § 500.2112. Additionally, a person aggrieved by a denial of insurance or by an incorrect premium charge, has the right to a managerial-level conference with the insurer. *Id.* § 500.2113(1). If not resolved at the conference, the matter will be decided by the Commissioner informally or, if further pursued, formally as a contested case. *Id.* § 500.2113(1),(3),(5). Similar duties to supply the consumer with information and specifically to provide competitive premium quotations of several insurers are prescribed for licensed insurance agents. *Id.* § 500.2116(1).

675. *Id.* § 500.2118(1). The act excludes from eligibility persons whose licenses are suspended or revoked, or whose policies have been cancelled for nonpayment or who have been convicted of various offenses. *Id.* § 500.2103. Other permissible underwriting criteria are: driving records of household members, speed modifications of vehicle, use and type (but not age) of vehicle. *Id.* § 500.2118(2). Claim experience of a person in relation to comprehensive coverage may be considered only where an insurer seeks to justify a required deductible. *Id.* § 500.2118(2)(h).

676. *Id.* § 500.2119(6). The Commissioner may prohibit use of an underwriting rule after a hearing is held on the rule. *Id.*

677. *Id.* § 500.2119(2).

678. *Id.* §§ 500.2122, .2123.

679. *Id.* §§ 500.2122(1), 2123(1).

680. *Id.* § 500.3037(1). Limited coverage applies to collision damage to the insured vehicle without deductibles when the operator was not substantially at fault.

681. *Id.* § 500.3102(1).

682. *Id.* § 500.3135.

683. *Id.* § 500.3135(2)(d). Liability is incurred for "[d]amages up to $400 to motor vehicles, to the extent that the damages are not covered by insurance." *Id.* The amendment also provided that damages would be assessed on a comparative fault basis, and would not

section remained unchanged.

Public Act 145 also directed insurers to participate in an organization to guarantee automobile insurance coverage to any applicant not able to acquire insurance by ordinary methods.[684]

4. Summary of Amendments—These provisions, taken as a whole, aim at correcting the shortcomings noted by the Supreme Court of Michigan in the *Shavers* opinion. The amending legislation focuses on increasing the availability of insurance, detailing the factors to be weighed in ratemaking and underwriting, better informing the consumer and opening more avenues for the insured to question his classification and premium. However, the amendments leave in place a very liberal personal injury protection scheme with unlimited benefits for allowable medical expenses and indexed work loss and survivors' benefits substantially exceeding the no fault plan of any other state. Thus Michigan's compulsory no fault insurance requirement is now clothed with the due process guarantees found wanting in the original statute.

F. Opinion on Remand to the Circuit Court

On November 21, 1979 the Supreme Court of Michigan issued its mandate to the Circuit Court of Wayne County to hear proof on issues relating to the 1979 Amendments.[685] All parties agreed that the new legislation corrected the deficiencies of the original act in failing to supply the consumer with sufficient information concerning the calculation of premiums. They also agreed that the amendments corrected the inadequacies in facilitating effective administrative review of insurers' determination of premiums and of

---

be assessed "in favor of" one who is "more than 50% at fault." *Id.* § 500.3135(3)(a). Actions under section 3135(2)(d) must be commenced in small claims court and will not be res judicata as to any other liability. *Id.* § 500.3135(4),(5). Furthermore, "[l]iability shall not be a component of residual liability, as prescribed in section 3131, for which maintenance of security is required by this act." *Id.* § 500.3135(3)(b).

684. *Id.* § 500.3301(1)(a). This organization is to be known as the "Michigan Automobile Insurance Placement Facility." *Id.* § 500.3301(2). For the organization of the Facility and the adoption of a plan of operation, see *id.* § 500.3310. The 1979 Act also supplies details of the Facility's operation to ensure equitable distribution of applicants, availability of insurance to qualified applicants, reasonable procedures for handling claims, responsiveness to the consumer and to supervision by the Commissioner. *Id.* §§ 500.3320, .3330, .3340, .3350, .3355, .3365, .3370, .3380, .3385.

685. *See* Shavers v. Attorney Gen., 412 Mich. 1105, 315 N.W.2d 130 (1982). Relevant regulations of the Insurance Commissioner had not been approved by the legislature; therefore, the case proceeded solely on the statutory questions resolved into four issues of law. Circuit Judge Horace W. Gilmore issued his opinion on June 16, 1980. Shavers v. Kelley, No. 73-248-068-CZ (Wayne County, Mich. Cir. Ct. June 16, 1980).

refusals or cancellations of insurance.[686]

Disputed questions, however, were whether the 1979 amendments had infused meaning in the exhortation that rates not be excessive, inadequate or unfairly discriminatory, and whether they had structured a fair ratemaking mechanism.[687] The trial court found that a main theme of the amendments was that "underwriting and rating decisions shall be based upon ascertainable uniform standards reasonably related to the concept of risk."[688]

The new statute particularized the factors that insurers were permitted to consider in making fair coverage and premium decisions. The legislature also identified criteria to test the adequacy and fairness of rates[689] to assess whether a fair degree of competition existed.[690] "If rates are unjustifiably high compared to the amount of losses or expenses for the insurance, underwriting returns will then be excessive. Under the statutory definition a 'reasonable degree of competition' may then be found not to exist, and a rate may be deemed 'excessive.' "[691]

The amendments also imposed duties on insurers to supply information concerning competitive rates,[692] to accept virtually all applicants and to arrange for insurance offerings by the Automobile Placement Facility[693] on the voluntary market where insufficient competition is present. The court found section 2111 especially important as it prescribed the sole rating factors governing the pricing of insurance:[694] "By setting ascertainable rating factors in Section 2111, Act 145 has moved to end unfair, subjective treatment of individuals."[695]

The aggregate of the above provisions led the circuit court to conclude that Public Act 145 imparted substantial meaning to the statutory standard that no fault rates shall not be excessive, inadequate or unfairly discriminatory. The amended legislation also provided the necessary procedure for motorists to seek administrative

---

686. *See* 315 N.W.2d at 132 (Levin, J., dissenting) (quoting the circuit court).

687. *Id.*

688. *Id. See* MICH. COMP. LAWS §§ 500.2109, .2110, .2111, .2118, .2119 (1983). *See also supra* notes 689-696 and accompanying text.

689. MICH. COMP. LAWS § 500.2109(1)(a)-(c) (1983).

690. *Id.* § 500.2109(2).

691. Shavers v. Kelley, No. 73-248-068-CZ, slip op. at 10 (Wayne County, Mich. Cir. Ct. June 16, 1980).

692. MICH. COMP. LAWS §§ 500.2112, .2116(1), .2118 (1983).

693. *See id.* §§500.2119(2), .2128.

694. No. 73-248-068-CZ, slip op. at 12.

695. *Id.*

review of premiums, refusals and cancellations.[696] Overall, the circuit court was satisfied that the legislature had imposed accountability on insurers and safeguarded fairness of ratemaking by emphasizing competition, and administrative and market place controls.

G. *Post*-Shavers *Cases*

1. *Offsets: Survivors' Benefits, Medicaid, Medicare—*

(a) O'Donnell v. State Farm Mutual Automobile Insurance Co. The question of whether governmental benefits must be set off against no fault benefits was left unresolved by *Shavers*. That provision of the no fault law had generally received a negative reaction from Michigan courts in other cases.[697] In *O'Donnell v. State Farm Mutual Automobile Insurance Co.*,[698] the Michigan Supreme Court, in a 4-3 decision, upheld section 3109(1) of the no fault act,[699] which provides: "[B]enefits provided or required to be provided under the laws of any state or the federal government shall be subtracted from the personal protection insurance benefits otherwise payable for the injury."[700] The governmental benefits involved were survivors' benefits under the Social Security Act, payable to the widow and children of the accident victim.[701] Michigan's no fault law limited the maximum payable survivors' benefits to $1000 per month. The insurance company asserted that section 3109(1) permitted it to subtract Social Security benefits from the $1000 no fault survivors' benefits and pay only the difference.[702]

The *O'Donnell* court upheld section 3109(1) on due process

---

696. The circuit court did not examine this issue directly. Rather, the court concentrated on fairness in availability of insurance and in ratemaking. However, sections 2122 and 2123 do set out insurers' duties, including providing proper and timely notice and reasons for termination or refusal to insure.

697. MICH. COMP. LAWS § 500.3109(1) (1983), which authorizes an insurer to set off benefits received from governmental sources previously had been held unconstitutional. *See* Mielke v. Michigan Millers Mut. Ins. Co., 82 Mich. App. 721, 267 N.W.2d 165 (1978), *rev'd*, 406 Mich. 858, 275 N.W.2d 555 (1979); Pollock v. Frankenmuth Mut. Ins. Co., 79 Mich. App. 218, 261 N.W.2d 554 (1977); Wysocki v. Detroit Auto. Inter-Ins. Exch., 77 Mich. App. 565, 258 N.W.2d 561 (1977), *rev'd*, 406 Mich. 857, 275 N.W.2d 551 (1979). *Contra* Greene v. State Farm Mut. Auto. Ins. Co., 83 Mich. App. 505, 268 N.W.2d 703 (1978); Smart v. Citizens Mut. Ins. Co., 83 Mich. App. 30, 268 N.W.2d 273 (1978).

698. 404 Mich. 524, 273 N.W.2d 829, *appeal dismissed*, 444 U.S. 803 (1979).

699. MICH. COMP. LAWS § 500.3109(1) (1983).

700. *Id.*

701. 273 N.W.2d at 832.

702. *Id. See* MICH. COMP. LAWS § 500.3108(1) (1983) ($1000 per month limitation).

grounds. As in *Shavers*, the court found a valid legislative purpose in requiring the setoff—to relieve all insureds of premium costs for duplicative benefits.[703] The purpose of payments to survivors under the Social Security Act—to alleviate the economic hardship occasioned by loss of the wage earner's contributions—paralleled section 3018's rationale of support for dependents. Payments to a beneficiary under both systems would result in double recovery for the same loss. The setoff required in section 3109(1) was not arbitrary but rather rationally related to a valid legislative goal, and hence not a violation of state or federal due process requirements.[704]

Secondly, the supreme court considered plaintiffs' claim that section 3109(1) discriminated against those receiving governmental benefits, and thus violated equal protection guarantees.[705] The section did not require that *private* benefits be set off against no fault entitlements.[706] Those with private insurance covering the same loss could receive a $1000 per month maximum for three years under no fault and, additionally, whatever benefits were due from the nongovernmental insurance.[707] The court viewed this classification and the resulting differential payments as a rational compromise by the legislature. By incorporating the setoff for governmental benefits, premium rates could be contained. By not requiring setoff for private health and accident benefits, persons with greater than $1000 per month income needs could purchase extra coverage and directly bear the burden of providing for those extra needs.[708] The distinction drawn by the legislature rested on a rational basis, and unless experience later proved the legislative premises invalid, section 3109(1) did not offend equal protection.[709]

---

703. 273 N.W.2d at 835.

704. *Id.* at 836.

705. *Id.*

706. *See* MICH. COMP. LAWS § 500.3109(1) (1983).

707. *See* 273 N.W.2d at 836.

708. *Id.* at 836-37.

709. *Id.* at 837. In fact, the legislature attempted to fine tune the set-off provision by enacting section 3109(1), which requires the personal protection insurer to offer reduced premium rates, deductibles and exclusions related to an insured's other health and accident coverage. This voluntary set-off applies only to the named insured, his or her spouse and relatives of either domiciled in the same household. The court also noted that the United States Supreme Court had upheld the required set-off of workers' compensation benefits (but not private benefits) from Social Security benefits otherwise payable. *Id.* at 838 (citing Richardson v. Belcher, 404 U.S. 78 (1971)). The court stated: "If the goals sought are legitimate, and the classification adopted is rationally related to the achievement of those goals, then the action of Congress is not so arbitrary as to violate the Due Process Clause of the Fifth Amendment." *Id.* at 84. It should be noted that the *O'Donnell* opinion is carefully

(b) Davey v. Detroit Automobile Inter-Insurance Exchange. Survivors' benefits were also at issue in a 1980 court of appeals case, *Davey v. Detroit Automobile Inter-Insurance Exchange.*[710] Under the statutory scheme existing prior to 1978, cost-of-living adjustment applied to work loss but not to survivors' benefits. The court in *Davey* held that this differential treatment violated equal protection. Therefore, the plaintiff could recover survivors' benefits equal to the benefits her husband would have received had he survived.[711] The decision authorized retroactive recovery on an indexed basis.

The legislature amended the survivors' benefits section in 1978,[712] entitling dependents of insureds who succumb to automobile accident injuries to benefits adjusted annually to reflect the cost of living. Thus, payments for survivors' and injured insureds' work loss benefits are now determined on the same scale.

(c) Workman v. Detroit Automobile Inter-Insurance Exchange. A different benefits issue was addressed by the Michigan Supreme Court in *Workman v. Detroit Automobile Inter-Insurance Exchange.*[713] There the court held that Medicaid was not a benefit provided or required to be provided under federal or state law.[714] Instead, it was assistance available under Michigan state law for the medically indigent.[715] At the time of the plaintiffs' injury in *Workman*, Michigan law defined a medically indigent per-

---

confined to the issue of federal Social Security survivors' benefits and does not examine other governmental benefits under the offset requirement. For details of the legislature's consideration of the issue of set-offs and the various bills introduced at the time the no fault act was under consideration, see Justice Williams' dissenting opinion. 273 N.W.2d at 846 n.13 (Williams. J., dissenting). Justice Williams also made several other points: The governmental benefits required to be set off by section 3109(1) are paid for by the wage earner through contributions (wage deductions) and are not "free" payments. Thus, they are in this sense similar to private insurance for which the insured pays. Section 3109(1) however, requires only governmental benefits to be set off. This set-off is based on the fund or source of the benefits, not on whether the beneficiary has contributed to the protective scheme. In applying to all governmental benefits, section 3109(1) might permit a pension benefit to be set off against no fault recovery. The burden of the setoff discriminatorily falls only on one class of no fault insureds—those contributing to and covered by government insurance. If any no fault premium reduction results, it advantages equally privately and nonprivately insureds and burdens only the class of governmental benefits recipients. Justice Williams thus concluded that section 3109(1) was facially unconstitutional. *Id.* at 838-41.

710. 98 Mich. App. 123, 296 N.W.2d 12 (1980).

711. 296 N.W.2d at 14.

712. Act of Oct. 16, 1978, Pub. L. No. 78-459, 1978 Mich. Pub. Acts 459 (codified at MICH. COMP. LAWS § 500.3108 (1983)).

713. 404 Mich. 477, 274 N.W.2d 373 (1979).

714. 274 N.W.2d at 382.

715. *Id.*

son as one to whom no legal obligation was owed for medical payments by a contractor, public or private.[716] Because the plaintiff could claim no fault personal protection coverage, she was "expressly precluded from qualifying as a medically indigent individual eligible for medical assistance under the state Medicaid program."[717] The *Workman* court found support for its conclusion in the actions of the Michigan legislature, which had addressed this issue in 1976, by amending the Social Welfare Act to exclude Medicaid eligibility when the claimant possesses no fault coverage.[718]

2. Offsets: Workers' Compensation Benefits—

(a) Moore v. Travelers Insurance Co. The Federal District Court for the Eastern District of Michigan held that the mandatory set-off of section 3109(1) was constitutional and that it covered workers' compensation benefits in *Moore v. Travelers Insurance Co.*[719] However, the court held that workers' compensation benefits may not be subtracted from personal injury no fault benefits unless they pertain to the same element of loss and the same time period.[720] Thus, workers' compensation payments for wage loss may be offset against no fault wage loss benefits applicable to the same months or years. On the other hand, workers' compensation payments for medical expenses could not be offset against claimed no fault wage loss benefits.[721]

(b) Mathis v. Interstate Motor Freight System. The relation of no fault benefits to workers' compensation benefits was examined again by the Michigan Supreme Court in *Mathis v. Inter-*

---

716. MICH. COMP. LAWS §§ 400.105, 106 (Supp. 1982-83).

717. 274 N.W.2d at 832. *But see* LeBlanc v. State Farm Mut. Auto. Ins. Co., 410 Mich. 173, 301 N.W.2d 775 (1981). In *LeBlanc*, the Michigan Supreme Court decided that Medicare payments to a senior citizen who was a no fault insured need not be deducted from no fault benefits. The court placed Medicare benefits under the permissive provisions of section 3109(a), rather than the mandatory set-off of section 3109(1). Under the former, insurers may offer reduced premium rates, coordinating no fault insurance with "other health and accident coverage on the insured." MICH. COMP. LAWS § 500.3109(a) (1983). But such coordination is at the option of the insured. *Id.* In *LeBlanc*, the insured had not elected to have Medicare coverage reflected in reduced no fault premiums and was, therefore, not subject to a dry-off of the Medicare payments.

718. 274 N.W.2d at 382 n.9. *See* MICH. COMP. LAWS § 400.106(1)(b) (Supp. 1982-1983).

719. 475 F. Supp. 891 (E.D. Mich. 1979). Other cases have upheld the constitutionality of section 3109(1) for workers' compensation offsets. *See* Lindsey v. Hartford Accident & Indem. Co., 90 Mich. App. 668, 282 N.W.2d 440 (1979); Hubert v. Citizens Ins. Co., 88 Mich. App. 710, 279 N.W.2d 48 (1979).

720. 475 F. Supp. at 894.

721. *Id.*

state Motor Freight System.[722] That litigation consolidated four cases raising the issue of recovery both under the Workers' Disability Compensation Act[723] ("WDCA") and the no fault law. In all of the cases, an employee was injured in the course of employment while occupying his employer's motor vehicle. The employees or their dependents received workers' compensation benefits and also sought to recover no fault personal protection benefits. The *Mathis* court concluded that the exclusive remedy provision of the WDCA applied to actions against the employer and did not bar an employee who had collected workers' compensation from receiving no fault benefits from the carrier insuring his employer's vehicle.[724] If the employer is self-insured under the no fault law, the employee may collect no fault benefits as well as workers' compensation from the employer. However, under section 3109(1) of the no fault act, workers' compensation benefits must be set off against no fault benefits.[725] The court rejected the plaintiffs' argument that the offset would violate equal protection, reasoning that the claims both for workers' compensation and no fault benefits stemmed from the same accident, and payments could result in duplicative recoveries.[726]

3. Offsets: Tort Recoveries—Section 3135 of the no fault act authorizes tort actions for noneconomic loss if the victim suffered death or serious injury, and for damages for allowable expenses, work loss and survivors' loss in excess of the limitations governing those losses.[727] Section 3116(1) states that recovery on a tort claim shall be subtracted from personal protection insurance benefits.[728] A lower court held that this section violated equal protection and due process clauses of the Michigan Constitution.[729] On by-pass leave to appeal, the Michigan Supreme Court perceived "an appar-

---

722. 408 Mich. 164, 289 N.W.2d 708 (1980).

723. MICH. COMP. LAWS §§ 418.101-941) (Supp. 1982-1983).

724. 289 N.W.2d at 714-15. MSection 3114(3) of the Michigan no fault act provides:

> An employee, his or her spouse, or a relative of either domiciled in the same household, who suffers accidental bodily injury while an occupant of a motor vehicle owned or registered by the employer, shall receive personal protection insurance benefits to which the employee is entitled from the insurer of the furnished vehicle.

MICH. COMP. LAWS § 500.3114(3) (1983).

725. 289 N.W.2d at 714.

726. *Id.* at 716 (applying O'Donnell v. State Farm Mut. Auto. Ins. Co., 404 Mich. 524, 273 N.W.2d 829 (1979), which allowed offset of survivor's benefits against no fault benefits).

727. *See* MICH. COMP. LAWS § 500.3135, .3107 (1983).

728. *Id.*

729. *See* Workman v. Detroit Auto. Inter-Ins. Exch., 404 Mich. 477, 274 N.W.2d 373, 377 (1979) (quoting from the trial court's opinion).

ent and patent absurdity: the Legislature, on the one hand, provides an injured person limited tort remedy in § 3135 of the act, while, on the other hand, providing that any tort recovery achieved pursuant to § 3135 will be effectively taken away under § 3116 of the act."[730]

To overcome the incompatibility of the two statutory provisions, the court probed the legislative intent and inferred a purpose to permit added compensation in cases of catastrophic injury and extraordinary economic losses:

> Accordingly, in light of § 3135, we construe § 3116 to mean that an insurance carrier paying personal injury protection benefits is entitled to reimbursement from the tort recovery of a person injured as a result of a motor vehicle accident only if, and to the extent that, the tort recovery includes damages for losses for which personal injury protection benefits were paid. Thus, since § 3135 abolishes tort remedy for losses covered under the personal injury protection insurance provisions of the act, an injured plaintiff should recover nothing for which the insurance carrier will have a right of reimbursement under § 3116. We believe this interpretation of § 3116 not only gives full effect to § 3135, but it also effectuates the essential purposes of this section, namely to prevent double recovery of economic loss by those persons who retain their right to sue in tort for economic loss under the act.[731]

Offset of a tort recovery was presented in a more complicated context in *Great American Insurance Co. v. Queen*.[732] Queen suffered injury in a motor vehicle accident in the course of his employment in 1976 and Great American subsequently paid him workers' compensation benefits. Queen then claimed benefits from his employer's no fault insurer. The no fault insurer subtracted the compensation benefits from the no fault benefits due. Under section 3135, Queen filed a claim against third party tortfeasors and settled with them for $18,500. Great American then brought suit against Queen and the third party tortfeasors, asserting a lien on the proceeds of the settlement. The courts below gave summary judgment for the defendants and the Michigan Supreme Court affirmed.[733]

The supreme court concluded that "the Legislature did not intend that the third-party recovery of a person injured in the course

730. 274 N.W.2d at 386.
731. *Id.*
732. 410 Mich. 73, 300 N.W.2d 895 (1980).
733. 300 N.W.2d at 897.

of his employment be subject to greater subrogation claims in favor of insurers."[734] Therefore, workers' compensation carriers are subject to the same reimbursement provisions as no fault insurers when compensation benefits substitute for no fault benefits.[735] The court reasoned that where workers' compensation benefits do not substitute for no fault benefits but actually exceed no fault benefits, the compensation carrier has the same right of reimbursement from third party tort recoveries as in the case of non-motor vehicle injuries.[736] In *Queen*, the carrier was not permitted reimbursement for medical treatment expenses, since the payment of workers' compensation benefits substituted for what the injured person would have received through no fault payments.

In *Queen*, as in *Mathis*,[737] the court attempted to resolve issues at the intersection of the WDCA and the no fault law. The distinguishing factor in *Queen*, however, was the presence of a third party tortfeasor from which Queen obtained tort damages in addition to receiving workers' compensation and no fault benefits. The WCDA provides:

> In an action to enforce the liability of a third party, the plaintiff may recover any amount which the employee . . . would be entitled to recover in an action in tort. Any recovery against the third party for damages . . ., after deducting expenses . . ., shall first reimburse the employer or carrier for any amounts paid or payable under this act. . . .[738]

Thus, under the WDCA, the plaintiff in a tort action may recover any amount of economic and noneconomic damages, while under no fault he is limited to recovery of noneconomic losses and economic losses that exceed the no fault ceilings. To equate the recovery rights of employee and nonemployee motor accident victims, the court in *Queen* announced a uniform rule that reimbursement out of a tort recovery is required only where an overlap or duplication has occurred.[739] Specifically, neither a no fault insurer nor a workers' compensation carrier is entitled to reimbursement of eco-

734. *Id.* at 901.

735. *Id. See* MICH. COMP. LAWS § 500.3116 (1983). *See also* Reliance Ins. Co. v. Messina Trucking, Inc., 83 Mich. App. 159, 268 N.W.2d 328 (1978) (holding that the compensation carrier could not recover benefits it had paid the insured for economic loss from the insured's tort recovery of noneconomic damages).

736. 300 N.W.2d at 901. *See* Pelkey v. Elsea Realty and Inv. Co., 394 Mich. 485, 232 N.W.2d 154 (1975) (a reimbursement case arising prior to the no fault law).

737. *See supra* notes 722-731 and accompanying text.

738. MICH. COMP. LAWS § 418.827(5) (Supp. 1982-1983).

739. 300 N.W.2d at 899-901.

nomic loss payments out of a victim's recovery of noneconomic damages from a tortfeasor.[740]

Justice Williams supplemented the court's opinion in *Queen* with an interesting analysis.[741] He concluded that the no fault act, which had been enacted after the WDCA, repealed by implication the conflicting section of the WDCA. He found a clear and irreconcilable repugnance between the WDCA provision and section 3135 of the no fault law.[742] He concluded that allowing both provisions to stand would violate the state constitutional guarantee of equal protection.

4. Motorcycles and Deductibles—In *Shavers v. Kelley*, the Michigan Supreme Court sustained the constitutionality of the no fault act's exclusion of motorcyclists.[743] That decision, however, did not resolve several recurrent issues of constitutional uncertainty. Under the statute, motorcyclists are required to maintain security against liability to others for negligently caused injuries, but not required to purchase no fault insurance.[744] If motorcyclists are nonetheless entitled to receive no fault benefits, automobile drivers might claim a denial of equal protection and due process because they must purchase no fault insurance under the act, and presumably their premiums would cover no fault benefits payable to motorcyclists who are not similarly obligated. In *Underhill v. Safeco Insurance Co.*,[745] the Michigan Supreme Court decided that no constitutional invalidity results from this differential treatment.

A second constitutional question raised in *Underhill* addressed a $5000 deductible included in a "limits of liability" clause applicable to injuries arising from the use of an owned motorcycle by the "named insured or a relative."[746] The plaintiff asserted that the no fault act unconstitutionally delegated authority to the Commissioner of Insurance to set deductibles without imposing proper statutory standards.[747] The Michigan Court of Appeals held the delegation unconstitutional but the Michigan Supreme Court re-

740. *Id.*

741. *Id.* at 905-910 (Williams, J., for affirmance).

742. *Id.* at 908.

743. 267 N.W.2d at 103-04. Motorcycles are excluded from the class of motor vehicles whose owners must maintain no fault security. MICH. COMP. LAWS § 500.3101(1),(2) (1983).

744. *See* MICH. COMP. LAWS § 500.3101(1),(2) (1983).

745. 407 Mich. 175, 284 N.W.2d 463 (1979).

746. 284 N.W.2d at 466-67. The insurance involved was the policy of the motorcycle operator's father who owned an automobile insured by Michigan Mutual Liability Company.

747. *Id.* at 471.

versed and remanded the issue.[748] On remand, the court of appeals sustained the $5000 deductible coverage provisions of the no fault act,[749] section 3109(3), which permitted a $300 deductible provision and stated that "[a]ny other deductible provisions require the prior approval of the commisioner."[750]

5. Miscellaneous—In *Dolson v. Assigned Claims Facility*,[751] the Michigan Court of Appeals sustained the one-year statute of limitations for recovery of personal injury protection benefits[752] against claims that the limitation violated the federal and state equal protection and due process guarantees.

XV. CONCLUSION

Future constitutional litigation in the field of no fault automobile accident insurance can be expected to pursue the individual rights approach emphasizing equal protecton and due process. A second generation of cases, introduced by the Michigan Supreme Court's decision in *Shavers*, may focus narrowly on possible specific inequities of no fault schemes and question broadly the practices and functioning of the entire state insurance system to ascertain its fidelity to due process. The assumption that such litigation will arise in the familiar structure of state regulation of insurance is reasonable but not absolutely ineluctable. Should the federal government re-enter the field, its course of action might be shaped by its interstate commerce power, by its power to protect individual rights from inimical state action or by both of these constitutional sources.

A. Resumption of Congressional Power Over Insurance

If the federal government evinced greater interest in expanding rather than contracting regulatory authority over the insurance industry, Congress could repeal the McCarran-Ferguson Act and subject insurers to the proscriptions of the antitrust laws. The *South-Eastern Underwriters* case[753] unequivocally identified insurance as an enterprise in interstate commerce. It would be clearly within the authority of Congress to regulate insurance in-

748. *Id.* at 471-72.
749. Porter v. Michigan Mut. Liab. Co., 97 Mich. App. 281, 293 N.W.2d 799 (1980).
750. MICH. COMP. LAWS § 500.3109(3) (1983).
751. 83 Mich. App. 596, 269 N.W.2d 239 (1978) (per curiam).
752. MICH. COMP. LAWS § 500.3174 (1983).
753. 322 U.S. 533 (1944). *See supra* notes 2-11 and accompanying text.

dustry practices in restraint of trade. The longstanding tradition of insurer membership in rating bureaus, which develop rates and rating systems adopted by many companies, illustrates widespread collaboration on "pricing the product."

State law, however, may explicitly protect these practices. Connecticut, for example, unabashedly authorizes insurers to "act in concert . . . with respect to any matters pertaining to the making of rates or rating systems . . . underwriting rules,"[754] and provides further that "[m]embers and subscribers of rating or advisory organizations may use the rates, rating systems, underwriting rules or policy or bond form of such organizations, either consistently or intermittently, but . . . shall not agree with each other or rating organizations or others to adhere thereto."[755]

If the McCarran-Ferguson antitrust exemption were removed, such state statutory authority for rate-fixing could scarcely survive. Insurance rating organizations and trade associations are national in scope and even if a company operates solely within one state, its conformity or nonconformity to multi-state concerted action may be said to affect interstate commerce. More than conspiratorial combinations in restraint of trade would be comprehended by repeal of the McCarran-Ferguson Act. Reassertion of the federal commerce power could spur inquiry into other private industry practices which burden or affect interstate commerce, and arouse a heretofore dormant body of litigation.

The probable effect of repealing the McCarran-Ferguson Act on state legislation and administrative regulation in the no fault sphere is unclear. The litigants in the cases reviewed in the foregoing discussion do not appear to have pursued the possible effects that assertion of the national commerce power might produce. The Massachusetts court in *Pinnick v. Cleary*[756] indicated that had an argument been advanced that the state no fault law, in purpose or effect, significantly affected interstate commerce, it might have considered employing a "less restrictive alternative test" to measure the statute's constitutionality.[757] In its analysis, the *Pinnick* court positioned that test between the compelling state interest and the mere rationality standard it actually applied. If, therefore, the immunity from federal regulation were removed and the com-

754. CONN. GEN. STAT. ANN. § 38-201d (West 1979).

755. *Id.* § 38-201f.

756. 360 Mass. 1, 271 N.E.2d 592 (1971). *See supra* notes 94-117 and accompanying text.

757. 271 N.E.2d at 601.

merce power fully asserted, courts reviewing the constitutionality of state insurance laws might measure them on a more finely calibrated scale.

Federal regulation flowing from the comerce power may possibly be constricted by the recent landmark case of *Fullilove v. Klutznick*.[758] In that case, the United States Supreme Court, wary of the implications of *National League of Cities v. Usery*,[759] observed that "[i]n certain contexts, there are limitations on the reach of the Commerce Power to regulate the actions of state and local governments. . . . To avoid such complications, we look to § 5 of the Fourteenth Amendment."[760] Section 5 is the "necessary and proper" clause which authorizes Congress to protect privileges and immunities, due process and equal protection guarantees against state action.[761] If Congress elects to utilize its section 5 power to control state legislative or administrative regulation of automobile insurance that it deems in violation of the fourteenth amendment, provisions of some no fault laws may be found unacceptable by national standards of consumer protection.

B. Settled Issues

Assuming that federal intervention does not occur and that state control over insurance continues in its present pattern, several issues have been settled by the major no fault decisions. Distinctive state constitutional provisions must be factored into any determination of which challenges are no longer practical to assert.

For example, implied consent to no fault coverage is not worth litigating in most jurisdictions. But in Kentucky, a specific constitutional provision that bars legislative abridgement of recoveries for injuries to persons or property provided the basis for a challenge to Kentucky's no fault law.[762] Illinois' constitutional prohibition of special legislation, when strictly construed by the Illinois Supreme Court, proved fatal to the classifications contained in Illinois' no fault law.[763] Thus, the viability of future constitutional attacks on no fault legislation may depend on the existence of un-

---

758. 448 U.S. 136 (1980).

759. 426 U.S. 833 (1976).

760. 448 U.S. at 143.

761. "The Congress shall have power to enforce, by appropriate legislation, the provisions of this article." U.S. CONST. amend. XIV, § 5.

762. *See supra* notes 299-309 and accompanying text.

763. Grace v. Howlett, 51 Ill. 2d 478, 283 N.E.2d 474 (1972). *See supra* notes 62-77 and accompanying text.

usual state constitutional provisions. Generally, however, state courts have sustained no fault legislation, applying a traditional rationality review.

Indeed, it does not seem fruitful to contest the major classifications contained in no fault laws. A line must be drawn somewhere between those persons, vehicles and events which are covered and those which are not. Almost always, parties can allege over- or under-inclusiveness, but it is well settled that courts will, barring clear arbitrariness, honor legislative delineations.

Challenges to curtailment of the right to jury trial have not succeeded and probably will not in the future, given the precedent of workers' compensation decisions. Some state courts, however, have reached a contrary conclusion where the right to jury trial is eliminated or diminished by the requirement of arbitration for claims not exceeding a certain monetary limit. The Illinois Supreme Court ruled the arbitration provision of the Illinois no fault statute unconstitutional on the ground that the arbitrator's award had the effect of a final judgment.[764] Similarly, the advisory opinion of New Hampshire's highest court stated that the proposed no fault statute would violate the right to jury trial because the party demanding a jury trial would be required to pay all accrued costs.[765]

It is also settled that state legislatures may modify or abrogate provisions of the common law despite right of access to courts and right of redress protections in their constitutions. A caveat must be issued, however, where the state court adopts the corollary that a reasonable alternative or substitute must be provided for common law rights that have been diminished or eliminated. The Florida Supreme Court has in the past taken a firm position on this issue.[766] Other states have considered and rejected the alternative remedy requirement, perhaps with some hesitation. In *Pinnick*, the Supreme Judicial Court of Massachusetts engaged in a "measure for measure" analysis of the gains and losses resulting from adoption of a no fault system,[767] out of deference to *New York Central Railroad v. White*.[768] The Connecticut Supreme Court in *Gentile*

764. *Id.*

765. Opinion of the Justices, 113 N.H. 205, 214, 304 A.2d 881, 887 (1973). *See supra* notes 79-83 and accompanying text.

766. *See supra* notes 201-208 and accompanying text.

767. 271 N.E.2d at 605.

768. 243 U.S. 184, 201 (1917). *See supra* notes 18-24 and accompanying text.

v. Altermatt,[769] made a similar analysis, but it observed that a requirement of a reasonable alternative did not mean "an exact equation of remedies."[770]

C. *Explorable Issues*

Equal protection and due process guarantees remain the obvious touchstones of constitutional challenges to no fault legislation. An interesting question is raised by Florida's abrogation of the collateral source rule for tort actions arising out of the operation of a motor vehicle.[771] Automobile accident plaintiffs may be denied equal protection if the rule remains operative for other tort actions. Judicial review may sustain the law, however, by analogizing it to other no fault limitations of tort recovery, as a rational means to avoid duplication of benefits.

Specific offsets of workers' compensation payments, disability benefits and Medicare have been tested and affirmed in Michigan.[772] In reaching those conclusions, however, the Michigan Supreme Court undertook a detailed analysis of the inequities in the consequences to recipients of governmental as opposed to private benefits, and of the facial incompatibility of certain provisions of state disability and no fault laws. That analysis emphasizes the necessity of examining offsets in the context of a state's statutory scheme of benefits. When a new system, such as no fault, is "patched in," existing benefits legislation may require subsequent adjustments to achieve coordination and integration. The Michigan experience in litigiating specific equal protection problems provides an example for testing the validity of various offsets. Particularly as the range of offsets increases in a mandatory no fault scheme, the first party beneficiary may wish to question the fairness, rationality and reconcilability of various deductions.

Another arguable issue is illustrated by the New Jersey exclusion from work loss benefits of those who are not "income producers" at the time of the accident.[773] The New Jersey Supreme Court in *Rybeck* did not rule on the constitutionality of denying work loss to unemployed persons.[774] At least for individuals recently or seasonally unemployed there appears to be a valid equal protection

769. 169 Conn. 267, 363 A.2d 1 (1975).
770. 363 A.2d at 15. *See supra* notes 268-279 and accompanying text.
771. *See supra* notes 154-155 and accompanying text.
772. *See supra* notes 722-742 and accompanying text.
773. *See supra* notes 412-415 and accompanying text.
774. *See* 358 A.2d at 832.

or due process argument that, having paid premiums for no fault insurance, they are entitled to some measure of benefits.

Monetary thresholds remain an anomaly. States that still cling to this debatable preventative of contrived injuries apparently do so without firm convictions. Most thresholds are quite low, communicating a lack of legislative enthusiasm for severe limitations on tort actions. To the extent that monetary thresholds persist, however, it can be questioned whether a figure that exludes all hospital costs, or a figure that is uniform throughout a state where regional costs of medical treatment vary substantially, is arbitrary and unfair.

D. The "New" Due Process

Beyond these questions targeting specific provisions of state statutes there remains an unsettled challenge of a different order of magnitude: to test government regulation of no fault insurance on consumer protection/due process grounds as did the Michigan Supreme Court in *Shavers*. Although it selected a rationality standard of constitutional review, the *Shavers* court did not exhibit the same degree of tolerance for the personal injury and the property damage provisions of the statute.

In upholding the mandatory property damage insurance, which affords no protection to the insured's vehicle but rather compensates for damage caused by the insured to the tangible property or properly parked vehicle of a third party, the *Shavers* court credited possible future improvements in the reparations system as justifying goals.[775] Where compulsory *personal injury* no fault insurance was concerned, however, the Michigan Supreme Court supplemented its rationality review by invoking the entitlement doctrine. The result of that analysis was a more sophisticated and demanding due process scrutiny. The *Shavers* court determined that the captive insurance consumer was entitled to fairness in the determination of rates and policy terms, to be informed of underwriting and rating factors affecting availability and pricing of insurance, to more responsive conduct on the part of insurers and to more effective control by administrative agencies. Thus, the entire regulatory mechanism as well as industry practices fell under judicial due process scrutiny. Accordingly, the court instructed the legislature to revise the no fault system or suffer invalidation of its

775. 267 N.W.2d at 102.

keystone provisions.

Application of the exacting *Shavers* standard to the automobile insurance schemes in other jurisdictions would certainly disclose flaws, although some statutes attempt to accord consumers protection.[776] The deterrent to raising a *Shavers*-type challenge to an entire state no fault system is the time and effort in constructing such a case. Unless a court is willing, as was the Michigan Supreme Court, to take judicial notice of a wide variety of long accepted industry practices, the practical difficulty of attacking the system per se may be insurmountable. Most importantly, for a *Shavers*-type challenge to succeed, a court must approach its review with a conviction that the consumer is entitled to accountability and fairness from insurers and administrators alike.

The reasoning of the Michigan Supreme Court suggests the innovative views expounded by Professors Reich and Van Alstyne in their writings of some years ago.[777] Professor Reich described the forms of government created wealth—income benefits, licenses, subsidies and jobs—and argued that these represented new forms of property requiring restraints against arbitrary official action. Professor Van Alstyne emphasized the need for fairness and for protection against "grossness" in administrative actions. The *Shavers* decision responds to these concerns. It may also presage a "new" due process which other jurisdictons may be willing to advance.

776. *See, e.g.*, CONN. GEN. STAT. ANN. §§ 38-343 to -345 (West 1979) (requiring insurers to provide insureds with information concerning rates, and affording an aggrieved insured the opportunity for a hearing before the Commissioner).

777. Reich, *The New Property*, 73 YALE L.J. 733 (1964); Van Alstyne, *Cracks in "The New Property": Adjudicative Due Process in the Administrative State*, 62 CORNELL L. REV. 445 (1977).

CP National Corp. v. Public Service Commission: The Jurisdictional Ambiguity Surrounding Municipal Power Systems

Municipally owned utilities throughout the country have often experienced strained relations with other utility systems. In Utah, those tensions have been exacerbated by two recent court decisions and an administrative determination. In *CP National Corp. v. Public Service Commission*,[1] the Utah Supreme Court upheld the dismissal of an action to condemn an investor-owned power system brought by a group of Southern Utah municipalities. The Utah Public Service Commission ("PSC") previously had approved the sale of the system to the Utah Power and Light Co. ("UP&L"), despite efforts by the municipalities to buy the system.[2] In a related dispute, *CP National Corp. v. City of St. George*,[3] the Federal District Court for the District of Utah enjoined the city of St. George from extending its electrical system to a recently annexed portion of the city which was being served by an investor-owned utility.

Those decisions emphasized the uncertain status of municipal utilities within Utah's utility regulatory scheme. The issues specifically considered were whether a municipal utility may expand services beyond corporate limits, whether the PSC may exercise jurisdiction over a municipally owned system and whether a municipally owned utility may extend services to an area annexed by the municipality although the area is served by a PSC certificated utility. Resolution of those issues is critical if recurring problems between municipal utilities and PSC certificated utilities are to be avoided as Utah cities and towns expand to accommodate growing populations.

]1. 638 P.2d 519 (Utah 1981).

2. Application of CP Nat'l Corp. and Utah Power & Light Co. for the Sale and Purchase of the Pub. Util. Elec. Business of CP Nat'l for Serv. in Washington, Iron and Kane Counties, Nos. 80-023-01, 80-035-02 (Utah P.S.C. June 4, 1981) [hereinafter cited as *PSC Sale Approval*].

3. No. C81-0182J (D. Utah Apr. 10, 1981) (order granting preliminary injunction).

I. BACKGROUND

A. *Electrical Utility Regulations and Municipal Power*

1. Development of Electricial Utility Regulations—Electricity became a common commodity during the late 1800's.[4] Originally, power systems were generally privately owned and operated. The limited generation and transmission capacity of those early systems permitted many suppliers to serve a single urban area[5] and made serving rural areas difficult.[6] As a result, rural municipalities, through tax appropriations or bond issues, developed publicly owned systems to provide power to town residents.[7] As technology increased electrical transmission and distribution capabilities, a single system could supply increasingly larger areas. The continued development of larger, more economical power systems resulted in the widespread consolidation of power companies.[8]

Public utilities were soon recognized as "natural" monopolies: capital intensive, service oriented industries capable of reducing average costs by increasing the units generated.[9] Competition between electric utilities involved a wasteful duplication of equipment that threatened the financial soundness of all competitors.[10] To retain the economic efficiencies inherent in a natural monopoly, yet avoid potential abuses,[11] government regulation was adopted as a substitute for market competition.[12] Regulation was intended to protect consumers from extortionate rates where competition was impractical, yet to permit suppliers a legitimate return on invested

---

4. The first electrical system was a street lighting system installed by Thomas Edison in New York City in 1882. R. HELLMAN, GOVERNMENT COMPETITION IN THE ELECTRIC UTILITY INDUSTRY 8 (1972).

5. *See* W. JONES, REGULATED INDUSTRIES 12 (1976).

6. *See* M. FARRIS & R. SAMPSON, PUBLIC UTILITIES 270-71 (1973).

7. E. VENNARD, GOVERNMENT IN THE POWER BUSINESS 29-30 (1968).

8. For example, during the period from 1882 to 1905, the city of Chicago granted 29 franchises to utility companies, but by 1897 virtually all competition was eliminated. *See generally*, R. HELLMAN, *supra* note 4, at 8-10 (discussing the consolidation of the original private power companies).

9. Most of the costs incurred in providing utility services are fixed, that is, they do not vary with the amount of output. Thus, increasing production reduces unit costs. *See* L. WHITE & A. STRICKLAND, REGULATION: A CASE APPROACH 4-7 (1976).

10. *See* W. JONES, *supra* note 5, at 4 (quoting H. ADAMS, *Relation of the State to Industrial Action* (1887), in TWO ESSAYS BY HENRY CARTER ADAMS (Dorfman ed. 1954)).

11. If a monopolistic firm is left unregulated, it will set rates to maximize profits, restrained only by consumer demand. *See* Demetz, *Why Regulate Utilities?*, 11 J. LAW & ECON. 55, 56 (1968).

12. *See* R. HELLMAN, *supra* note 4, at 10-11.

capital.[13] Government regulation also provided a means of furthering the public interest in supplying adequate utility services at reasonable rates to all individuals.[14]

2. Municipal and Private Power Today—Municipally owned power systems were unable to take advantage of the technological developments that increased the efficiency and capacity of electrical generating systems; therefore, they struggled to survive among the rapidly expanding private systems.[15] Many municipal utilities withdrew from providing power, while others took advantage of the economic efficiency of large investor-owned utilities by purchasing power at wholesale prices from those systems.[16] Approximately 2200 local publicly owned power systems presently serve 13.5% of the country's power customers.[17] An additional ten percent of the nation's power consumers are served by rural electrical cooperatives,[18] and the remaining seventy percent are served by investor-owned systems.[19]

Proponents of municipal power contend that local government ownership provides customers a more direct voice in utility management decisions than that afforded customers served by inves-

---

13. *See* 1 A. PRIEST, PRINCIPLES OF PUBLIC UTILITY REGULATION 3-4 (1969). Henry Adams observed: "The control of the state over industries should be coextensive with the application of the law of increasing returns in industries." H. ADAMS, *supra* note 10 at 4.

14. The Supreme Court stated:

> Property does become clothed with a public interest when used in a manner to make it of public consequence, and affect the community at large. When, therefore, one devotes his property to a use in which the public has an interest, he, in effect, grants to the public an interest in that use, and must submit to be controlled by the public for the common good, to the extent of the interest he has thus created. Munn v. Illinois, 94 U.S. 113, 126 (1876).

The Court in *Munn* upheld a state statute that fixed maximum rates for the storage of grain by public warehouses. *Munn* was one of the first major challenges to the power of government to regulate specific aspects of business.

15. E. VENNARD, *supra* note 7, at 31-33.

16. *Id.*

17. Bergman, *U.S. Electric Utility Statistics*, PUB. POWER, Sept.-Oct. 1981 at 65. Approximately three-fourths of the publicly held systems serve communities with populations of 10,000 or less. AMERICAN PUBLIC POWER ASS'N, THE PUBLIC BENEFITS OF PUBLIC POWER 5 (1981). However, many large cities, including Los Angeles, Seattle, San Antonio and Memphis, have municipal power systems. *Id.*

18. Bergman, *supra* note 17, at 65. Because the PSC has jurisdiction over rural electrical associations in Utah, *see* UTAH CODE ANN. §§ 54-2-1(20), (30) (1974), the jurisdictional issues discussed in this Comment will affect rural electrical associations as though they were investor-owned utilities. See *infra* notes 42-60 and accompanying text. For an overview of rural electrical association regulation, see Note, *Regulation of Rural Electrical Cooperatives*, 1966 UTAH L. REV. 102.

19. Bergman, *supra* note 17, at 65.

tor-owned systems.[20] Municipal power advocates further argue that electrical services meet an essential community need warranting local control, that the economic benefits derived from operating a municipal system directly enhance the local community and economy[21] and that most municipal systems provide power at lower rates than those charged by investor-owned utilities.[22] The rate advantage can be attributed to several factors. First, research has shown that publicly owned utilities have lower unit costs than investor-owned systems.[23] Second, municipal utilities are often financed through less expensive tax-exempt bonds.[24] Third, publicly held utilities, as non-profit institutions, are exempt from income taxation.[25] Finally, the federal government has given municipalities a preferred status over investor-owned utilities for purposes of allocating inexpensive wholesale power from federally owned facilities.[26]

Opponents of municipal power have disputed the benefits derived from local government ownership by contending that municipal power management is sometimes influenced by political rather than economic factors, to the detriment of city residents.[27] They also note that investor-owned utilities contribute substantially to local tax revenues while the tax-free status of municipal utilities shifts the tax burden to other taxpayers[28] and that investor-owned

---

20. AMERICAN PUBLIC POWER ASS'N, *supra* note 17, at 7.

21. *See* Nazaruk, *Sioux Center's Sages*, PUB. POWER, Jan.-Feb. 1982 at 16-17 (benefits include charitable contributions, improved city services and industrial development).

22. AMERICAN PUBLIC POWER ASS'N, *supra* note 17, at 8-9. Salt Lake City residents receive power from Utah Power & Light Company, an investor-owned utility, and are billed 7.7 cents for each kilowatt hour. UTAH POWER & LIGHT COMPANY, ELECTRIC SERVICE SCHEDULE No. 1 (Nov. 1, 1982). In comparison, Murray City, a nearby city that has municipal power, charges residential customers approximately 5.8 cents for each kilowatt hour and a three dollar service charge. Murray, Utah, Ordinance 761 (Oct. 1, 1982). Bountiful, a city north of Salt Lake City that also has municipal power, charges residential customers approximately five cents for the first 200 kilowatt hours and three cents for each additional hour. BOUNTIFUL CITY LIGHT AND POWER, SCHEDULE No. 1 (July 1, 1980).

23. AMERICAN PUBLIC POWER ASS'N, *supra* note 17, at 10. Locally owned power systems had lower unit costs in production, distribution, accounting and collecting, sales and administration. *Id.*

24. *Id.* at 9; E. VENNARD, *supra* note 7, at 40-41.

25. E. VENNARD, *supra* note 7, at 40-41.

26. Reclamation Project Act of 1939, § 9(c), 43 U.S.C. § 485h(c) (1976). The intent of Congress was to break investor-owned utility control over power development. *See* Fereday, *The Meaning of the Preference Clause in Hydroelectric Power Allocation Under the Federal Reclamation Statutes*, 9 ENVTL. L. 601 *passim* (1979).

27. *See* M. FARRIS & R. SAMPSON, *supra* note 6, at 275-76.

28. Manning, *Should Government Operate Utility Services?*, 71 PUB. UTIL. FORT. 28, 31-32 (1963).

utilities provide a sound, low risk investment for private capital.[29]

Today, investor-owned public utilities are regulated primarily through state utility commissions.[30] Before providing services to a particular area, a public utility must obtain certification from the state commission. This certification process assures that adequate service is provided throughout the state without needless duplication of equipment.[31] State utility commissions also determine the reasonableness of rate adjustments, assure that service provided to the certificated area is nondiscriminatory, monitor accounting procedures and oversee property and equipment management.[32]

3. Public Utility Commissions and Regulation of Municipal Utilities—Authority for municipalities to provide utility services is generally found in statutory grants,[33] or sometimes in state constitutions.[34] In some states, municipal utilities are authorized to serve customers living beyond corporate limits.[35] In others, geographic or quantitative limits are imposed on the extraterritorial services of municipal utilities.[36]

Similarly, state utility commission jurisdiction over municipal

---

29. Sporn, *Observations on Private Versus Public Power*, 53 PUB. UTIL. FORT. 717, 717-18, 720 (1954).

30. See W. JONES, *supra* note 5, at 39-41. Initial regulatory efforts consisted of local governments issuing franchises to utility companies that desired to serve a particular area. Political corruption, coupled with the inability of city councils to handle regulatory issues, proved this method ineffective. See *id.* at 25-30; L. WEISS & A. STRICKLAND, *supra* note 9, at 7. Subsequently, state commissions, which were initially created to regulate railroads, were given expanded powers to regulate other utility systems such as electric utilities. The Massachusetts Board of Railroad Commissioners, established in 1869, was the first state utility commission. Initially, state commissions assumed only an advisory role, but as the need to regulate natural monopolies increased, commissions were given regulatory jurisdiction. See W. JONES, *supra* note 5, at 31, 39-44. For an overview of each of the fifty state utility commissions, see *Public Utility Regulation in Texas—A Symposium*, 28 BAYLOR L. REV. 771, 1157 app. (1976).

31. *See, e.g.*, Utah Gas Serv. Co. v. Mountain Fuel Supply Co., 18 Utah 2d 310, 422 P.2d 530, 532-33 (1967).

32. *See Public Utility Regulation in Texas—A Symposium, supra* note 30.

33. *See, e.g.*, COLO. REV. STAT. § 31-12-101(34) (1973); N.J. STAT. ANN. § 40:62-12 (West 1967); PA. STAT. ANN. tit. 53, § 23153 (Purdon 1957); TEX. REV. CIV. STAT. ANN. art. 1108 (Vernon 1963); WYO. STAT. § 15-7-101(v),(vi) (1977).

34. *See, e.g.*, ARIZ. CONST. art. 2, § 34, art. 13, § 5; CAL. CONST. art. 11, § 19; MICH. CONST. art. VII, § 24; OHIO CONST. art. XVIII, § 4.

35. *See, e.g.*, Crandall v. Town of Safford, 47 Ariz. 402, 56 P.2d 660 (1936); CAL. CONST. art. 11, § 19; NEB. REV. STAT. § 19-2701 (1977); N.J. STAT. ANN. § 40:62-21 (West 1967); OR. REV. STAT. § 225.030 (1981); WYO. STAT. § 15-7-201 (1977).

36. *See, e.g.*, MICH. CONST. art. VII, § 24 (allowed to sell 25% of product beyond limits); ALA. CODE § 11-50-1 (1975) (serve "surrounding territory"); ARK. STAT. ANN. § 73-264 (1974) (contiguous rural area); MINN. STAT. ANN. §§ 455.29, .32 (1963) (may sell surplus product only within 30 miles of corporate limits except fourth class cities which are not limited by distance).

and privately owned utilities varies from state to state. Some states provide statutory exemptions for municipally owned utilities from state regulatory commission jurisdiction,[37] others place municipal utilities within state commission jurisdiction,[38] and a third group grants the state commission jurisdiction over municipal utilities to the extent that services are rendered beyond corporate limits.[39] The rationale for granting limited jurisdiction over a municipality that provides utility services beyond its boundaries is that it is assuming the role of a utility normally regulated by the state.[40] Absent state commission jurisdiction, customers beyond corporate limits would be deprived of both political and regulatory recourse against the municipal utility.[41]

B. Utility Regulation in Utah

The state legislature established the Utah Public Service Commission to regulate "every public utility" providing service within the state.[42] All utilities within PSC jurisdiction must obtain a certificate of convenience and necessity before constructing or operating a utility system.[43] The PSC is authorized to approve rate adjustments,[44] resolve issues regarding property use and disposition,[45] determine asset valuation,[46] and approve system improvements and repairs.[47]

1. Municipal Utilities' Exemption from PSC Jurisdiction—The PSC's powers are limited by the Utah Constitution, which forbids the legislature to delegate "to any special commis-

---

37. *See, e.g.*, KAN. STAT. ANN. § 66-104 (1980); N.C. GEN. STAT. §§ 62-3(23)(d), 30 (1975); N.M. STAT. ANN. § 62-6-4 (Supp. 1981).

38. *See, e.g.*, N.Y. PUB. SERV. LAW § 66 (McKinney 1955 & Supp. 1981-82); WIS. STAT. ANN.§§ 196.01(1), .02(1) (West 1957 & Supp. 1981-82).

39. *See, e.g.*, City and County of Denver v. Public Utils. Comm'n, 101 Colo. 38, 507 P.2d 871 (1973); KAN. STAT. ANN. § 66-101, 104 (1980); N.J. STAT. ANN. § 40:62-24 (West 1967); PA. CONS. STAT. ANN. §§ 1102(a)(5), 1301, 1501 (Purdon 1979).

40. *See* City of Lamar v. Town of Wiley, 80 Colo. 18, 248 P. 1009, 1010 (1926).

41. *Id.*

42. UTAH CODE ANN. § 54-4-1 (Supp. 1981).

43. *Id.* § 54-4-25. For analyses of what constitutes "convenience and necessity," see Salt Lake & U.R.R. v. Public Serv. Comm'n, 106 Utah 403, 149 P.2d 647, 649 (1944); Utah Light & Traction Co. v. Public Serv. Comm'n, 101 Utah 99, 118 P.2d 683 *passim* (1941); Mulcahy v. Public Serv. Comm'n, 101 Utah 245, 117 P.2d 298, 299-301 (1941).

44. UTAH CODE ANN. § 54-3-1 to -3 (1974 & Supp. 1981).

45. *Id.* §§ 54-4-13, -30 (1974).

46. *Id.* §§ 54-4-21, -22.

47. *Id.* § 54-4-8.

sion . . . any power to make, supervise or interfere with any municipal improvement, money, property or effects."[48] The Utah Constitution also explicitly grants municipal corporations authority to "maintain or operate . . . public utilities local in extent and use."[49] The state legislature has codified the right of municipalities to provide utility services for both cities[50] and towns.[51]

The Utah Supreme Court consistently has interpreted those provisions as exempting municipal utilities from PSC jurisdiction. In *Logan City v. Public Utilities Commission*,[52] the Utah Supreme Court addressed the issue of whether the Public Utility Commission[53] could regulate electric rates charged by a municipal system. Holding that it could not, the court reasoned that the Utah constitutional provision forbidding legislative delegation to state commissions of control over municipal affairs[54] barred the Commission from exercising jurisdiction over municipal utilities.[55] The court stated: "[T]he legislature under the guise of police power may not . . . violate rights and privileges guaranteed or safeguarded by the Constitution."[56]

Municipal utilities in Utah also have been granted the right to sell "surplus" product beyond municipal limits.[57] In *County Water System, Inc. v. Salt Lake City*,[58] the Utah Supreme Court held that the sale of surplus water beyond corporate limits did not subject a city to PSC jurisdiction, concluding that the sale of surplus was consistent with normal municipal functions.[59] The court rejected the claim that absent PSC jurisdiction, municipal utilities could initiate destructive competition with regulated utilities beyond corporate limits, stating that the authority of municipalities to sell utility services beyond corporate boundaries is limited to the disposal of surplus.[60]

---

48. UTAH CONST. art. VI, § 29.
49. *Id.* art. XI, § 5(b).
50. UTAH CODE ANN. § 10-8-14 (1973).
51. *Id.* § 10-13-14, -14.5 (1973 & Supp. 1981).
52. 72 Utah 536, 271 P. 961 (1928).
53. A predecessor of the Public Service Commission. UTAH CODE ANN. § 54-1-2 (1974).
54. UTAH CONST. art. VI, § 29. *See supra* note 48 and accompanying text.
55. 72 Utah at 556, 271 P. at 972.
56. *Id.* at 561, 271 P. at 970.
57. UTAH CODE ANN. §§ 10-8-14, 10-13-14, -14.5 (1973 & Supp. 1981).
58. 3 Utah 2d 46, 278 P.2d 285 (1954).
59. *Id.* at 53, 278 P.2d at 289-90.
60. *Id.* at 52, 278 P.2d at 289 ("[cities] have no authority to sell water outside the city limits except . . . to sell surplus product").

2. Municipal Annexation and Acquisition—An unresolved issue in Utah's regulatory scheme is whether a municipality may extend utility services to an annexed area when a PSC certificated utility already serves that area. The Fifth District Court for Washington County, Utah, has held that the constitutional right of a municipality to provide utility services is "superior" to the PSC's authorization of a non-municipal utility to serve an area prior to annexation.[61] The court stated that the PSC certificated utility could continue to serve the annexed area only after receiving a franchise from the municipality to do so.[62]

Prior to the decision of the state district court, the municipality petitioned the PSC to enjoin the certificated utility from serving additional customers in the annexed area.[63] The PSC declined, stating that it did not have the authority to take that action.[64] The Utah Supreme Court upheld the PSC decision, reiterating that the law forbids PSC jurisdiction over municipal utility affairs.[65] The supreme court also noted that the certificated utility should not have its capital investment negated by municipal annexation without an opportunity to adjudicate its rights.[66] In the absence of PSC jurisdiction over municipal utilities, however, it is not certain in which forum the certificated utility's rights are to be adjudicated under Utah law.

The extent of the property condemnation and acquisition powers of municipal utilities is another inherent issue in the annexation dispute. The Utah Constitution allows municipalities "to acquire by condemnation, or otherwise, within or without the corporate limits, property necessary for [any public services], subject to restrictions imposed by general law."[67] The eminent domain statutes authorize condemnation of property for "public uses for the benefit of any county, city or incorporated town,"[68] and of "electric light and electric power lines, and sites for electric light

---

61. Dixie Rural Elec. Assoc. v. City of St. George, No. Civ. 5571, slip op. at 6 (Utah 5th Dist. Ct. July 22, 1977).

62. "The authority issued to the plaintiff, Dixie REA, by the [PSC], does not grant Dixie REA the exclusive right as against the City to serve customers which may be located in the city." *Id.*

63. City of St. George v. Public Serv. Comm'n, 565 P.2d 72, 73 (Utah 1977).

64. *Id.*

65. *Id.* at 73-74. *See supra* notes 48-55 and accompanying text.

66. 565 P.2d at 73.

67. UTAH CONST. art. XI, § 5(b).

68. UTAH CODE ANN. § 78-34-1(3) (Supp. 1981).

and power plants."[69] In addition, cities are authorized to "purchase, receive, hold, sell, lease, convey and dispose of property, real and personal, for the benefit of the city, both within and without its corporate boundaries."[70] Whether those acquisition and condemnation provisions permitted the acquisition of an ongoing electric system was uncertain prior to the decision in *CP National Corp. v. Public Service Commission.*[71]

II. CP National Corp. Sale and City of St. George Annexation

A. *Facts of the Controversy*

CP National Corporation (CPN), a California based investor-owned utility, had been certificated by the PSC to supply power to three southern Utah counties.[72] The CPN system was primarily a transmission and distribution system with access to generating facilities only through UP&L transmission lines.[73]

When CPN expressed a desire to sell its entire Utah system, several southern Utah municipalities formed the Southwest Power Agency (Agency) and attempted to purchase CPN's system.[74] The Agency was created pursuant to the Interlocal Cooperation Act, which allows municipalities to exercise collectively powers granted to each individually, yet prohibits the exercise of any greater power.[75] Of the eighteen municipalities included,[76] only five had existing power systems.[77] Those municipal power systems purchased power from investor-owned utilities such as UP&L or from federal generating facilities. None of the municipalities owned transmission lines, but instead each municipality relied on UP&L to transmit power to its distribution system.[78]

The Agency wanted to purchase the CPN transmission and

69. *Id.* § 73-34-1(8).
70. *Id.* § 10-8-2 (1973).
71. 638 P.2d 519 (Utah 1981).
72. *PSC Sale Approval, supra* note 2, at 4, 7.
73. *Id.* at 11.
74. *Id.* at 5-6.
75. Utah Code Ann. §§ 11-13-14 to -15 (Supp. 1981); CP Nat'l Corp. v. Public Serv. Comm'n, 638 P.2d 519, 521 (Utah 1981).
76. Cedar City, Brian Head, Enoch, Paragonah, Parowan, Kanab, Enterprise, Hurricane, Irvins, LaVerkin, Leeds, New Harmony, St. George, Santa Clara, Springdale, Toquerville, Virgin and Washington. *PSC Sale Approval, supra* note 2, at 14.
77. Enterprise, Hurricane, Paragonah, Parowan and St. George. *Id.*
78. *Id.* at 11.

distribution system to increase the capabilities of those municipalities presently providing power, and to enable others to establish municipal power systems.[79] The municipalities also wanted to develop power systems to be eligible to purchase inexpensive power from federal generating plants pursuant to the federal preference given municipal power systems.[80] Although parts of the CPN system were located beyond any municipal boundaries, the Agency wanted the entire system intact. Its plans included individual municipal control of the distribution system within the boundaries of each municipality, and joint control of the interconnecting transmission system to serve not only each municipality but also those residents in the unincorporated areas previously served by CPN.[81] However, negotiations between the Agency and CPN for sale of the system failed.[82]

UP&L, a PSC certificated utility, subsequently entered into a tentative agreement to purchase the CPN system. UP&L desired to acquire the system to enhance its existing Utah network.[83] Although UP&L allowed CPN to renegotiate with the Agency, again no agreement was reached.[84] A firm agreement was then made between UP&L and CPN and was submitted to the PSC for approval. The Agency subsequently initiated condemnation proceedings against CPN and UP&L to acquire the entire CPN system for the Agency's use. That action was dismissed by the trial court and appealed to the Utah Supreme Court.[85]

Pending resolution of the condemnation proceeding, St. George, one of the municipalities, annexed an area served by CPN. After efforts to purchase that portion of the CPN system failed, the city began running power lines parallel to CPN's to provide power to the annexed area through its municipal system.[86] CPN brought suit in federal district court seeking damages and injunctive relief against the city.[87]

---

79. *Id.*

80. *See supra* note 26 and accompanying text.

81. CP Nat'l Corp. v. Public Serv. Comm'n, 638 P.2d 519, 521 (Utah 1981).

82. *Id.*

83. *PSC Sale Approval, supra* note 2, at 6-7.

84. *Id.* at 12.

85. 638 P.2d at 520.

86. Memorandum in Support of Motion for Preliminary Injunction at 2-3, CP Nat'l Corp. v. City of St. George, No. C81-0182J (D. Utah Apr. 10, 1981) (order granting preliminary injunction).

87. CP Nat'l Corp. v. City of St. George, No. C81-0182J (D. Utah Apr. 10, 1981) (order granting preliminary injunction).

B. The Approval of the Sale to UP&L by the PSC

The PSC approved the sale from CPN to UP&L,[88] pursuant to a statutory requirement that the sale be "in the public interest."[89] The PSC's approval was given despite evidence offered by the Agency of its desire and ability to purchase the CPN system.[90] The Agency contended that because the respective municipalities had expressed a desire to operate their own system, the PSC "must yield to [that] public mandate" rather than allow an investor-owned utility to serve the area.[91] The PSC acknowledged the "substantial interest" of many citizens in southern Utah to have municipal power, yet the commission was reluctant to act on that basis alone because of its desire to remain neutral in the private versus public power debate.[92] The PSC resolved those conflicting pressures by adopting a "public interest" standard that required UP&L to show that the sale would result in "positive benefits" for the public.[93] The PSC found the economic and practical benefits that would result from UP&L's purchase and management of the system satisfied the positive benefits standard.[94] In contrast, the PSC expressed doubts about the Agency's ability to finance and manage the system.[95]

Although it approved the sale, the PSC ordered UP&L to grant an option to purchase the portion of the CPN system within its corporate boundaries to each municipality.[96] The option was intended to avoid duplication of facilities should any of the municipalities elect to develop its own power system.[97] The options were conditioned on the purchasing municipality's agreeing to acquire its own power supply in lieu of long term wholesale contracts with UP&L.[98] The PSC feared that the municipalities would otherwise purchase their systems and obtain power from UP&L at federally

88. *PSC Sale Approval, supra* note 2, at 31.

89. UTAH CODE ANN. § 54-4-30 (1974).

90. *PSC Sale Approval, supra* note 2, at 15.

91. *Id.* at 13-14.

92. *Id.* at 17. *See also supra* notes 20-29 and accompanying text.

93. *PSC Sale Approval, supra* note 2, at 17-18. The PSC rejected a standard that would have required a showing only that the sale would have no adverse impact because of the divisive controversy generated by the sale. *Id.*

94. *Id.* at 18-24. The factors enumerated by the PSC included lower rates than those charged by CPN, greater operating efficiency and local tax revenues. *Id.*

95. *Id.* at 24-27. The factors included uncertain sources of capital, no firm plan for operation and maintenance, threat of higher rates and uncertainty of power sources. *Id.*

96. *Id.* at 27-28.

97. *Id.* at 27, 30.

98. *Id.* at 29.

regulated wholesale rates applicable to municipal utilities.[99] Since those rates are lower than UP&L's rates for retail customers, the municipalities would have been in a position to force UP&L's retail customers to subsidize the municipalities' power costs.[100]

C. *The Utah Supreme Court Decision on Agency Condemnation*

In *CP National Corp. v. Public Service Commission*,[101] the Utah Supreme Court affirmed the district court's dismissal of the condemnation action, holding that the municipalities lacked statutory authority to condemn the CPN system.[102] The court construed the section of the eminent domain statute[103] as allowing condemnation of real property interests only.[104] The court then concluded that "[t]he taking of an ongoing public utility business is more than the taking of real or even tangible personal property."[105] Further, the court narrowly construed the provision in the eminent domain statute that authorizes the condemnation of "electric light and electric power lines, and sites for electric light and power plants"[106] as permitting condemnation only of "'lines and sites' for a power plant."[107]

The court found support for its decision in the statute that allows cities to sell surplus utility services beyond corporate limits.[108] The court read that statute as forbidding the Agency to "purposely engage" in the distribution of non-surplus power to

---

99. The Federal Power Commission was given the power to determine rates for interstate sales of electrical power. Public Utility Act of 1935, ch. 687, 49 Stat. 803, 850-51 (codified as amended in 16 U.S.C. §§ 824d-24e (1976 & Supp. V 1981)). That power was later transferred to the Federal Energy Regulatory Commission. Department of Energy Organization Act, Pub. L. No. 95-91, 91 Stat. 565, 583-84 (codified in 42 U.S.C. § 7172(a)(1)(B) (Supp. IV 1980)). The Federal Energy Regulatory Commission has promulgated regulations that govern rate determination. *See* 18 C.F.R. §§ 2, 35, 41 (1982).

100. *PSC Sale Approval*, *supra* note 2, at 30-31.

101. 638 P.2d 519, 524 (Utah 1981).

102. *Id.* at 524.

103. UTAH CODE ANN. § 78-31-1(3) (Supp. 1981).

104. 638 P.2d at 523. The Agency contended that Utah's eminent domain statute required only a showing that acquiring the CPN system would be for a "public use" benefiting municipal residents. *Id.* at 521-22. The Agency also cited the statute that authorizes cities to obtain property both within and without their boundaries. *Id.* at 522 (citing UTAH CODE ANN. § 10-8-2 (1973)).

105. 638 P.2d at 523.

106. UTAH CODE ANN. § 78-34-1(8) (Supp. 1981).

107. 638 P.2d at 524.

108. *Id.* (citing UTAH CODE ANN. § 10-8-14 (1973)). *See supra* notes 57-60 and accompanying text.

those living in unincorporated areas previously served by CPN.[109] Citing the case law that places municipal utilities beyond PSC jurisdiction, the court concluded that because "[c]ustomers who are non-residents of the municipality would be left at the mercy of officials over whom they have no control at the ballot box, and they could not turn to the [PSC] for relief," the Agency, therefore, could not extend its municipal power system to unincorporated areas by condemning the existing power system.[110]

D. The Granting of the Preliminary Injunction by the Federal District Court

In a related action filed during the pendency of the condemnation proceeding, *CP National Corp. v. City of St. George*,[111] CPN asked for injunctive relief from the Federal District Court of Utah, contending that the city of St. George was violating its property rights by running parallel lines through the annexed area of the city.[112] The city responded by stating that injunctive relief should not be given because CPN was unable to show a likelihood of prevailing on the merits.[113] The city contended that the certificate of convenience and necessity issued by the PSC did not grant CPN an exclusive right to provide power within the annexed area and thus the city's effort to provide power within the annexed area did not violate the rights as granted to CPN by the PSC.[114] Further, the city noted that it was acting pursuant to its constitutional authority to provide utility services to its residents.[115] Despite the city's contentions, a preliminary injunction was granted.[116] Further litigation on the merits became unnecessary because of CPN's subsequent sale to UP&L and the option granted thereby to St. George.[117]

---

109. 638 P.2d at 524.

110. *Id.*

111. No. C81-0182J (D. Utah Apr. 10, 1981) (order granting preliminary injunction). *See supra* notes 86-87 and accompanying text.

112. Memorandum in Support of Motion for Preliminary Injunction at 2, CP Nat'l Corp. v. City of St. George.

113. Defendant's Memorandum in Opposition to Plaintiff's Motion for Preliminary Injunction at 2, CP Nat'l Corp. v. City of St. George.

114. *Id.* at 3-5.

115. *Id. See supra* note 61 and accompanying text.

116. CP Nat'l Corp. v. City of St. George, No. C81-0182J (D. Utah Apr. 10, 1981) (order granting preliminary injunction).

117. *See* CP Nat'l Corp. v. City of St. George, No. C81-0182J (D. Utah Mar. 5, 1982) (stipulation and order for dismissal with prejudice). *See supra* notes 96-100 and accompanying text.

III. ANALYSIS

A. PSC Approval of the CPN Sale and Municipal Authority to Supply Services Beyond Corporate Limits

The PSC acted appropriately and pursuant to its statutory mandate in approving the sale to UP&L.[118] The Agency failed to acquire the CPN system because CPN rejected it as a buyer,[119] not because the PSC denied it the right to purchase the system.[120] Moreover, the PSC demonstrated sensitivity to the Agency's objectives by granting the option alternative.[121]

An underlying issue, not addressed by the PSC, but implicit in the Agency's effort to purchase the CPN system, was whether a Utah municipality could regularly supply extraterritorial utility services. Statutory and case law apparently limit the sale of utility services beyond municipal boundaries to the disposal of surplus.[122] The Agency's plans to provide services to unincorporated areas exposed an ambiguity in Utah's regulatory scheme. If the right to provide power were recognized, the extra-municipal customers would have no regulatory or political recourse against the municipal utility. Yet if the PSC denied the municipality authority to provide power, it would violate the constitutional provision forbidding state commission oversight of municipal affairs.[123] The option granted by the PSC struck a balance that appears to be consistent with the intent of state law. The municipalities could own the CPN system within their boundaries, but the facilities outside corporate limits would be owned by UP&L and regulated by the PSC. Thus, the authority of municipalities within their boundaries was respected, while the state regulation of utility services to people living in unincorporated areas was preserved.

The complexity of the PSC's compromise in the CPN case suggests that maintaining the strict jurisdictional separation be-

---

118. UTAH CODE ANN. § 54-4-30 (1974).

119. *PSC Sale Approval*, *supra* note 2, at 12.

120. A supporter of the Agency's effort to purchase the CPN system said, "The big question here is whether three commissioners appointed by the governor can say these towns are full of crap and that father knows best." Wall St. J., Jan. 13, 1981, at 35, col. 3.

121. The PSC stated: "[In requiring] UP&L to extend an opportunity to the municipalities for the purchase of their own distribution system, our decision has been influenced by the stated purpose of the mayors to acquire their own municipal electric distribution system. . . ." *Id.* at 30.

122. *See supra* notes 57-60 and accompanying text.

123. UTAH CONST. art. VI, § 29. *See supra* note 48 and accompanying text.

tween municipal and investor-owned utilities may unduly restrict the development of municipal utility systems. To allow municipalities to maximize the benefits of municipal power,[124] Utah law should be amended to authorize cities and towns to regularly provide utility services beyond corporate limits.[125] That change would allow municipal utilities to meet growing resident demand for electricity, increase efficiency, lower costs, and cooperate with other municipal systems to improve service for municipal residents, even if such efforts included serving customers beyond corporate limits.

Realization of the benefits and efficiencies of expanded municipal utility systems could be facilitated by adopting an alternative method of regulating municipal utilities. By granting the PSC jurisdiction over municipal utilities to the extent they serve areas beyond municipal boundaries,[126] customers beyond municipal limits would be assured access to a government entity to resolve disputes with the municipality.[127] Moreover, PSC jurisdiction ensures that utility facilities beyond corporate limits are not inefficiently duplicated.[128]

Adoption of such a regulatory scheme in Utah would require an amendment of the state constitutional provision forbidding state commission control over municipalities.[129] That provision should be altered to grant the PSC jurisdiction over municipal utilities to the extent that services are provided beyond corporate limits. That constitutional change would also remove the ambiguity inherent in the option agreement mandated by the PSC decision.[130] Under existing law, no specific statute grants the PSC the

---

124. *See supra* notes 20-26 and accompanying text.

125. *See supra* note 35 and accompanying text.

126. *See supra* note 39 and accompanying text.

127. *See* City of Lamar v. Town of Wiley, 80 Colo. 18, 248 P. 1009, 1010 (1926).

128. *See PSC Sale Approval, supra* note 2, at 10.

129. UTAH CONST. art. VI § 29. *But see* City of Lamar v. Town of Wiley, 80 Colo. 18, 248 P. 1009, 1010 (1926). In *City of Lamar*, the Colorado Supreme Court held that the Colorado Public Utilities Commission had jurisdiction over a municipally owned utility providing services to another city, despite a similar constitutional provision disallowing commission oversight of municipal utilities. *See also* City and County of Denver v. Public Utils. Comm'n, 181 Colo. 38, 507 P.2d 871, 872-74 (1973) (approving the *City of Lamar* holding). The Utah Supreme Court, however, has rejected the *City of Lamar* holding. County Water System, Inc. v. Salt Lake City, 3 Utah 2d 46, 278 P.2d 285, 289 (1954). The court's rejection relied on the fact that the court in *City of Lamar* did not address "the precise problem as to what effect should be given the constitutional provision prohibiting the delegation of control [of municipalities]." *Id.* at 51, 278 P.2d at 289. In addition, the court noted a Colorado case seemingly limiting the scope of *City of Lamar*. *Id.* (referring to City of Englewood v. City and County of Denver, 123 Colo. 290, 229 P.2d 667, 673 (1951)).

130. *See supra* note 123 and accompanying text.

right to order a certificated utility to divest itself of property interests.[131] The only authority for the divestiture order is the PSC's general jurisdictional grant that authorizes it "to do all things, whether herein specifically designated or in addition thereto, which are necessary or convenient in the exercise of such power and jurisdiction."[132]

The option agreement's requirement that municipalities eventually eliminate long term contracts with UP&L raises the additional question of whether the PSC can exercise jurisdiction over a municipality by limiting the city's access to power suppliers. The Utah Supreme Court has implicitly held that if the PSC granted an investor-owned utility the exclusive right to sell wholesale power to a municipal system, it would be considered an effort by the PSC to regulate the municipal system.[133] The PSC order requiring municipalities to eliminate long term contracts with UP&L could be viewed as an effort to regulate them by limiting access to wholesale suppliers. Adoption of the constitutional change proposed above, however, would resolve this question by giving the

---

131. *See PSC Sale Approval*, *supra* note 2, at 27: "As a general rule, we will not compel the sale of a utility company or of its properties nor will we undertake to establish ourselves as a super board of directors to make determinations with respect to the disposition of utility properties beyond the scope of applicable law."

132. UTAH CODE ANN. § 54-4-1 (Supp. 1981).

133. Utah Power & Light Co. v. Public Serv. Comm'n, 122 Utah 284, 249 P.2d 951 (1952). The court upheld a PSC order requiring UP&L to provide wholesale power to a municipal utility system pursuant to the municipality's request. The investor-owned utility, previously providing power to the municipal system, claimed that the order violated its right to provide power to that area. The court responded by stating that the PSC could not "grant to any public utility the area the city occupies as the exclusive territory of such utility, for to do so would be in effect to regulate the city by hampering its right to obtain power by any means or from anyone it desires." *Id.* at 289, 249 P.2d at 953.

An additional issue stemming from the UP&L options ordered by the PSC is whether the condition placed on the municipalities would be considered an effort by UP&L to eliminate municipal competition in violation of federal antitrust laws. The United States Supreme Court held that an investor-owned utility is not immune from antitrust violations when refusing to wholesale or transmit power to municipalities. Otter Tail Power Co. v. United States, 410 U.S. 366, 372-75 (1973). The Court, however, has articulated an exception to the applicability of antitrust laws. If the anti-competitive conduct is initiated by a state officer pursuant to legislation, antitrust laws may be inapplicable. Parker v. Brown, 317 U.S. 341 (1943). The PSC order to limit long-term contracts between UP&L and municipalities may fall within that state action rule, thus not subjecting UP&L to prosecution. *See* Gas Light Co. v. Georgia Power Co., 440 F.2d 1135, 1139-40 (5th Cir. 1971) (state utility commission approved rate schedule held to be within *Parker* doctrine); Washington Gas Light Co. v. Virginia Elec. & Power Co., 438 F.2d 238, 251-52 (4th Cir. 1971) (state utility commission acquiescence held to be within *Parker* doctrine). *But see* Canter v. Detroit Edison Co., 428 U.S. 579 (1976) (*Parker* doctrine not applied to state utility commission acquiescence to utility anti-competitive marketing scheme formulated by utility itself).

PSC limited jurisdiction over municipal utilities to allow such an order.

B. Municipal Condemnation of Existing Utility Facilities

The ability of municipal utilities to condemn existing utility facilities remains uncertain after the Utah Supreme Court's decision that the Agency could not condemn the CPN system. The court's decision limits municipal condemnation of an existing power system to the system's physical components. The court emphasized that the municipalities were not authorized to condemn CPN's supply and transmission contract rights.[134]

Even in its treatment of physical asset condemnation, the decision is ambiguous. The court's reading of the statutory provision allowing municipal condemnation of "electric light and electric power lines, and sites for electric light and power plants"[135] as allowing condemnation of "lines and sites for a power plant"[136] seemingly limits municipal condemnation of existing facilities to condemnation for the development of a power plant. Under the court's interpretation, it is unclear whether a municipality may condemn existing physical facilities when only developing a transmission and distribution system. The statutory language, although read narrowly by the court, appears broad enough to authorize municipal condemnation for that purpose.

Despite the narrowness and uncertainty of the court's decision, there is still another statutory obstacle that may preclude municipal condemnation of an existing utility's facilities. Utah's eminent domain statutes state that before property presently appropriated to a public use is condemned, the prospective use of the property must be shown to be "a more necessary public use."[137] The court specifically reserved the issue of whether the municipality's use of the CPN system was "a more necessary public use."[138] The Utah Supreme Court has previously stated, however, that property devoted to a public use may not be condemned to be used

---

134. CP Nat'l Corp. v. Public Serv. Comm'n, 638 P.2d 519, 522-23 (Utah 1981): The difficulty still remaining, however, is that here the municipalities want to take over an existing power system already constructed, maintained, and operating as an ongoing business. They do not simply seek to acquire property. . . . They seek to acquire virtually the entire CPN business including the ability to substitute themselves in CPN's place in its supply and transmission contracts with others.

135. UTAH CODE ANN. § 78-34-1(8) (Supp. 1981).

136. 638 P.2d at 524.

137. UTAH CODE ANN. §§ 78-34-3, -4 (1977 & Supp. 1981).

138. 638 P.2d at 524.

for the same purpose.[139] In a similar case, the Supreme Court of Oklahoma has held that absent specific statutory authorization, a municipality may not condemn an investor-owned power system to use it for the same purpose.[140] The Oklahoma court stated that a contrary holding would allow the municipality to destroy the rights of the investor-owned utility without benefiting the public.[141] Thus, unless Utah municipalities are able to show that the policies favoring municipal ownership of utilities constitute "a more necessary public use," condemnation of existing utility property interests may be unavailable under present eminent domain statutes.

C. Municipal Annexation of an Area Served by PSC Certificated Utility

Annexation by a Utah municipality providing utility services of an area presently served by a PSC certificated utility raises the difficult issue, absent an agreement between the parties, of determining which utility has the right to serve the annexed area. The constitutional provision authorizing municipalities to provide utility services for municipal residents[142] supports the view that the municipality ultimately has the right to serve the annexed area. It would be unjust, however, to ignore the certificated utility's rights and allow its capital investment in the area to be negated by municipal annexation.[143] Yet allowing the respective utilities to compete within the annexed area would not only undermine the previous investment of the certificated utility but would also lead to a wasteful duplication of facilities within the annexed area.[144]

---

139. Utah Copper Co. v. Stephen Hayes Estate, Inc., 83 Utah 545, 31 P.2d 624, 628-29 (1934).

140. City of Pryor Creek v. Public Serv. Comm'n, 536 P.2d 343, 346-47 (Okla. 1975). The Utah Supreme Court cited this case to support the proposition that a specific statutory authorization is needed to condemn an electrical utility system. 638 P.2d at 524. However, the *CP National* court did not emphasize the case's underlying issue of whether a municipality may condemn an existing utility system to use for the same purpose absent specific statutory authorization.

141. 536 P.2d at 346.

142. UTAH CONST. art. XI, § 5(b).

143. *See* Town of Culpepper v. Virginia Elec. & Power Co., 215 Va. 189, 207 S.E.2d 864, 866-69 (1974) (declaratory judgment in annexation dispute holding that utility commission certification was a property right entitled to court protection).

144. *See* CP Nat'l Corp. v. City of St. George, No. C81-0182J (D. Utah Apr. 10, 1981) (order granting preliminary injunction preventing city from competing within an annexed area). *But see* Union Rural Elec. Ass'n v. Town of Frederick, 629 P.2d 1093 (Colo. App. 1981); Caddo Elec. Coop. v. State *ex rel.* Whelan, 391 P.2d 234 (Okla. 1964) (limiting competition to pre-existing customers of certificated utility).

Municipal condemnation of the certificated utility's system within the annexed area is an alternative that might ensure adequate compensation for the certificated utility. The ambiguities concerning municipal condemnation authority resulting from the Utah Supreme Court's decision in the CPN case[145] and from the requirement that the municipality show that its use of the utility facilities would be "a more necessary public use"[146] may render condemnation a nonviable alternative.

The legislature should therefore act to protect the rights of the certificated utility when a portion of its service area is annexed by an adjacent municipality. One statutory alternative would be to grant municipalities explicit authority to condemn a PSC certificated utility's system within the annexed area.[147] If pursued, that alternative would ensure that the certificated utility receives adequate compensation for its investment in the annexed area. Municipal condemnation may not adequately protect the certificated utility if the condemnation renders useless its remaining system which lies beyond the new municipal limits. Similarly, if the municipality decides not to condemn the system, but rather extends its utility system into the annexed area in an effort to eliminate the certificated utility through competition, the certificated utility would be unjustly damaged and inefficient equipment duplication would result.

Rather than relying on the uncertainties of municipal condemnation to resolve the dispute, a preferable alternative would be to submit the conflict to PSC oversight. If the parties were unable to reach an agreement, the PSC could then determine a fair purchase price.[148] Additionally, if the PSC determined that the acquisition would render the certificated utility's remaining equipment useless, the PSC could require the municipality to acquire the *entire* certificated system in that area, even if the system extended beyond the annexed area. The proposed jurisdictional change sug-

---

145. 638 P.2d at 524.

146. UTAH CODE ANN. §§ 78-34-3, -4 (1977 & Supp. 1981). *See supra* notes 137-41 and accompanying text.

147. *See* IDAHO CODE § 61-333B (1976). Note also that a specific provision allowing condemnation of an existing utility system would likely satisfy the requirement that use by the municipality be a more necessary use. *See* North Salt Lake v. St. Joseph Water Co., 118 Utah 600, 223 P.2d 577, 582 (1950) (holding that specific statute that allowed condemnation of existing water utility system evidenced legislative intent that the municipality's use is more necessary).

148. *Cf.* MO. ANN. STAT. § 394.080(4) (allowing for PSC resolution if municipal utility and rural electrical association are unable to conclude on acquisition price).

gested earlier, which would allow municipal utilities to regularly render service beyond corporate limits subject to PSC regulation, would facilitate this alternative. Enactment of additional statutory provisions precluding the municipality from extending utility services into the annexed area *until* the certificated system has been acquired,[149] would eliminate the threat of destructive competition within the annexed area.

Enactment of statutes requiring municipal acquisition of existing electric systems and PSC oversight of the acquisitions would provide an efficient means of resolving the annexation problem. This approach would allow the municipalities to provide municipal utility services in annexed areas, avoid needless duplication of equipment and allow certificated utilities to recover the value of their systems.

IV. Conclusion

The CPN sale to UP&L and the annexation conflict between CPN and the city of St. George are illustrative of problems that can arise under Utah's present utility regulatory scheme. The present jurisdictional line between municipal utilities and the PSC can lead to seemingly unresolvable disputes. By allowing municipalities to regularly serve customers living beyond corporate limits subject to PSC jurisdiction, municipal utilities would be given needed leeway for expansion and consolidation beyond corporate limits. PSC supervision of municipalities providing extraterritorial service would prevent wasteful equipment duplication and would ensure that all customers served by a municipal system would have protection by a government entity. In addition, requiring municipal acquisition of a certificated utility system within an area annexed by the municipality, subject to PSC oversight, would facilitate efficient resolution of disputes between municipal entities and PSC certificated utilities which arise when a municipality expands its boundaries through annexation.

Kevin G. Glade

149. Arizona has a provision that precludes extension of services until the certificated system has been acquired. *See* Ariz. Rev. Stat. § 9-516 (1977).

Nordgren v. Mitchell: Indigent Paternity Defendants' Right to Counsel

An indigent's right to state-provided counsel has long been a subject of controversy.[1] Although it is well settled that an indigent criminal defendant cannot be incarcerated unless free counsel was available to him at trial,[2] the right to appointed counsel in non-criminal or "quasi-criminal" proceedings is severely limited and hazily defined.[3] It comes as no surprise that the right of indigent defendants to counsel in paternity actions is unsettled.[4]

A majority of American jurisdictions that have dealt with this issue now provide the putative father free counsel in paternity actions when he is unable to afford his own.[5] In *Nordgren v. Mitch-*

---

1. *See generally* Johnson & Schwartz, *Beyond* Payne: *The Case for a Legally Enforceable Right to Representation in Civil Cases for Indigent California Litigants* (pt.1), 11 LOY. L.A. L. REV. 249 (1978) (discussing various doctrinal grounds supporting extension of right to counsel of civil litigants); Note, *The Right to Counsel in Civil Litigation*, 66 COLUM. L. REV. 1322 (1966) [hereinafter cited as Note, *The Right to Counsel*] (examining factors favoring and factors imposing limits on the right to civil counsel); Note, *The Emerging Right of Legal Assistance for the Indigent in Civil Proceedings*, 9 U. MICH. J.L. REF. 554 (1976) (discussing possibility of granting right in civil cases where "fundamental interests" are at stake).

2. Scott v. Illinois, 440 U.S. 367, 373-74 (1979).

3. *See generally* Johnson & Schwartz, *supra* note 1, at 260-63 (discussing courts' vacillation between focus on meaningful opportunity to be heard and focus on significance of threatened deprivation).

4. *See generally* H. KRAUSE, ILLEGITIMACY: LAW & SOCIAL POLICY 109-12 (1971) (discussing whether paternity actions should be denominated "civil" or "criminal"); Note, *The Right to Appointed Counsel in Paternity Actions*, 19 J. FAM. L. 497 (1981) (compiling relevant U.S. Supreme Court and state court cases); Note, *Paternity—The Right of the Putative Father to Counsel in a Paternity Action*—Hepfel v. Bashaw, 279 N.W.2d 342 (Minn. 1979), 6 WM. MITCHELL L. REV. 208 (1980) (discussing reasoning behind *Hepfel* court's holding requiring appointment of counsel).

5. Nine jurisdictions now grant indigent paternity defendants a right to counsel by judicial decision. *See* Reynolds v. Kimmons, 569 P.2d 799 (Alaska 1977); Salas v. Cortez, 24 Cal. 3d 22, 593 P.2d 266, 154 Cal. Rptr. 529 (counsel provided if state is a party), *cert. denied*, 444 U.S. 900 (1979); Kennedy v. Wood, 439 N.E.2d 1367 (Ind. App. 1982); Artibee v. Cheboygan Circuit Judge, 397 Mich. 54, 243 N.W.2d 248 (1976); Hepfel v. Bashaw, 279 N.W.2d 341 (Minn. 1976) (counsel provided if state is a party); M. *ex rel.* T. v. S., 169 N.J. Super. 209, 404 A.2d 653 (1979); Wake County *ex rel.* Carrington v. Townes, 281 S.E.2d 765 (N.C. App. 1981); Corra v. Coll, 451 A.2d 480 (Pa. Super. 1982); State *ex rel.* Graves v. Daugherty, 226 S.E.2d 142 (W. Va. 1980). Seven other states statutorily grant indigent paternity defendants the right to counsel. ILL. ANN. STAT. ch. 40 § 1355 (Smith-Hurd 1980); MONT. CODE ANN. § 40-6-119 (1981); NEV. REV. STAT. § 126.201 (1979); N.Y. FAM. CT. ACT § 262(a)(viii) (McKinney Supp. 1981); N.D. CENT. CODE § 14-17-18 (Supp. 1977); WYO. STAT. § 14-2-116 (1977). *See* IND. CODE. ANN. § 34-1-1-3 (1973) (applies to civil actions generally).

ell,[6] however, the federal district court for Utah ruled that the minimum constitutional requirements of due process do not dictate appointment of counsel in all paternity cases.[7] Instead, the court indicated that state trial courts should determine, on a case-by-case basis, whether counsel should be appointed in paternity suits.[8] The *Nordgren* court's holding was based on a balancing of the private interests at stake in paternity suits, the public interests involved and the possibility of an erroneous outcome if counsel is denied, against a historical presumption against a right to counsel in noncriminal cases.[9] The court concluded that the trial court's balancing of these factors in individual paternity cases would lead to appointment of counsel in some but not all cases.[10]

I. Background: The Right to Counsel

The sixth amendment to the United States Constitution provides: "In all criminal prosecutions, the accused shall enjoy the right . . . to have the assistance of counsel for his defence."[11] For almost 150 years, the amendment was understood to mean only that an accused person who hired counsel on his own was entitled to have that counsel's advice at trial.[12] It was not until 1938, in *Johnson v. Zerbst*,[13] that the United States Supreme Court interpreted the sixth amendment to mean that no federal court could deprive a criminal defendant of life or liberty unless counsel was

---

See also Uniform Parentage Act § 19 (1979) (requiring right to counsel in paternity suits).

One state has adopted permissive language, *see* Hawaii Rev. Stat. §§ 584-16, -19 (1976), and five states have held that no right to counsel exists in paternity actions. *See* State *ex rel.* Hamilton v. Snodgrass, 325 N.W.2d 740 (Iowa 1982); Franks v. Mercer, 401 So. 2d 470 (La. Ct. App. 1981); State v. Green, 277 N.C. 188, 176 S.E.2d 756 (1970); Sheppard v. Mack, 68 Ohio App. 2d 95, 427 N.E.2d 522 (1980); State v. Walker, 87 Wash. 2d 443, 553 P.2d 1093 (1976).

6. 524 F. Supp. 242 (D. Utah 1981), *appeal docketed*, No. 81-2283 (10th Cir. Nov. 12, 1981).

7. *Id.* at 245.

8. *Id.*

9. *Id.* at 244-45.

10. *Id.*

11. U.S. Const. amend. VI.

12. In England, the right to counsel originally was recognized only in civil actions. Criminal defendants received counsel only if the judges thought it necessary in order to clarify the legal issues in the case. It was not until 1836 that an English statute granted a right of counsel to felony defendants. *See* An Act for Enabling Persons Indicted of Felony to Make Their Defense by Counsel or Attorney, 1836, 6 & 7 Will. 4, ch. 114; Note, *The Right to Counsel in Civil Litigation*, *supra* note 1, at 1328.

13. 304 U.S. 458 (1938).

provided at trial.[14] Contemporaneously, the Court gradually began extending the right to counsel to state criminal proceedings, under the due process clause of the fourteenth amendment.[15] Initially, in *Powell v. Alabama*,[16] the Court required state courts to provide counsel only in capital cases when the defendant otherwise was unable to obtain counsel or was unable, because of "ignorance, feeblemindedness, illiteracy, or the like," to conduct his own defense.[17]

The subsequent extension of the right to counsel in state criminal proceedings followed an uncertain course. In *Betts v. Brady*,[18] the Court adopted a case-by-case approach to determine whether indigent defendants accused of noncapital offenses were entitled to appointed counsel in state proceedings.[19] The Court refused, however, to interpret due process as requiring counsel in all criminal cases, reasoning that the right to counsel was not fundamental and essential to a fair trial.[20] In *Gideon v. Wainwright*,[21] decided in 1963, the Court rejected the case-by-case approach. Dismissing *Betts* as an "anachronism,"[22] the Court held that the right to counsel was indeed fundamental[23] and extended the right to counsel to all defendants "charged with crime" in state courts.[24] Finally, in *Scott v. Illinois*,[25] the Court established incarceration as the touchstone for determining the right to appointed counsel in state proceedings,[26] thus making the minimum state requirements under

---

14. *Id.* at 463.

15. Although arguments have also been made on equal protection grounds, *see* Powell v. Alabama, 287 U.S. 45, 50 (1932), the right is usually granted on due process grounds. *Id.* at 60; Gideon v. Wainwright, 372 U.S. 335 (1963). *Cf.* San Antonio Indep. School Dist. v. Rodriguez, 411 U.S. 1, 20 (1973) (stating that denial of a right because of inability to pay must be absolute to be constitutionally significant on equal protection grounds). Thus, equal protection requirements are satisfied by a minimally adequate hearing. *See* Note, *Indigent Prisoner Defendants' Rights in Civil Litigation*: Payne v. Superior Court, 90 HARV. L. REV. 1029, 1037 (1977). *But cf.* Griffin v. Illinois, 351 U.S. 12, 19 (1956) (granting criminal defendants a right to have a transcript of trial or its equivalent constitutionally required because "[t]here can be no equal justice where the kind of trial a man gets depends upon the amount of money he has").

16. 287 U.S. 45 (1932).

17. *Id.* at 71-73.

18. 316 U.S. 455 (1942).

19. *See id.* at 471-73.

20. *Id.* at 471.

21. 372 U.S. 335 (1963).

22. *Id.* at 345.

23. *Id.* at 342-43.

24. *Id.* at 344-45.

25. 440 U.S. 367 (1979).

26. *Id.* at 373-74.

due process essentially coextensive with the federal requirements.[27]

Although classification of a case as criminal has traditionally been viewed as a precondition to obtaining court-appointed counsel,[28] the Supreme Court emphasizes the nature of the threatened deprivation, rather than the nominal classification of the action, as the dispositive factor.[29] The Court established the standards for deciding when due process requires procedural protections such as the right to counsel in *Mathews v. Eldridge*.[30] Referring to the flexible nature of due process,[31] the *Mathews* Court stated that identification of the specific procedural requirements mandated by due process requires the balancing of three factors: the private interests at stake in the litigation, the risk of an erroneous deprivation of those interests in light of the value of any procedural safeguards, and the government's financial and administrative interests.[32] Today, the *Mathews* analysis remains the preferred test;[33] it provides the standard by which the existence and scope of the paternity defendant's rights must be judged.

II. THE *Nordgren* CASE

The indigent paternity defendant's right to counsel was considered by the Utah Federal District Court in *Nordgren v. Mitchell*.[34] Steven Nordgren, an indigent inmate at the Utah State Penitentiary,[35] was unable to obtain representation for his paternity defense in the state district court[36] from either the Legal Aid Soci-

---

27. *See supra* text accompanying note 14.

28. *See* Note, *The Right to Counsel*, *supra* note 1, at 1330.

29. For example, in *In re* Gault, 387 U.S. 1 (1967), the Court held that a defendant in a nominally civil juvenile court proceeding was entitled to counsel in a delinquency action resulting in commitment until age 21. *Id.* at 36-37. *See generally* Johnson & Schwartz, *supra* note 1, at 261-63 (discussing the nature of the deprivation resulting from civil lawsuits).

30. 424 U.S. 319 (1976). In *Mathews*, the plaintiff contended that due process required a full evidentiary hearing before the Social Security Administration could terminate his disability benefits. The Court held that the fiscal and administrative burden of a hearing would be disproportionate to the additional benefits it would provide. *Id.* at 347-48.

31. *Id.* at 334.

32. *Id.* 334-35.

33. *See* Lassiter v. Department of Social Servs., 452 U.S. 18 (1981); Little v. Streater, 452 U.S. 1 (1981).

34. 524 F. Supp. 242 (D. Utah 1981), *appeal docketed*, No. 81-2283 (10th Cir. Nov. 12, 1981).

35. *Id.* at 242. Nordgren actually brought the action in conjunction with three other plaintiffs in similar situations. *Id.*

36. The child's mother brought the suit under pressure from the state even though Nordgren was indigent and the child was ten years old when the suit was filed. Telephone

ety of Salt Lake or Utah Legal Services[37] and asked the court to appoint counsel in his behalf. When the court declined and the Utah Supreme Court denied review, Nordgren sued in federal district court seeking a declaratory judgment ordering the state to appoint counsel.[38] The district court refused.[39]

Employing the analysis of *Mathews v. Eldridge*,[40] the *Nordgren* court held that due process does not establish a right to court-appointed counsel for paternity defendants in all cases.[41] Instead, it held the Constitution to require, at minimum, that the state trial court determine the existence of that right in each case individually. The court went on to suggest that the state court employ the *Mathews* test, balancing the private interests at stake, the governmental interest affected and the risk that the procedures used will lead to an erroneous outcome against a historical presumption that an indigent litigant in a noncriminal case has no right to appointed counsel.[42] The court's analysis of those factors in the context of paternity suits relied heavily on two recent United States Supreme Court decisions—*Lassiter v. Department of Social Services*[43] and *Little v. Streater*.[44]

In *Lassiter*, the Court held that the right to appointed counsel in a parental status termination proceeding[45] should be determined on a case-by-case basis.[46] The Court specifically addressed each of

---

interview with Brian M. Barnard, attorney for plaintiffs (Sept. 30, 1982). In order to qualify for federal assistance, *see infra* note 100 and accompanying text, the state requires that paternity actions be instigated on behalf of all illegitimate minor welfare recipients and provides that the mother must cooperate as a condition of receiving welfare. Brief of Appellants at 9-10, Nordgren v. Mitchell, No. 81-2283 (10th Cir. filed Nov. 12, 1981).

37. 524 F. Supp. at 242-43. Both Legal Services and the Legal Aid Society asserted that they lacked the resources to handle prison cases and regularly refused to do so. Brief of Appellants, *supra* note 36, at 32.

38. 524 F. Supp. at 242-43.

39. *Id.* at 245.

40. 424 U.S. 319 (1976); *see supra* notes 29-32 and accompanying text.

41. 524 F. Supp. at 245.

42. *Id.* at 244-45.

43. 452 U.S. 18 (1981).

44. 452 U.S. 1 (1981).

45. In *Lassiter*, a North Carolina state court had found the defendant's son to be neglected and had given custody of him to the Durham County Department of Social Services. 452 U.S. at 20. A year later, in 1976, the defendant was convicted of second degree murder and began serving a sentence of 25 to 40 years in prison. *Id.* In 1978, the department petitioned the state court to have the defendant's parental rights terminated, which would allow foster parents to adopt the boy, alleging that the defendant had not contacted the department concerning the child since 1975. *Id.* at 20-21.

46. *Id.* at 31-32. The Court stated: "We . . . leave the decision whether due process calls for the appointment of counsel for indigent parents in termination proceedings to be

the *Mathews* factors, finding the parent's interest "extremely important" and the state's financial interest "relatively weak."[47] It found, however, that the risk of error was not insupportably high in all cases, given the state's procedural safeguards and the nature of the action.[48] The Court then introduced a fourth factor into the analysis—a historical presumption against a right to appointed counsel—and held that defendants threatened with termination of their parental rights have no absolute right to appointed counsel.[49] The Court observed, however, that the factors could, in quite limited cases, be so distributed as to overcome the historical presumption, and thus invited resort to the case-by-case approach.[50]

In *Little*, the Court addressed each *Mathews* factor in the context of a paternity action, holding that indigent paternity defendants were entitled to free blood grouping tests.[51] The Court

---

answered in the first instance by the trial court, subject, of course, to appellate review." *Id.* (citation omitted).

47. *Id.* at 31. The Court found the state's interest in an accurate and just determination to parallel that of the indigent parent, diverging only where financial considerations were concerned, and went on to state that the state's pecuniary interest, although legitimate, was "hardly significant enough to overcome private interests as important as those here." *Id.* at 28.

48. *Id.* at 28-31. The Court noted the procedures under North Carolina state law that are designed to assure accurate decisions, *id.* at 28-29, and the small likelihood that difficult points of evidentiary or substantive law would arise, but continued:

> Yet the ultimate issues with which a termination hearing deals are not always simple, however commonplace they may be. Expert medical and psychiatric testimony, which few parents are equipped to understand and fewer still to confute, is sometimes presented. The parents are likely to be people with little education, who have had uncommon difficulty in dealing with life, and who are, at the hearing, thrust into a distressing and disorienting situation.

Id. at 30. The Court thus conceded that the factors could, in some cases, combine to overwhelm the uncounseled parent, but concluded that the risk is not so high as to require appointment of counsel in all cases. *Id.* at 30-31.

49. *Id.* at 31-32.

50. *Id.* The Court stated:

> If, in a given case, the parent's interests were at their strongest, the state's interests were at their weakest, and the risks of error were at their peak, it could not be said that the *Eldridge* factors did not overcome the presumption against the right to appointed counsel, and that due process did not therefore require the appointment of counsel.

Id. at 31.

51. *Id.* at 16-17. Blood grouping tests compare various elements of the child's blood with those same elements in the blood of the putative father. *Id.* at 7. Depending on the correlation of those elements, the tests either exclude certain men absolutely from being the fathers or indicate the probability of a given man having fathered the child. For a discussion of those probabilities, *see infra* note 92. *See generally* Terasaki, *Resolution by HLA Testing of 1000 Paternity Cases not Excluded by ABO Testing*, 16 J. FAM. L. 543 (1978) (presenting data on the reliability of HLA testing).

equated the paternity defendant's interests with those of the parent in *Lassiter*,[52] noted the valuable evidentiary safeguard that blood tests represent[53] and again cited the relative insignificance of the state's financial interest.[54] The Court concluded that denial of blood tests to a defendant under an evidentiary burden to disprove paternity[55] denied him a meaningful opportunity to be heard.[56]

Although identifying *Little* as expounding the interests of the parties and their alignment in the litigation,[57] the *Nordgren* court found *Lassiter* to involve substantially the same balance of interests it faced and thus to be controlling.[58] The court noted that the state's pecuniary interest alone could not outweigh the defendant's interest in obtaining appointed counsel and that the procedural complexity of the action would sometimes make the risk of an erroneous determination insupportably high. For those reasons, and because the state court could appoint counsel if it determined that the trial was developing into one involving complicated legal or evidentiary problems, the court chose to allow the trial court to balance those interests on a case-by-case basis.[59]

III. Analysis: The Requirements of Due Process

A. *The Nature of Paternity Actions*

Although a few states characterize actions to establish paternity as criminal[60] or "quasi-criminal,"[61] most, including Utah,[62]

---

52. 452 U.S. at 13.

53. *Id.* at 7-8.

54. *Id.* at 15-16.

55. The Court noted that under Connecticut law, the putative father is charged with the burden of disproving paternity "if [the] mother or expectant mother continues constant in her accusation." *Id.* at 11. (Conn. Gen. Stat. § 46b-160 (1981)). The Court noted further that the father was required to disprove paternity by evidence other than his own testimony. *Id.* at 11-12. That burden, combined with the burden of bearing the cost of blood tests, denied due process. *Id.*

In most states, however, the mother or other complaining party has the burden of proving paternity. *See, e.g.*, Huntingdon v. Crowley, 64 Cal. 2d 647, 414 P.2d 382, 51 Cal. Rptr. 254 (1966); People *ex rel.* Abraham v. Bolen, 6 Ill. App. 3d 445, 284 N.E.2d 692 (1972); Edick v. Martin, 34 A.D.2d 1096, 312 N.Y.S.2d 427 (1970).

56. 452 U.S. at 16. *See infra* text accompanying notes 62-67.

57. 524 F. Supp. at 245.

58. *Id.*

59. *Id.*

60. *E.g.*, Commonwealth v. MacKenzie, 368 Mass. 613, 619 n.5, 334 N.E.2d 613, 616 n.5 (1975).

61. *E.g.*, Hunter v. State, 293 Ala. 226, 230, 301 So. 2d 541, 544-45 (1974); Artibee v. Cheboygan Circuit Judge, 397 Mich. 54, 57, 243 N.W.2d 248, 249 (1976). The Supreme Court of the United States has also referred to the "quasi-criminal" nature of paternity

characterize them as civil.[63] Because of that classification, it has been asserted that granting a right to counsel in these "civil" actions threatens "setting off an avalanche by moving one key stone,"[64] breaching the longstanding prohibition against appointing counsel in all civil cases.[65] This concern appears unfounded, however, because paternity actions, as will be shown below, are easily distinguished from other types of civil actions.

B. The Right of Access to the Courts

In *Boddie v. Connecticut*,[66] the United States Supreme Court stated that "absent a countervailing state interest of overriding significance, persons forced to settle their claims of right and duty through the judicial process must be granted a *meaningful* opportunity to be heard."[67] In *Boddie*, the Supreme Court held that the constitutional guarantee of free access to the courts has no meaning when states impose restrictions which, as a practical matter, deny that access to certain individuals.[68] The restriction in *Boddie* was a filing fee that denied indigents the opportunity to bring divorce actions.[69]

The relationship of *Boddie* to the paternity defendant's right to counsel is clear. For a paternity defendant's access to the courts to have meaning, he must not only have an opportunity to get into court, but must also have an opportunity to present evidence and arguments supporting his position. An indigent defendant, often of

---

actions. *See* Little v. Streater, 452 U.S. at 10.

62. State v. Judd, 27 Utah 2d 79, 81, 493 P.2d 604, 605 (1972).

63. *E.g.*, State v. Long, 12 Ariz. App. 170, 172, 468 P.2d 621, 623 (1970); Platt v. Ponder, 233 Ark. 682, 684, 346 S.W.2d 687, 689 (1961); People *ex rel.* Harris v. Williams, 8 Ill. App. 3d 821, 823-24, 291 N.E.2d 323, 325 (1972); Cohen v. Burns, 149 Ind. App. 604, 606, 274 N.E.2d 283, 284 (1971); Kimble v. Keefer, 11 Md. App. 48, 50, 272 A.2d 668, 668 (1971); Snay v. Snarr, 195 Neb. 375, 377, 238 N.W.2d 234, 235 (1976); *In re* Torino v. Cruz, 82 Misc. 2d 684, 369 N.Y.S.2d 291, 298 (Fam. Ct. 1975); State v. Bowen, 80 Wash. 2d 808, 810, 498 P.2d 877, 879 (1972); X. v. Y., 482 P.2d 688, 690 (Wyo. 1971). For a discussion of reasons for denominating paternity actions as civil, see H. KRAUSE, *supra* note 4, at 109-11. *See generally* Note, *The Nature of Paternity Actions*, 19 J. FAM. L. 475 (1981) (discussing the various interests at stake and procedures involved in paternity actions).

64. Artibee v. Cheboygan Circuit Judge, 397 Mich. 54, 62, 243 N.W.2d 248, 251 (1976)(Coleman, J., dissenting).

65. *See* Lassiter v. Department of Social Servs., 452 U.S. 18, 25 (1981).

66. 401 U.S. 371 (1971).

67. *Id.* at 377 (emphasis added).

68. *Id.* at 380-81.

69. *Id.* at 372. In determining the extent of due process, the Court was influenced by "the basic position of the marriage relationship in this society's hierarchy of values." *Id.* at 374. Certainly, the relationship of parent and child occupies an equally basic position.

limited education, is likely to be incapable of presenting his position persuasively when faced with overwhelming opposition[70] in a system of technical legal concepts. He is unlikely to be aware of his right to a free blood grouping test,[71] his right to object on the basis of the statute of limitations[72] or the technical requirements of trial that are clearly beyond the knowledge of the average layman. The indigent, therefore, lacks a meaningful opportunity to be heard. The presence of counsel, however, would guarantee that the defendant's rights are protected and that all relevant issues in the case are raised.

C. *Balancing the Interests Involved*

The due process analysis of *Mathews v. Eldridge*[73] requires a balancing of the interests of the private parties, the interests of the state and the danger of an erroneous outcome.[74] An analysis of those factors in the context of paternity actions indicates that the interests of the private parties clearly outweigh the state's pecuniary interest in denying counsel.

1. Private Interests—The private interests at stake in paternity litigation are commanding. For example, the alleged father has great social interests at stake. An erroneous determination that he has fathered an illegitimate child could have profound personal effects and could seriously endanger his relationship with his family.[75] Once the defendant has been adjudged to be the father, he is under an obligation to provide support.[76] Failure to provide that

---

70. *See infra* notes 100-103 and accompanying text.

71. Utah statutes require that blood grouping tests be granted to indigent paternity defendants free of charge. It is unclear, however, whether the defendant must request the test. The section of the judicial code dealing with evidence states that "the court *shall* order . . . blood tests," UTAH CODE ANN. § 78-25-18 (1977)(emphasis added), implying that the court has a duty to order tests. The section dealing specifically with paternity, however, states that "[t]he court, upon its own initiative . . . *may*, or upon the motion of any party to the action . . . *shall* order . . . blood tests," *id.* § 78-45a-10 (emphasis added), which suggests that the court is not required to order blood tests except on request of the parties.

72. In Utah, the father is liable only for support expenses occurring during the four years preceeding commencement of the action. *Id.* § 78-45a-3. Under former law, which conflicted with the above provision, the statute of limitations was four years from the child's date of birth. *Id.* § 77-60-15 (1978) (repealed 1980).

73. 424 U.S. 319 (1976).

74. *See supra* text accompanying notes 30-33 and 47-54. For cases applying the *Mathews* analysis, see Lassiter v. Department of Social Servs., 452 U.S. 18, 27 (1981); Little v. Streater, 452 U.S. 1, 13 (1981).

75. Salas v. Cortez, 24 Cal. 3d 22, 28, 593 P.2d 226, 230, 154 Cal. Rptr. 529, 533, *cert. denied*, 444 U.S. 900 (1979).

76. UTAH CODE ANN. § 78-45a-1 (1977). The obligation to provide support includes a

support is a crime to which stiff penalties attach.[77] Because the determination of paternity is res judicata in any later action to compel support payment,[78] the putative father's personal and financial[79] freedom are also at stake in a paternity action.

In addition to the father's interests, the child has correspondingly important interests that will be protected to the same degree. If, as is often stated, paternity actions have the best interests of the child as their purpose,[80] those interests are best served by an accurate paternity determination.[81] An erroneous determination that one individual is the father may preclude enforcement of the support obligation against the actual father, who may be better able to provide support.[82] Moreover, an inaccurate determination of paternity may create animosity in the alleged father toward the child, which could be emotionally damaging.[83] Thus, the private interests at stake in paternity litigation are of extreme importance.[84]

---

duty to provide for expenses of pregnancy and confinement of the mother, educational expenses of the child and funeral expenses of a deceased child. *Id.* § 78-45a-5.

77. A first conviction for failure to provide support is a Class A misdemeanor, *id.* § 76-7-201(2) (1978), punishable by imprisonment up to one year, *id.* § 76-3-204(1), and a $1000 fine, *id.* § 76-3-301(3). A second conviction is a third degree felony, *id.* § 76-7-201(3), punishable by imprisonment up to five years, *id.* § 76-3-203(3), and a $5000 fine, *id.* § 76-3-301(2). In nonsupport cases, however, the judge is given the option of fashioning a scheme of periodic payments in lieu of the criminal sanctions, *id.* § 76-7-202; thus, the threat of imprisonment is most useful in gaining leverage over the adjudged father.

78. Reynolds v. Kimmons, 569 P.2d 799, 801-02 (Alaska 1977).

79. In addition to the general support obligation, *see supra* note 76, a determination of paternity qualifies the child as a "child" for inheritance purposes. *See* UTAH CODE ANN. § 75-2-109(1)(b)(ii) (1978). Thus, in addition to his intestate share, *id.* § 75-2-103, the child is entitled to a family allowance, *id.* § 75-2-403, a homestead allowance, *id.* § 75-2-401, and a share of exempt property if there is no surviving spouse, *id.* § 75-2-402. He also is entitled to claim his intestate share as a pretermitted heir if not mentioned in the father's will. *Id.* § 75-2-302.

80. H. KRAUSE, *supra* note 4, at 113-15.

81. Salas v. Cortez, 24 Cal. 3d 22, 33, 593 P.2d 226, 233, 154 Cal. Rptr. 529, 536-37, *cert. denied*, 444 U.S. 900 (1979).

82. *See id.* at 33, 593 P.2d at 233, 154 Cal. Rptr. at 537.

83. *Cf.* H. KRAUSE, *supra* note 4, at 271 ("[there is] obvious value [in] fostering a home environment that is conducive to the proper social and personal development of a child").

84. In *Salas*, the court stated:

> "If the child is to have anything, it must have a *right* to have his paternity ascertained in a fair and efficient manner." It is in the child's interest not only to have it adjudicated that *some* man is his or her father and thus liable for support, but to have some assurance that the correct person has been so identified. When the state initiates paternity proceedings, whether on behalf of the mother . . . or the child . . ., the state owes it to the child to ensure that an accurate determination of parentage will be made.

24 Cal. 3d at 33-34, 593 P.2d at 234, 154 Cal. Rptr. at 537 (citation omitted, emphasis in

2. *Public Interests*—The state's interest, the second factor in the *Mathews* analysis, closely parallels the interests of the private parties, diverging only where financial considerations are involved.[85] Although the state has an economic interest in removing the child from its welfare rolls, that interest is legitimate only when the removal is for proper reasons. An inaccurate determination of paternity may accomplish the objective, but is hardly defensible. The state's interest, therefore, differs from that of the child and father only where the cost of providing counsel is concerned. The United States Supreme Court has consistently held that the state's pecuniary interest, although legitimate, is insignificant when compared to important private interests.[86] Clearly, the need to save money is secondary to the need to ensure protection of the alleged father and the child.[87]

An accurate outcome reached with the aid of appointed counsel may, in fact, further the state's financial interests. When the adjudged father feels that he has been treated fairly, he may be encouraged to voluntarily pay the required support—a result that bitterness is unlikely to accomplish. Voluntary support, in turn, will not only save the state the expense of supporting the child, but may also save it the cost of prosecuting the father for nonsupport.[88] Moreover, the state may save money by actually reducing the number of paternity cases that go to trial, thereby easing the caseload on the courts.[89] The presence of counsel may discourage mothers from pursuing unsound claims or may encourage the defendants to settle cases when victory at trial is improbable.

3. *The Danger of an Erroneous Outcome*—The third *Mathews* factor—the probability of an erroneous adjudication if the right in question is denied—was viewed by the *Nordgren* court as

---

original).

85. *Lassiter*, 452 U.S. at 27-28.

86. *Id. See also Mathews*, 424 U.S. at 348. *Cf.* Sherbert v. Verner, 374 U.S. 398, 406-07 (1963) (denial of unemployment benefits to an individual who refused to work on Sundays on religious grounds abridged first amendment rights).

87. An average non-jury paternity trial lasts less than two hours. Gliaudys, *Paternity: A Reluctant Fatherhood*, 53 CAL. ST. B.J. 318, 319 (1978). Based on experience with court-appointed counsel in other contexts, the cost of appointing counsel in indigent paternity cases is significant, but not extraordinarily so. *See* Note, *Dollars and Sense of an Expanded Right to Counsel*, 55 IOWA L. REV. 1249 (1970) (discussing cost of expanding the right to counsel in criminal cases).

88. Salas v. Cortez, 24 Cal. 3d at 33, 593 P.2d at 234, 154 Cal. Rptr. at 537; Wake County *ex rel.* Carrington v. Townes, 281 S.E.2d 765, 772 (N.C. App. 1981).

89. Salas v. Cortez, 24 Cal. 3d at 33, 593 P.2d at 234, 154 Cal. Rptr. at 537.

variable[90] and appears determinative of whether a right to counsel exists for paternity defendants. Because the private interests are large and the competing state interests small, only a low likelihood of error can justify denying these defendants the right to counsel.[91] As the *Nordgren* court recognized,[92] the most reliable evidence in determining paternity is the blood grouping test.[93] Utah provides free blood tests to defendants who request them.[94] Among those defendants who are not absolutely excluded by the tests,[95] however, the risk of error is insupportably high.

In the vast majority of states that classify paternity actions as civil,[96] the safeguards against erroneous outcomes available in criminal cases are not present. In criminal cases, the presumption of innocence and high burden of proof operate to protect the defendant.[97] The requirement that the prosecutor reveal evidence favoring the defendant, the careful screening of the case by personnel sworn to protect the defendant and the requirement that only the defendant can appeal[98] further safeguard the defendant's rights. Civil paternity actions do not have those protections; thus, the procedural safeguards against an erroneous determination of paternity are relatively low.

Most significantly, the alignment of the parties in paternity actions threatens to result in a high number of erroneous results. Although the mother can bring suit privately, paternity actions are currently most often brought by or in conjunction with the state[99]

---

90. *See* Nordgren v. Mitchell, 524 F. Supp 242, 245 (D. Utah 1981), *appeal docketed*, No. 81-2283 (10th Cir. Nov. 12, 1981).

91. *See supra* note 50 and accompanying text.

92. 524 F. Supp. at 245.

93. In one study, the complete battery of HLA blood grouping tests was given to 1000 paternity defendants. Of the 1000, 25% were definitely excluded. Of the remaining 750, 16% had a 99% probability of being the father, 67% had a 95% probability and 86% had a 90% probability. Terasaki, *supra* note 51, at 552-53.

94. The Utah statutory provisions conflict as to whether the defendant must request the test or whether the court is under an affirmative duty to order the tests in each case. *See supra* note 71. The United States Supreme Court recently held that due process requires granting blood tests to paternity defendants. *See* Little v. Streater, 452 U.S. 1 (1981).

95. That was the situation in *Nordgren*—blood tests were administered but were inconclusive. Telephone interview, *supra* note 36.

96. *See, e.g.*, State v. Long, 12 Ariz. App. 170, 172, 468 P.2d 621, 623 (1970); People *ex rel.* Jones v. Schmitt, 101 Ill. App. 2d 183, 186, 242 N.E.2d 275, 276 (1968); Cohen v. Burns, 149 Ind. App. 604, 606, 274 N.E.2d 283, 284 (1971); Snay v. Snarr, 195 Neb. 375, 377, 238 N.W.2d 234, 235 (1976); X. v. Y., 482 P.2d 688, 690 (Wyo. 1971).

97. Johnson & Schwartz, *supra* note 1, at 265-66.

98. *Id.*

99. In Utah, for example, the county attorney must act for the petitioner upon request, UTAH CODE ANN. § 77-31-12 (Supp. 1981), and the county may bring the suit itself if

because, under federal law, a state plan for child support cannot qualify for federal Aid to Families with Dependent Children unless that plan provides that the state will undertake to establish paternity.[100] When the state intervenes, the private resources of the mother and the state's massive financial and legal resources and investigative powers[101] may be pitted against an indigent defendant who is unlikely to understand the procedural complexities of a court action and who is unlikely to be excluded by blood grouping tests.[102] Clearly, the indigent defendant can be overwhelmed, with the result that the risk of an erroneous finding against the defendant is insupportably high.[103] That risk tips the *Mathews* balance conclusively in favor of providing counsel.[104]

D. The Case-by-Case Approach

Although the case-by-case approach may have some value when the state is not a party to the action,[105] its use when the state

---

it has furnished support to the child, *id.* §§ 77-31-8, 78-45a-2 (1977). That is apparently the common situation. *See* Note, *supra* note 63, at 479.

100. 42 U.S.C. § 654(4)(A) (1976).

101. Federal law authorizes appropriation of money to the states to carry out the function of establishing paternity, *id.* § 651, and requires that the state grant access to all "files and records maintained by any of the departments, agencies, or instrumentalities of the United States or of any State." *Id.* § 653(e)(1).

102. *See supra* note 93.

103. Compare Justice Blackmun's statement in the context of the right to counsel in parental termination proceedings:

> Faced . . . with an adversary—the State—that commands great investigative and prosecutorial resources, with standards that involve ill-defined notions of fault and adequate parenting, and with the inevitable tendency of a court to apply subjective values or to defer to the State's "expertise," the defendant parent plainly is outstripped if he or she is without the assistance of "the guiding hand of counsel."

Lassiter v. Department of Social Servs., 452 U.S. 18, 59 (1981) (Blackmun, J., dissenting) (citations omitted).

104. The incarceration of the defendant at the Utah State Penitentiary was an additional factor that weighed in favor of a right to counsel in *Nordgren*. *See* 524 F. Supp at 242. In addition to Nordgren's limited access to legal resources, he was unable to seek out witnesses or conduct investigations. *See generally* Payne v. Superior Court, 17 Cal. 3d 908, 923, 553 P.2d 565, 576, 132 Cal. Rptr. 405, 416 (1976) (requiring counsel for indigent prison inmate in civil damages action).

Because of the special problems faced by incarcerated defendants, commentators have argued for an extended right to counsel for them in most civil actions. *E.g.*, Note, *supra* note 15 at 1038-39. Further, the United States Supreme Court has given special consideration to the incarcerated status of defendants in other contexts. *See, e.g.*, Johnson v. Avery, 393 U.S. 483 (1969) (requiring assistance for illiterate inmates in preparing petitions for post conviction relief).

105. For example, in a non-state proceeding, the mother might choose to be unrepresented. In that kind of case, the risk of overwhelming the defendant would be low and the proceeding might remain somewhat informal, which would benefit all parties. Suits against

is a party wastes judicial resources. The state's hiring of an attorney who is likely a specialist in paternity actions[106] tips the *Mathews* balance so seriously against the defendant that no amount of mitigating circumstances can possibly compensate for the resulting imbalance. Because the case-by-case approach will, therefore, lead to the same result whenever the state is a party, it is inefficient and unproductive.[107]

The United States Supreme Court has frequently referred to the flexible nature of due process as justifying the case-by-case identification of the procedures due process requires,[108] but that flexibility requires case-by-case consideration of different constitutional contexts, not of different litigants.[109] Thus, the Court has considered different classes of litigants, not different litigants individually.[110] By applying the case-by-case approach to individual litigants, both the *Nordgren* and *Lassiter* courts have reverted to the discredited standard of *Betts v. Brady*.[111] Not only does an individualized case-by-case analysis make the application of due pro-

---

indigent defendants, however, are unlikely to be purely private actions, because the mother is unlikely to bring suit against someone who does not have the means to pay support.

106. One specialist in paternity has said that "[t]he deputy district attorney who handles child support matters within the context of paternity cases becomes . . . a specialist. To avoid imbalance in the system and preserve the defendant's rights, an informed and skillful defense attorney is essential." Gliaudys, *supra* note 87, at 324.

107. In Gagnon v. Scarpelli, 411 U.S. 778 (1973), the United States Supreme Court, despite its holding in Gideon v. Wainwright, 373 U.S. 335 (1963), *see supra* text accompanying notes 21-24, adopted a case-by-case approach to granting counsel in parole revocation hearings. 411 U.S. at 788. In addition to noting the absence of formal rules in the hearing and the presentation of facts to a qualified board of experts, the Court stressed that the state was not represented by counsel. Instead, an informal hearing was held between the defendant and his probation officer. *Id.* at 789. In paternity cases, the formality of the proceeding and the presence of the state's counsel mandates greater protection of the defendant's interests.

108. *See, e.g.*, Lassiter v. Department of Social Servs., 452 U.S. 18, 31 (1981).

109. *Id.* at 49 (Blackmun, J., dissenting).

110. *Id.* at 49-50. The Supreme Court observed:

> The Court's own precedents make this clear. In *Goldberg v. Kelly* [397 U.S. 254, 264 (1970)], the Court found that the desperate economic conditions experienced by welfare recipients *as a class* distinguished them from other recipients of governmental benefits . . . . In *Mathews v. Eldridge* [424 U.S. 319, 339-45 (1976)], the Court concluded that the needs of Social Security disability recipients were *not* of comparable urgency . . . . These cases established rules translating due process in the welfare context as requiring a pretermination hearing but dispensing with that requirement in the disability benefit context. A showing that a particular welfare recipient had access to additional income, or that a disability recipient's eligibility turned on testimony rather than written medical reports, would not result in an exception from the required procedural norms.

Id. (emphasis in original).

111. *See supra* text accompanying notes 18-24.

cess inconsistent, but it is impossible to review economically or fairly on appeal. As Justice Blackmun, dissenting in *Lassiter*, stated:

> [T]he case-by-case approach advanced by the Court itself entails serious dangers . . . . The pleadings and transcripts of an uncounselled termination proceeding at most will show the obvious blunders and omissions of the defending parent. Determining the difference legal representation would have made becomes possible only through imagination, investigation, and legal research focused on the particular case . . . . [F]ailures to challenge the state's evidence or to develop a satisfactory defense . . . often cut to the essence of the fairness of the trial, and a court's inability to compensate for them effectively eviscerates the presumption of innocence.[112]

IV. Conclusion

Paternity defendants in state prosecuted actions should be entitled to counsel, not on a case-by-case basis, but on the basis of their membership in a class of defendants entitled to that right. In adopting the case-by-case approach, the *Nordgren* and *Lassiter* courts have embraced a mode of analysis previously abandoned by the United States Supreme Court in other right-to-counsel cases. When paternity cases are analyzed in light of the Court's usual approach to due process problems, it becomes clear that paternity defendants must be granted counsel to minimize the substantial risk of erroneous results and the correspondingly serious financial and social damage those determinations could inflict. The state's interest in economical proceedings is hardly commensurate with the interest in correct results that the state shares with both the putative father and the child whose interests it seeks to protect. Surely, when the state itself seeks to impose such a fundamental relationship upon parties otherwise unable to defend themselves, appointed counsel should be provided.

Rodney R. Parker

112. 452 U.S. at 50-51 (Blackmun, J., dissenting).

H.L. v. Matheson—A Minor Decision About Parental Notice

In *H.L. v. Matheson*,[1] the United States Supreme Court upheld a Utah statute[2] that requires parental notification before an unmarried, unemancipated minor can obtain an abortion. The Court strictly limited its holding to the facts of the case, upholding the statute as applied to the particular class of unemancipated minors living with and dependent on their parents, and making no claim or showing of advanced maturity or unusual parental relations.[3] *Matheson*'s limited holding has little precedential force, but sweeping dicta in the Court's opinion suggest that in future cases, a presumption[4] against maturity and unusual parental relations will require plaintiffs to prove those characteristics. When a statute impinges on protected individual liberties, the state has traditionally borne the burden of showing that the statute is rationally related to some state interest.[5] The presumptions raised in *Matheson*, however, threaten the constitutional guarantee of decisional privacy by placing the burden on individual plaintiffs to show that the statute has no rational relationship as applied to them.

The plaintiff in *Matheson* was a dependent fifteen-year-old girl who lived with her parents in Utah. After discovering she was pregnant, the plaintiff consulted with a social worker and a physician. The physician advised the plaintiff that an abortion would be in her best medical interest. Because of Utah's parental notice statute, however, he refused to perform the abortion without first notifying her parents.[6] When the plaintiff challenged the statute in state court, her case was dismissed and the Utah Supreme Court upheld the dismissal.[7] The United States Supreme Court affirmed, holding that as applied to the class before it, the statute plainly served important state interests, was narrowly drawn to protect only those interests, and did not violate any constitutional

---

1. 450 U.S. 398 (1981).
2. UTAH CODE ANN. § 76-7-304 (1978).
3. 450 U.S. at 413.
4. "Presumption" is used to indicate a shift in the burden of proof. *See infra* notes 65-67 and accompanying text.
5. *See infra* notes 31-67 and accompanying text.
6. H__ L__ v. Matheson, 608 P.2d 907, 908 (Utah 1979), *aff'd*, 450 U.S. 398 (1981).
7. *Id.* at 908, 913.

guarantee.[8]

I. The Privacy Right in the Abortion Decision

The right to decisional privacy was first recognized in *Griswold v. Connecticut*.[9] In that case, a Connecticut statute that forbade the use of contraceptives was struck down as violative of a constitutional right of decisional privacy. The Court reasoned that peripheral rights arise from the specific guarantees expressed in the Bill of Rights.[10] The right of a married couple to decide whether to use contraceptives, was found to lie "within the zone of privacy created by several fundamental constitutional guarantees."[11]

The right to decisional privacy was first applied to the abortion decision in *Roe v. Wade*.[12] In *Roe*, a pregnant single woman challenged the constitutionality of a Texas statute that made abortion a crime except when performed as a medical necessity to save the mother's life.[13] The Court recognized that any regulation limiting "fundamental rights," such as the right of privacy, could be justified only by a "compelling state interest."[14] The Court found a number of state interests that supported anti-abortion statutes, including safeguarding the mother's health, maintaining medical standards, and protecting potential life.[15] The Court reasoned, however, that those interests are not compelling throughout the

8. 450 U.S. at 413.

9. 381 U.S. 479 (1965). The adjective "decisional" is used to distinguish the privacy right involved in *Griswold* from, for example, the right to privacy from unwarranted governmental searches or seizures.

10. *See id.* at 484. For example, the Court concluded: "[t]he right of freedom of speech and press includes not only the right to utter or to print, but the right to distribute, the right to receive, the right to read and freedom of inquiry, freedom of thought, and freedom to teach." *Id.* at 485. *See generally*, Comments on the *Griswold* Case, 64 Mich. L. Rev. 197 (1965).

11. 381 U.S. at 485. Among the guarantees creating zones of privacy are the first amendment right to assemble, the third amendment prohibition against quartering soldiers in a person's home without consent during peace time, the fourth amendment right of security against unreasonable searches and seizures, the fifth amendment right against self-incrimination and the residual rights of the ninth amendment. *Id.* at 484.

12. 410 U.S. 113 (1973). For a criticism of the *Roe* decision see Ely, *The Wages of Crying Wolf: A Comment on* Roe v. Wade, 82 Yale L.J. 920 (1973). For a pre-*Roe* discussion of the potential application of the decisional privacy right to the abortion decision see Means, *The Phoenix of Abortional Freedom: Is a Penumbral or Ninth-Amendment Right About to Arise from the Nineteenth Century Legislative Ashes of a Fourteenth-Century Common-Law Liberty?*, 17 N.Y.L.F. 335 (1971).

13. Tex. Penal Code Ann. §§ 1191-94, 1196 (Vernon 1961).

14. 410 U.S. at 155.

15. *Id.* at 154.

pregnancy, but become compelling at some point during the pregnancy.[16] Because an early abortion is generally less hazardous than giving birth, the Court reasoned that the state has no compelling health reason to regulate abortion until "approximately the end of the first trimester."[17] Therefore, the Court in *Roe* held that through the first trimester, the abortion decision is entirely up to the woman and her physician, to be made in light of the physician's best medical judgment after consultation with the woman.[18]

The Court further held that the state's interest in the health of the mother becomes compelling at approximately the end of the first trimester; from that point forth the state "may regulate the abortion procedure to the extent that the regulation reasonably relates to the preservation and protection of maternal health."[19] Permissible regulation would include establishing licensing requirements for the hospital and person performing the abortion, establishing minimum qualifications for the person performing the abortion, and specifying the type of facility where abortions may be performed.[20]

The Court found that the state's interest in protecting fetal life becomes compelling at viability.[21] During the period after viability, the state "may go so far as to proscribe abortion except when it is necessary to preserve the life or health of the mother."[22] In sum, *Roe* precludes the state from regulating the abortion decision during the first trimester, allows the state to regulate abortions to the extent reasonably necessary to protect maternal health in the second trimester, and permits the state to proscribe abortion during the final trimester except where a physician determines it is necessary to preserve the life or health of the mother.[23]

In its first major post-*Roe* confrontation with abortion regula-

---

16. *See id.*

17. *Id.* at 163.

18. *Id.*

19. *Id.*

20. *Id. Cf.* Doe v. Bolton, 410 U.S. 179, 193-95 (1973) (holding unconstitutional a Georgia law that required performance of procedure in state-licensed hospital, but not limited to the second or third trimester).

21. 410 U.S. at 163. The Court noted that a fetus is viable when it is "potentially able to live outside the mother's womb, albeit with artificial aid." *Id.* at 160 (citation omitted). The Court further accepted the medical viewpoint that viability usually occurs at twenty-eight weeks, but may occur as early as twenty-four weeks. *Id.* (citing L. HELLMAN & J. PRITCHARD, WILLIAMS OBSTETRICS 493 (14th ed. 1971)).

22. *Id.* at 163-64.

23. *Id.* at 164-65.

tions, *Planned Parenthood v. Danforth*,[24] the Court held that a state may not require a blanket parental consent for a minor's abortion.[25] The *Danforth* statute required an unmarried woman under age 18 to obtain the written consent of a parent or person *in loco parentis* prior to obtaining an abortion during the first twelve weeks of pregnancy.[26] The Court recognized that if a state cannot regulate or proscribe abortion during the first trimester, it cannot rationally grant power to a third person to do so.[27]

The Court further reasoned that "[a]ny independent interest the parent may have in the termination of the minor daughter's pregnancy is no more weighty than the right of privacy of the competent minor mature enough to have become pregnant."[28] Earlier cases had established that minors are protected by the same constitutional guarantees as adults.[29] Drawing on that rationale, the Court in *Danforth* stated that "rights do not mature and come into being magically only when one attains the state-defined age of majority."[30] The Court, however, did not forget prior decisions establishing a significant state interest in the well-being of its minors.[31] Because of that state interest, the *Danforth* Court recognized that state regulation inhibiting a minor's right to privacy is subject to a relatively low level of judicial scrutiny, requiring only that the restriction serve "any significant state interest . . . that is not present in the case of an adult."[32] In failing to specify by what means the restriction must "serve" the interest, the Court impliedly established a rational relationship test.[33]

Bellotti v. Baird,[34] the Court's next abortion decision regard-

---

24. 428 U.S. 52 (1976).

25. *Id.* at 74.

26. *Id.* at 84-89 (text of statute in appendix to opinion).

27. *See id.* at 69, 75. Spousal consent requirements were also determined to be unconstitutional by that rationale. *Id.*

28. *Id.* at 75 (dictum).

29. *See In re* Gault, 387 U.S. 1, 13 (1967) ("Neither the fourteenth amendment nor the Bill of Rights is for adults alone").

30. Planned Parenthood v. Danforth, 428 U.S. 52, 74 (1976).

31. *See, e.g.*, Ginsberg v. New York, 390 U.S. 629, 640 (1968) (plurality opinion) (particular state interest was preventing harm to minors from exposure to obscenity).

32. Planned Parenthood v. Danforth, 428 U.S. 52, 75 (1976). *See also* Prince v. Massachusetts, 321 U.S. 158, 170 (1944) ("The power of the state to control the conduct of children reaches beyond the scope of its authority over adults"). In contrast, state regulations limiting the fundamental rights of adults may be justified only by a "compelling state interest." Roe v. Wade, 410 U.S. 113, 155 (1973).

33. *See* Carey v. Population Servs. Int'l, 431 U.S. 678, 705-07 (1977); Ginsberg v. New York, 390 U.S. 629, 639-41 (1968).

34. 443 U.S. 622 (1979) (two pluralities) (*Bellotti II*). In *Bellotti II*, the Court reiter-

ing minors, recognized the state's special interest "in encouraging an unmarried pregnant minor to seek the advice of her parents in making the important decision whether or not to bear a child."[35] The statute challenged in *Bellotti*[36] required parental consent, but if consent was denied, the statute authorized the abortion "by order of a judge . . . for good cause shown, after such hearing as he deems necessary."[37] The Court found the statute unconstitutional because it allowed a judge to deny authorization even if the minor was found to be mature and competent to make the decision, and because it required parental consultation in every instance regardless of whether it would be in the minor's best interest and regardless of the minor's maturity and competence to make the abortion decision.[38]

Although it held the statute unconstitutional, the Court elaborated on a possibility raised in *Danforth* that, under certain circumstances, a parental consent requirement might be constitutional.[39] The Court indicated that a valid consent statute would provide a viable alternative procedure for obtaining abortion authorization.[40] According to the Court, a pregnant minor could obtain an abortion, without parental consent, if she could show either: "(1) that she is mature enough and well enough informed to make her abortion decision, in consultation with her physician, independently of her parent's wishes; or (2) that even if she is not able to make this decision independently, the desired abortion would be in her best interests."[41]

Justice White in dissent provided the Court's first indication that a parental *notification*, as distinguished from a *consent* statute might be unconstitutional. Analyzing the plurality opinion's re-

---

ated that "the States may validly limit the freedom of children to choose for themselves in the making of important, affirmative choices with potentially serious consequences." *Id.* at 635 (dictum). That decision represented the second time the Massachusetts statute was before the Court. In Bellotti v. Baird, 428 U.S. 132 (1976) (*Bellotti I*), the Court vacated the decision of the district court and remanded for certification of questions to the Massachusetts Supreme Judicial Court. *Bellotti II* determined the constitutionality of the Massachusetts statute as construed by the Massachusetts Supreme Judicial Court.

35. *Bellotti II*, 443 U.S. at 639.

36. MASS. GEN. LAWS ANN. ch. 112, § 125 (West Supp. 1979).

37. *Id.*

38. *See Bellotti II*, 443 U.S. at 644-51.

39. *See id.* at 643. The Court in *Danforth* held that a blanket parental consent requirement is unconstitutional, not that every such requirement is unconstitutional. Planned Parenthood v. Danforth, 428 U.S. 52, 74 (1976).

40. 443 U.S. at 643.

41. *Id.* at 643-44 (dictum) (footnote omitted).

quirement that the proceeding and any appeals must be conducted with anonymity, White reasoned that the total anonymity would render a blanket parental notice statute unconstitutional.[42] Parental notice statutes were in fact held unconstitutional in lower courts on the ground that no state interest would be served by requiring notice to the parents of an emancipated or mature minor.[43]

II. *H.L. v. Matheson*

In *H.L. v. Matheson*,[44] the issue of notice, as distinguished from consent, was directly addressed by the United States Supreme Court. At trial the plaintiff sought a declaratory judgment that a Utah statute requiring parental notification in every case was unconstitutional. The plaintiff also sought an injunction against its enforcement. After a hearing in which the plaintiff responded monosyllabically to questions, the trial court denied both a temporary restraining order and a preliminary injunction and entered a judgment dismissing the plaintiff's action.

On appeal, the Utah Supreme Court unanimously affirmed the trial court's action and upheld the constitutionality of the statute.[45] The court determined that the notice statute served two "significant state interest[s] that [are] not present in the case of an adult."[46] Those interests include safeguarding the health of the pregnant minor,[47] and encouraging the minor to consult with her

---

42. *See id.* at 657 (White, J., dissenting).

43. *See* Wynn v. Carey, 582 F.2d 1375, 1390 (7th Cir. 1978); Margaret S. v. Edwards, 488 F. Supp. 181, 205 (E.D. La. 1980); Planned Parenthood Ass'n v. Ashcroft, 483 F. Supp. 679, 697 (W.D. Mo. 1980), *modified on other grounds*, 655 F.2d 848 (8th Cir. 1981); Leigh v. Olson, 497 F. Supp. 1340, 1351 (D.N.D. 1980); Women's Servs. v. Thone, 483 F. Supp. 1022, 1048 (D. Neb. 1979); Akron Center for Reproductive Health, Inc. v. City of Akron, 479 F. Supp. 1172, 1202 (D. Ohio 1979), *modified on other grounds*, 651 F.2d 1198 (6th Cir. 1981); Women's Community Health Center, Inc. v. Cohen, 477 F. Supp. 542, 546 (D. Me. 1979). *See generally*, Note, *Parental Notice Statutes: Permissible State Regulation of a Minor's Abortion Decision*, 49 FORDHAM L. REV. 81 (1980); Comment, *The Maine Abortion Statutes of 1979: Testing the Constitutional Limits*, 32 ME. L. REV. 315 (1980).

44. 450 U.S. 398 (1981).

45. H__ L__ v. Matheson, 604 P.2d 907, 912 (Utah 1979), *aff'd*, 450 U.S. 398 (1981).

46. *Id.* at 912.

47. The Utah Supreme Court explained the relation of parental notice to the minor's health as follows: "[I]t is the parent of the minor, who would frequently possess additional information, which might prove invaluable to the physician in exercising his 'best medical judgment.'" *Id.* at 909-10. The United States Supreme Court reiterated this state interest noting that "[p]arents can provide medical and psychological data, refer the physician to other sources of medical history, such as family physicians, and authorize family physicians to give relevant data." 450 U.S. at 411.

parents prior to making the important decision about abortion.[48]

In a plurality decision, the United States Supreme Court affirmed.[49] The Court decided the issue narrowly, upholding a statutory requirement of parental notification by the physician of an unmarried, unemancipated pregnant minor seeking an abortion who makes no claim or showing of maturity or of adverse relations with her parents.[50]

In her appeal to the Supreme Court, the plaintiff first argued that the statute was overbroad because it could be applied to all unmarried minor girls, including those who are mature.[51] The Court dismissed that argument because the plaintiff failed to allege or proffer evidence that either she or any member of her class was mature or emancipated.[52] The Court then listed the trial court's findings that plaintiff was "unmarried, fifteen years of age, reside[d] at home and [was] a dependent of her parents."[53] Accordingly, the Court concluded there was "an insufficient basis for a finding that she is either mature or emancipated."[54] Implicit in that analysis is a presumption that pregnant minors seeking an abortion are immature.[55]

Having narrowed the class before it to immature unemancipated minors neither claiming nor showing unusual maturity or unusual parental relations, the Court reviewed its previous decisions on the right to privacy as they applied to the plaintiff. The Court's analysis concentrated on *Danforth* and *Bellotti*, distinguishing the issues of consent and notice and reiterating its view that a minor often is unable to make an important decision in her best interests.[56] The Court emphasized the value of the "mature advice and emotional support" that the girl's parents could provide.[57] The Court found that "[t]he Utah statute gives neither parents nor judges a veto power over the minor's abortion decision,"[58]

---

48. 604 P.2d at 912.
49. 450 U.S. 398 (1981) (plurality opinion).
50. *Id.* at 413.
51. *Id.* at 405.
52. *Id.* at 405-06.
53. *Id.* at 406.
54. *Id.*
55. This presumption seems to contradict the reasoning of the Court in *Danforth* where the Court protected the privacy of minors "mature enough to have become pregnant." Planned Parenthood v. Danforth, 428 U.S. 52, 75 (1976). *See also infra* note 71 and accompanying text.
56. 450 U.S. at 409-10.
57. *Id.* at 410 (quoting *Bellotti II*, 443 U.S. at 640-41).
58. *Id.* at 411.

and that when applied to immature and dependent minors, the statute "plainly serves the important considerations of family integrity and protecting adolescents."[59] The Court thus established a second implicit presumption: notice to parents serves the state interest in preserving family integrity.

Finally, the Court found that the statute served another state interest "by providing an opportunity for parents to supply essential medical and other information to a physician,"[60] thereby aiding in the determination of any long-term medical, emotional, and psychological consequences of an abortion. With that finding, the Court presumed that notice would serve the state's interests in the health of the minor. Notwithstanding the narrow holding of *Matheson*, the Court's analysis seems to have established three broad presumptions: (1) unmarried, unemancipated minors living at home are not mature enough to make the abortion decision without parental consultation; (2) the state interest in family integrity is served by parental notice; and (3) the state interest in the minor's health is served by parental notice.

III. Impact of the Case

Matheson acknowledges the established constitutional right to decisional privacy[61] in the abortion decision.[62] It further acknowledges that, although minors are protected by such constitutional rights, a state statute that is rationally related to a significant state interest can overcome the qualified constitutional rights of minors when the state interest in protecting minors is different from its interest in protecting adults.[63] Because the Court's holding simply followed those established rules, an examination of the *Matheson* case might properly end here. The impact of the three presumptions raised in dicta, however, may have more far-reaching consequences and merits further examination.

The *Matheson* Court placed difficult barriers in the path of future plaintiffs who do not fit the Court's limited holding. In the future, plaintiffs must allege and prove their maturity or that

59. *Id.*

60. *Id.*

61. *See id.* at 407-08. *See also supra* notes 9-11 and accompanying text.

62. *See* 450 U.S. at 407-08. *See also supra* notes 12-13 and accompanying text.

63. *See* 450 U.S. at 409. *See also Bellotti II*, 443 U.S. 622 (1979); Carey v. Population Servs. Int'l, 431 U.S. 678, 692 (1977); *Bellotti I*, 428 U.S. 132 (1976); Planned Parenthood v. Danforth, 428 U.S. 52, 74 (1976); Ginsberg v. New York, 390 U.S. 629, 637 (1968); *In re* Gault, 387 U.S. 1, 13 (1967).

neither family integrity nor the minor's health interests would be served by parental notice.[64] Moreover, the Court's presumptions depart from the traditional requirement that the state either justify a statute which "burdens the exercise of a fundamental right,"[65] or explore less burdensome alternatives.[66] In view of the intrinsically unique factors involved in one's personal health and family relations, the Court's presumptions evidence a disturbing deterioration of the fundamental rights of minors.

IV. Analysis and Recommendations

A. *The Presumptive Derogation of a Constitutional Right*

When the state seeks to restrict a constitutional right, the burden has traditionally been the state's to show that its interest is rationally served by the statute.[67] By raising presumptions as to the beneficial results of the Utah parental notification statute, the *Matheson* Court shifted the burden to the minor to show that no state interest is served in her case. That departure from the traditional state burden denigrates the right of decisional privacy as applied to minors. Privacy becomes not a right that minors share with adults, but merely a privilege that minors must earn by showing that the state has no significant interest in their individual situations.

The well-being of the immature minor[68] and the preservation

---

64. 450 U.S. at 403 n.8, 406.

65. Carey v. Population Servs. Int'l, 431 U.S. 678, 696 (1977) (plurality opinion).

66. *See* Shelton v. Tucker, 364 U.S. 479, 488 (1960): "[E]ven though the governmental purpose be legitimate and substantial, that purpose cannot be pursued by means that broadly stifle fundamental personal liberties when the end can be more narrowly achieved."

67. See Carey v. Population Servs. Int'l, 431 U.S. 678 (1977), where the Court stated:

> [W]hen a State . . . burdens the exercise of a fundamental right, its attempt to justify that burden as a rational means for the accomplishment of some significant state policy requires more than a bare assertion, based on a conceded complete absence of supporting evidence, that the burden is connected to such a policy.

Id. at 696 (footnote omitted).

68. If the minor is mature, the state interest in protecting the minor is no different than it would be for an adult; therefore, no rational relationship between the notice statute and the state interest can exist. *See Bellotti II*, 443 U.S. 622 (1979). The *Matheson* Court noted that the Utah statute cannot be constitutionally applied to emancipated minors. 450 U.S. at 406. The United States District Court for Utah held that the statute "does not apply to emancipated minors and that, if so applied it would be unconstitutional." *Id.* (citing L.R. v. Hansen, No. C-80-0078J (D. Utah Feb. 8, 1980)). The *Matheson* Court further stated that "since there was no appeal from [the district court's] ruling, it is controlling on the State." *Id.* at 406. The Court also assumed that the statute would be construed to exempt demonstrably mature minors in a proper case, basing that assumption on *Bellotti I*, 428 U.S. 132, 146-48 (1976).

of family integrity are certainly important state interests.[69] It is inappropriate, however, to presume broadly that such interests will be served by parental notification statutes. Maturity, physical or mental well-being, and the degree of family integrity are factors too variable in nature to be generally presumed. Maturity is an absolutely individual characteristic. Indeed, Judge Dembitz of the Family Court of the State of New York asserts that because "[m]inors who seek abortions express reasons similar to those of adults, and demonstrate a comparable appreciation of their dilemmas,"[70] those minors seeking abortion ought to be presumed mature.[71] The privacy rights of minors must not be abrogated by presuming a rational relationship between state interests and statutes regulating abortions for minors.[72] The burden is properly on the state to provide evidence of rationality,[73] and therefore the state should have to provide evidence that the individual minor is immature.

Similarly, the determination of the effect of parental notification on the well-being of a minor will vary from case to case. The decision to have an abortion undoubtedly will cause some psychological trauma, but possibly no more than requiring the minor to notify her parents.[74] It can hardly be presumed generally that notification will cause less harm. Once again the state should have to provide evidence that notification will cause the lesser harm;[75] otherwise, the statute may in fact not serve the state's interest, but defeat it.

Whether parental notification will serve the state interest in preserving family integrity depends on the minor's actual familial relationships. Certainly "parental involvement in a minor's abortion decision, if compassionate and supportive, [is] highly desira-

---

69. *See* H. L. v. Matheson, 450 U.S. 398 (1981).

70. Dembitz, *The Supreme Court and a Minor's Abortion Decision*, 80 COLUM. L. REV. 1251, 1256 (1980) ("Like adults, they are attempting to escape from a potentially shattering, life-long handicap").

71. *See id.* at 1255-56 ("[A] minor's very decision to seek an abortion shows deliberation, a sense of responsibility and foresight as to consequences—qualities lacking in numerous teenagers, who ignore their pregnancies, fantasizing that they will magically disappear").

72. *See* Planned Parenthood v. Danforth, 428 U.S. 52, 75 (1976).

73. *See* Carey v. Population Servs. Int'l, 431 U.S. 678, 696 (1977). *See also supra* note 67.

74. *See* Dembitz, *supra* note 70, at 1255.

75. This might be accomplished by a psychological determination of the minor's moral attitudes about abortion and a sociological and psychological determination of the possible long-term psychological effect of destroying the familial relationship in situations where that is determined to be the likely result.

ble."[76] When that reaction is not likely, however, parental involvement is most probably undesirable. If the minor's views on abortion are antithetical to those of her parents, she may be faced with the choice of leaving home if she decides to have the abortion.[77] Even if her parents concur that an abortion is desirable, her unchaste behavior may disillusion her parents to the extent that there is an irretrievable breakdown of previous family trust and harmony.[78]

A minor not wishing to tell her parents of her abortion decision probably has a good reason for not doing so. A general presumption that she does not have a good reason, in light of the wide variance of familial relationships, is a non sequitur. The interest of the state in preserving family integrity would probably not be served by statutorily mandated parental notification in many cases. Accordingly, Judge Dembitz has stated:

> In fact, young unmarried women facing the choice between an unwanted pregnancy and an abortion . . . customarily consult their mothers; a minor who, faced with pregnancy or abortion, objects to parental notification is therefore probably aware from parental threats and attitudes . . . that she is in acute danger of expulsion from her home, or of scorn, degradation, physical abuse, or punishment if her parents learn of her abortion decision.[79]

Thus, the burden should once again be on the state to show that the requirement of parental notification will serve its interest before the minor's right to decisional privacy can be overcome.

B. The Requirement to Pursue Less Burdensome Alternatives

The interests of the state in protecting the health and well-being of pregnant minors and their families can be better served by alternatives that would be less oppressive to the minor's constitutional rights. The Supreme Court has said that "even though the governmental purpose be legitimate and substantial, that purpose cannot be pursued by means that broadly stifle fundamental personal liberties when the end can be more narrowly achieved."[80] The ends of protecting the mother's health and preserving family integrity can be achieved more narrowly through a case-by-case

---

76. *Bellotti I*, 443 U.S. at 640 n.20.
77. *See* Dembitz, *supra* note 70, at 1257.
78. *See id.*
79. *Id.*
80. Shelton v. Tucker, 364 U.S. 479, 488 (1960).

determination.

With the burden on the state to show a minor's immaturity, a hearing could be required in order to assess the minor's ability to understand and make a mature decision regarding her situation. An individual assessment of maturity would allow a minor to maintain her constitutional rights unless the state can prove that she is immature. That approach would permit a true balancing of a minor's constitutional rights against significant state interests.

Similarly, a two-step process with the burden on the state could be used to ensure that the minor's health interests would be served by parental notice. First, a psychological and sociological examination of the minor could be made to enable a balancing of any psychological risks to the minor in making the abortion decision without parental input against the trauma she might experience if her parents were given notice that she is pregnant and seeking an abortion. A second phase of this hearing process would involve an evaluation of the comparative risks to the minor's physical health posed by making the abortion decision with and without parental input. That determination might include any threat of physical harm being inflicted by the minor's parents if they are informed about her situation. Such information might be obtained from the psychological-sociological examination. Additionally, the court would have to weigh the value of any medical background information that the minor's parents could provide against the harm that might be done through disclosure. That determination would likely depend on the adequacy of medical information obtainable confidentially from the minor and her school or family physician. Although the court is not qualified to make the expert psychological, sociological and medical determinations involved in this two-step process, once the expert determinations are made, it could properly weigh the different expert findings and decide whether, on the evidence presented, the state has shown that notification of the minor's parents would protect her health.

The hearing process could be used to allow the state to show that family integrity would be served by providing notice to the minor's parents. Evidence on that issue could be obtained by those experts conducting the psychological and sociological examinations of the minor described above. Their inquiry would focus on the nature of the minor's relationship to her parents and the basis for her reluctance to speak with them about the abortion decision. If the finding is that family integrity would be served, notice would be required. If the finding is that family integrity is weak or would

be weakened if notice were given, the minor's fundamental right to decisional privacy would outweigh the state's interest in family integrity. Again, the court would likely require expert assistance in making that determination.[81]

V. Conclusion

The *Matheson* decision may have been correct as applied to the plaintiff before the Court, but the opinion contains no supporting data by which to judge. Although the case of a minor does not require the strict "compelling interest" scrutiny required for a state to overcome an adult's constitutional right, it nonetheless requires a rigorous standard of judicial review. A state statute that encumbers the constitutional right of a minor must be rationally related to the accomplishment of a significant state interest. The interests of a state in the well-being of its immature minors and in family integrity are undeniably significant, but the burden must remain with the state to show that notice to the minor's parents will serve those interests.

Whether a minor's right to privacy in the abortion decision is outweighed by state interests is determined by individual circumstances; thus blanket presumptions are offensive to the constitutional rights of minors. To restrict constitutionally guaranteed rights by unsupported presumptions is to enervate the constitutional protections and to chill the spirit of liberty. In the future, the Court should only interfere with such rights after great thought and reasoned analysis, and it should, as it advises legislatures, "act with particular sensitivity"[82] in this area.

David W. Scofield

81. That is why the plaintiff's objections to questions in the district court hearing in *Matheson* regarding the minor's reasons for not wanting to inform her parents seem to have been properly upheld as irrelevant to the constitutional issue. Those questions could not relate to the constitutional issue; only a professional determination by psychologists and sociologists would apply to the court's determination of whether the state has met its burden of showing a rational relationship.

82. *Bellotti II*, 443 U.S. at 642 (dictum).

New Rewards for Released-Time: *Lanner* v. *Wimmer* Expands Constitutional Church-State Involvement

The establishment clause of the first amendment, despite its two-hundred-year history, was first construed by the United States Supreme Court only three decades ago.[1] That construction, however, did not silence all speculation about the meaning of the clause.[2] Instead, it reflected two competing interpretations of the prohibition against an established religion: absolute separation of church from state, or some accommodation of church by state.[3] Shortly thereafter, in cases challenging the constitutionality of "released-time" religious education programs,[4] the competing separation and accommodation constructions emerged clearly, with both being relied on by the Court.[5] The confusion that resulted from those ambiguous standards[6] was the legacy of the Tenth Circuit

---

1. *See* Everson v. Board of Education, 330 U.S. 1, 15 (1947).

2. *See, e.g.*, School Dist. of Abington Township v. Schempp, 374 U.S. 203, 254-65 (1963) (Brennan, J., concurring). *See generally* E. CORWIN, THE CONSTITUTION AND WHAT IT MEANS TODAY 269-77 (13th ed. 1973) (tracing the origin and application of various interpretations of the religion clauses).

3. *Everson* dicta was strongly separationist but the decision itself allowed accommodation of church needs by the state. *See infra* notes 28-30 and accompanying text.

4. "Released-time" programs release students to churches for religious education during time they would otherwise be in public schools. L. PFEIFFER, GOD, CAESAR, AND THE CONSTITUTION 180-82 (1975).

5. *Compare* Zorach v. Clauson, 343 U.S. 306 (1952) (government may impartially cooperate with religion) *with* McCollum v. Board of Education, 333 U.S. 203 (1948) (government should not participate at all in religion).

6. Neither courts nor commentators have easily explicated the Court's standards. *See* Meek v. Pittenger, 421 U.S. 349, 358 (1975); Lemon v. Kurtzman, 403 U.S. 602, 612, 614 (1971); Smith v. Smith, 523 F.2d 121, 123-24 (4th Cir. 1975), *cert. denied*, 423 U.S. 1072 (1976); State *ex rel.* Holt v. Thompson, 6 Wis. 2d 659, 225 N.W.2d 678, 681-83, 681 n.5 (1975); Choper, *The Establishment Clause and Aid to Parochial Schools*, 56 CAL. L. REV. 260 (1968); Giannella, *Religious Liberty, Nonestablishment, and Doctrinal Development* (pt. 2), 81 HARV. L. REV. 513 (1968); Kauper, *Church and State: Cooperative Separatism*, 60 MICH. L. REV. 1 (1961); Kruse, *The Historical Meaning and Judicial Construction of the Establishment of Religion Clause of the First Amendment*, 2 WASHBURN L.J. 65 (1962); Moehlman, *The Wall of Separation: The Law and the Facts*, 38 A.B.A.J. 81 (1952); Nowak, *The Supreme Court, the Religion Clauses and the Nationalization of Education*, 70 NW. L. REV. 883 (1976); Taylor, *Equal Protection of Religion: Today's Public School Problem*, 38 A.B.A.J. 277 (1952); Note, *The "Released Time" Cases Revisited: A Study of Group Decisionmaking by the Supreme Court*, 83 YALE L.J. 1202 (1974); Comment, 50 MICH. L. REV. 1359 (1952); Comment, 26 S. CAL. L. REV. 186 (1953).

Court of Appeals in *Lanner v. Wimmer*.[7] In *Lanner*, the court found portions of a Utah school district's released-time program unconstitutional but simultaneously expanded permissible church-state involvement in education.

I. BACKGROUND

A. *Origins of the Religion Clauses*

Unlike many other constitutional values, the first amendment's guarantees of free exercise of religion and freedom from establishment[8] are primarily American in origin.[9] Ironically, however, a European heritage of religious persecution was perpetuated at first in many of the colonies.[10] This religious conflict engendered strong desires for religious freedom,[11] an impetus fueled by successful experiments in toleration and disestablishment in several of the colonies.[12] The Virginia experience[13] contributed most significantly to the shaping of the Constitution's religion clauses. Writings of Thomas Jefferson and James Madison first helped convince Virginians of the need for religious liberty[14] and later served as the philosophical bases for the Bill of Rights' religious guarantees.[15]

---

7. 662 F.2d 1349 (10th Cir. 1981).

8. The pertinent part of the first amendment reads: "Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof. . . ." U.S. CONST. amend. I.

9. R. MILLER & R. FLOWERS, TOWARD BENEVOLENT NEUTRALITY: CHURCH, STATE, AND THE SUPREME COURT 1 (1977).

10. *See* F. CURRAN, CATHOLICS IN COLONIAL LAW 1-8 (1963); W. TORPEY, JUDICIAL DOCTRINES OF RELIGIONS RIGHTS IN AMERICA 8-11 (1948).

11. *See* R. HOFSTADTER, AMERICA AT 1750: A SOCIAL PORTRAIT 187-204 (1971); R. MILLER & R. FLOWERS, *supra* note 9, at 2-4; W. TORPEY, *supra* note 10, at 13-14.

12. *See generally* R. HOFSTADTER, *supra* note 11, at 194-204 (discussing generally the sources of religious freedom in the United States); H. LASKI, THE AMERICAN DEMOCRACY 28 (1948) (citing the Rhode Island and Pennsylvania experiences); W. TORPEY, *supra* note 10, at 10-12 (discussing toleration in Rhode Island, Maryland, Pennsylvania and Delaware); W. SWEET, THE STORY OF RELIGION IN AMERICA 77-82 (M. Meyers ed. 1950) (citing Rhode Island and Maryland).

13. *See* T. JEFFERSON, THE COMPLETE JEFFERSON 946 n.1 (1943); R. MILLER & R. FLOWERS, *supra* note 9, at 4; THE MIND OF THE FOUNDER: SOURCES OF THE POLITICAL THOUGHT OF JAMES MADISON 7-8 (1973).

14. T. JEFFERSON, A BILL FOR ESTABLISHING RELIGIOUS FREEDOM, in THE COMPLETE JEFFERSON, *supra* note 13, at 946 (freedom of religion is a natural right, which should not be fettered by any civil obligation; church and state will each operate better if separate from the other); J. MADISON, A MEMORIAL AND REMONSTRANCE, in LETTERS AND OTHER WRITINGS OF JAMES MADISON 162 (1884) (the state has no authority to direct religion thoughts or actions; a just government should protect citizens' religious freedom with an "equal hand," not "invading the equal rights of any Sect, nor suffering any Sect to invade those of another").

15. *See* Everson v. Board of Education, 330 U.S. 1, 13 (1947) (dictum). Even a dis-

Familiarity with the establishment clause's origin, however, has not guaranteed its uniform interpretation. Two conflicting theories emerged even among early commentators.[16] One theory held that the establishment clause required only that government treat all sects the same, not that it be indifferent or unsupportive.[17] The other looked to Jefferson's interpretation, which required a "wall of separation" between church and state.[18] Neither theory was tested against specific facts in the nineteenth century, in part because the clause was held unenforceable against state action.[19]

Even when the Supreme Court commenced incorporation of religious liberty into the fourteenth amendment, it did not squarely decide between separation or accommodation. Actually, the establishment clause itself was not expressly incorporated until well into the twentieth century.[20] In 1940, the Supreme Court held

---

senter, Justice Rutledge, termed the two documents the "warp and woof of our constitutional tradition [of religious liberty]." *Id.* at 39 (Rutledge, J., dissenting). *See also* Reynolds v. United States, 98 U.S. 145, 162-64 (1878) (Constitution's religious freedoms defined by the works of Jefferson and Madison). *See generally* R. MILLER & R. FLOWERS, *supra* note 9, at 4-6 (discussing Jefferson's and Madison's shaping of Virginia religious freedom and their roles in the passage of the Bill of Rights).

16. *See* E. CORWIN, *supra* note 2, at 269-70.

17. J. STORY, COMMENTARIES ON THE CONSTITUTION OF THE UNITED STATES 627-34 (5th ed. 1891); T. COOLEY, THE GENERAL PRINCIPLES OF CONSTITUTIONAL LAW IN THE UNITED STATES OF AMERICA 224-25 (3rd ed. 1898).

18. The phrase originated in an 1802 letter from Jefferson to the Danbury Baptist Association expressing Jefferson's "sovereign reverence" for the religion clauses, which protect religious opinions from the powers of government by "building a wall of separation between church and state." Letter from Thomas Jefferson to Danbury Baptist Association (Jan. 1, 1802), *reprinted in* 8 THE WRITINGS OF THOMAS JEFFERSON 113 (Washington ed. 1861). In a later address, however, Jefferson said he did not believe "that instruction in religious opinion and duties was meant to be precluded by the public authorities" and urged that religious schools be established "on the confines of the University, so as to give their students ready and convenient access" to secular studies. Address by President Jefferson to the President and Directors of the University of Virginia (Oct. 7, 1822), *reprinted in* THE COMPLETE JEFFERSON, *supra* note 13, at 957-58.

19. *See* T. COOLEY, *supra* note 17, at 224. *See generally* Permoli v. New Orleans, 44 U.S. (3 How.) 589, 609 (1845) (dictum) (U.S. Constitution does not protect against state infringement on religious liberty); Barron v. Baltimore, 32 U.S. (7 Pet.) 243 (1833) (dictum) (Bill of Rights applies to federal, not state government). Establishment persisted in a variety of forms in New England until 1833 when Massachusetts finally adopted a separation provision. R. MILLER & R. FLOWERS, *supra* note 9, at 4.

20. In Meyer v. Nebraska, 262 U.S. 390 (1923), the Court included freedom of worship in the list of fundamental rights embraced by fourteenth amendment "liberty." *Id.* at 399. In Pierce v. Society of Sisters, 268 U.S. 510 (1925), the Court ruled that parochial school attendance satisfied state compulsory attendance laws, denying any "general power of the State to standardize its children by forcing them to accept instruction from public teachers only." *Id.* at 535. For an overview history of the incorporation of the religion clauses, see School Dist. of Abington Township v. Schempp, 374 U.S. 203, 253-58 (1963) (Brennan, J., concurring).

that the protections of the free exercise clause were enforceable against state action.[21] Not until 1947, however, in *Everson v. Board of Education*,[22] did the Court finally rule that the establishment clause should have the same application and broad interpretation as the free exercise clause, since the two clauses worked together to create full religious freedom.[23]

B. Everson, McCollum, Zorach: *Separation or Accommodation Under the Establishment Clause*

1. Everson v. Board of Education—The Supreme Court in *Everson* reviewed a taxpayer's challenge to a New Jersey statute and local school board resolution reimbursing parents for bus fare paid to transport children to public and parochial schools.[24] The taxpayer claimed that the program violated the establishment clause because the law required residents to pay taxes that furthered the goals of the Roman Catholic church.[25]

Justice Black, writing for the majority, utlimately rejected the first amendment challenge,[26] but his opinion announced the incorporation of the establishment clause and construed it as requiring a strict separation of church and state.[27] Reviewing the history of the religion clauses and emphasizing the roles of Jefferson and Madison in their formulation,[28] Justice Black concluded:

> The 'establishment of religion' clause of the First Amendment means at least this: Neither a state nor the Federal Government can set up a church. Neither can pass laws which aid one religion, aid all religions, or prefer one religion over another. Neither can force nor influence a person to go to or remain away from church against his will or force him to profess a belief or disbelief in any religion. No

---

21. Cantwell v. Connecticut, 310 U.S. 296, 303-04 (1940). *See also* L. TRIBE, AMERICAN CONSTITUTIONAL LAW 813-14 (1978) (discussing briefly the incorporation of the religion clauses).

22. 330 U.S. 1 (1947) (5-4 decision).

23. *Id.* at 15. *See* School Dist. of Abington Township v. Schempp, 374 U.S. 203, 257-58 (1963) (Brennan, J., concurring); R. MILLER & R. FLOWERS, *supra* note 9, at 7; L. TRIBE, *supra* note 21, at 814-15.

24. Because the school district itself operated no bus service, public and parochial school students alike rode public buses. 330 U.S. at 3; *id.* at 19-20 (Jackson, J., dissenting).

25. *Id.* at 5. The taxpayer also contended that this use of tax monies was a taking of private property for private use, without due process of law. Justice Black's majority opinion rejected that challenge, finding that the use was not solely private but had a secular public welfare purpose. *Id.* at 5-7.

26. *Id.* at 18.

27. *Id.* at 15.

28. *Id.* at 8-15.

> person can be punished for entertaining or professing religious beliefs or disbeliefs, for church attendance or non-attendance. No tax in any amount, large or small, can be levied to support any religious activities or institutions, whatever they may be called, or whatever form they may adopt to teach or practice religion. Neither a state nor the Federal Government can, openly or secretly, participate in the affairs of any religious organizations or groups and *vice versa*. In the words of Jefferson, the clause against establishment was intended to erect 'a wall of separation between church and State.'[29]

Despite that paragraph's strong separationist stance, the Court upheld the payments. Justice Black reasoned that the benefits were going to "citizens," not a church, and that the free exercise clause prohibits denial, on religious grounds, of public welfare benefits to anyone.[30]

2. *McCollum v. Board of Education*—Only a year later, in *McCollum v. Board of Education*,[31] the Court addressed the "released time" issue for the first time and invalidated a Champaign, Illinois program[32] that allowed weekly religion classes in public school buildings during school hours.[33] Illinois' compulsory education law required children aged seven to sixteen to be in school during regular school hours.[34] Students not participating in the religion classes were therefore required to stay in the school and continue their secular studies.[35] Parental consent was required for par-

---

29. *Id.* at 16. Although this paragraph is technically dicta, commentators point to it as the basis not only of McCollum v. Board of Education, 333 U.S. 203 (1948), which immediately followed *Everson*, but of other cases in the decades since, during which time the membership of the Court has totally changed. Fey, *An Argument for Separation*, in THE WALL BETWEEN CHURCH AND STATE 26, 32 (1963). For the source of the "wall of separation" phrase, see *supra* note 18.

30. 330 U.S. at 16. Justice Black likened the transportation aid to fire and police protection to which all citizens and organizations are entitled. *Id.* at 17-18. Justice Jackson, in dissent, found the "undertones" of the majority opinion "utterly discordant" with its conclusion. Because reimbursement went only to those attending nonprofit schools, not just any private school, he felt that a general public welfare purpose was precluded by the very terms of the statute and resolution. *Id.* at 18-20 (Jackson, J., dissenting). Justice Rutledge's dissent found permanent separation of church and state mandated by the history of the religion clauses. He argued that not the amount involved, but the principle, invalidated the program. *Id.* at 31-33, 49, 60-62 (Rutledge, J., dissenting).

31. 333 U.S. 203 (1948).

32. A state statute allowed local school boards to determine the use of public school buildings within the district. *Id.* at 205.

33. *Id.* at 207, 209. The religion classes lasted thirty minutes in the lower grades, forty-five in the higher. *Id.* at 207-08.

34. *Id.* at 205.

35. *Id.* at 209. The record is not clear on the form of these studies. Justice Frankfurter stated that nonparticipating children were "required to attend a regular school class, or a

ticipation in the religion classes, with request cards distributed by the schools and returned by the children or the public school teachers to the religion teachers.[36] Absences were recorded by religion teachers on forms provided by the school.[37] Although subject to approval and supervision by the school superintendent, the religion teachers were selected and hired by the program's sponsoring religious council.[38] Appellant McCollum, a taxpayer and parent, sought to prohibit the released-time program as a violation of the first and fourteenth amendments.[39]

Justice Black, again writing for the Court, found two elements of the Champaign program constitutionally objectionable. First, the released-time program imvolved the use of tax-supported public buildings.[40] Second, the program employed the state's compulsory education machinery to provide pupils for the religion classes.[41] Relying on his separationist language in *Everson*, Justice Black held that such "close cooperation" between religious authorities and the school system was impermissible under the first and fourteenth amendments.[42]

3. Zorach v. Clauson—In *Zorach v. Clauson*,[43] decided only three years later, the Court seemingly abandoned the separationist position of *McCollum* and *Everson*. In *Zorach*, the Court reviewed a first amendment challenge to a New York City program that released public school students for one hour a week to attend religious classes or devotional exercises.[44] The religion classes were held off the school grounds and no public funds were involved.[45]

---

study period during which [they were] often left to [their] own devices." *Id.* at 227 (Frankfurter, J., concurring). Justice Reed described the alternative as "study or released time periods." *Id.* at 238 (Reed, J., dissenting).

36. *Id.* at 207 n.2.

37. *Id.* at 209 n.5.

38. *Id.* at 207.

39. *Id.* at 205.

40. *Id.* at 209.

41. *Id.* at 212.

42. *Id.* at 209-10. Justices Frankfurter and Jackson concurred. Viewing the public school as a symbol of secular unity, Justice Frankfurter found reason to invalidate the program because the "momentum of the whole school atmosphere and school planning" supported the religious instruction. *Id.* at 280. Thus the program engendered subtle feelings of coercion and a heightened consciousness of religious differences. *Id.* at 213-17, 227-30 (Frankfurter, J., concurring). Justice Reed's dissent, by contrast, sketched a history of church-state cooperation and concluded that a state should have great leeway in handling these kinds of social issues. *Id.* at 244-56 (Reed, J., dissenting).

43. 343 U.S. 306 (1952) (6-3 decision).

44. *Id.* at 308.

45. *Id.* at 308-09.

Students were released upon written request of their parents, and although the churches made weekly attendance reports to the schools, the schools did not enforce attendance requirements.[46] Challengers argued that the program was an establishment of religion, similar to that in *McCollum*,[47] because the public school was used by the churches to support the program.[48]

In his majority opinion, Justice Douglas rejected the challenge, distinguishing the *Zorach* and *McCollum* programs; the *Zorach* program, he found, involved neither public facilities nor public funds.[49] The decision found no free exercise issue and, addressing the establishment clause challenge, dismissed the idea that coercion was used to obtain students for the program.[50] Although the opinion specifically did not overrule *McCollum*,[51] it ignored the "wall of separation" metaphor prominent in that case and in *Everson*. Instead, Justice Douglas emphasized the need for cooperation and accommodation:[52]

> We are a religious people whose institutions presuppose a Supreme Being. We guarantee the freedom to worship as one chooses. We make room for as wide a variety of beliefs and creeds as the spiritual needs of man deem necessary. We sponsor an attitude on the part of governmeent that shows no partiality to any one group and that lets each flourish according to the zeal of its adherents and the appeal of its dogma. When the state encourages religious instruction or cooperates with religious authorities by adjusting the schedule of public events to sectarian needs, it follows the best of our traditions. For it then respects the religious nature of our people and accommodates the public service to their spiritual needs.[53]

---

46. *Id.* at 308 n.1.

47. *Id.* at 309.

48. *Id.* Appellants also alleged that the program was administered in a coercive manner but the allegation did not implicate the schools, so the New York Court of Appeals denied their offer of proof. *Id.* at 311 n.7.

49. *Id.* at 308-09.

50. *Id.* at 311. Justice Douglas was quite literal in his definition of coercion: "No one is forced to go to the religious classrooms and no religious exercise or instruction is brought to the classrooms of the public schools . . . . [T]he school authorities are neutral in this regard . . . ." *Id.* However, since appellants were not allowed to prove their coercion allegation, *see supra* note 48, Justice Douglas's conclusion was not supported by specific evidence.

51. *Id.* at 315.

52. *Id.* at 313-15.

53. *Id.* at 313-14. At the same time, however, the opinion qualifies its accommodation: "There cannot be the slightest doubt that the First Amendment reflects the philosophy that church and state should be separated." *Id.* at 312. But it also tempers that qualification: "The First Amendment, however, does not say that in every and all respects there shall be a separation of church and State." *Id.*

Despite three strong dissents focused primarily on the coercion issue,[54] Justice Douglas found the program an example of constitutionally permissible church-state accommodation.[55]

The *Zorach* opinion did not attempt to harmonize *McCollum*; instead it simply acknowledged, then ignored, the earlier case. The Court in *McCollum*, emphasizing separation of church and state, invalidated a released-time program making use of public classrooms and compulsory education laws. The Court in *Zorach*, emphasizing accommodation of church by state, upheld a released-time program making similar use of compulsory education laws but held off school premises. The two cases not only differ in outcome but reflect seemingly irreconcilable bases of decision.[56]

C. *The* Lemon *Test: Attempt at Resolution*

To resolve the confusion engendered by *McCollum*, *Zorach* and other religion cases, the Supreme Court in *Lemon v. Kurtzman*[57] established a tripartite test for evaluating permissible church-state involvement.[58] The *Lemon* test requires: first, that the statute authorizing the involvement have a secular purpose; second, that it neither inhibit nor advance religion; and third, that it not promote excessive state entanglement with religion.[59] A program that violates any prong of the test is invalid.

The *Lemon* Court tested for unconsititutional entanglement by examining "the character and purposes of the institutions that are benefited, the nature of the aid that the State provides, and the resulting relationship between the government and the reli-

---

54. Justice Black argued that the compulsory school machinery aided the religious organizations in *Zorach* as it did in *McCollum*, exalting the orthodox over the unbelievers. *Id.* at 317-19 (Black, J., dissenting). Both Justices Black and Frankfurter warned of the potential for bitter and divisive community controversy. *Id.* at 320, 323 (Black, Frankfurter, J.J., dissenting). Justice Jackson's dissent found the *Zorach* program unconstitutionally coercive because church attendance was one of only two possible ways of using compulsory education time. *Id.* at 323-24 (Jackson, J., dissenting).

55. *Id.* at 314-15.

56. *See* commentary cited *supra* note 6.

57. 403 U.S. 602 (1971). *Lemon* invalidated statutes which dispensed state money to parochial schools to fund teachers and secular instructional materials involved in secular subjects. The programs required state monitoring to ensure that the statutes' secular and financial requirements were met.

58. The test evolved from standards enunciated in other Supreme Court decisions. *See* Walz v. Tax Comm'n, 397 U.S. 664, 664 (1970); Board of Education v. Allen, 392 U.S. 236, 243 (1968); School Dist. of Abington Township v. Schempp, 374 U.S. 203, 222 (1963).

59. 403 U.S. at 612-13.

gious authority."[60] As articulated and applied in *Lemon*, the entanglement test guards against two evils: the danger of ever-increasing governmental involvement[61] and the potential for political divisiveness.[62] The *Lemon* test attempts to remedy the uncertainty left after *Zorach*[63] by harmonizing strict separation and accommodation.

D. *Applying the* Lemon *Test to Released-Time Programs*

The *Lemon* test was applied to a released-time program in *Smith v. Smith*,[64] a Fourth Circuit case sustaining a Harrisonburg, Virginia released-time program for elementary school students. As in the *Zorach* program, the weekly religion classes at issue in *Smith* were held off school premises, with enrollment conducted by the churches and reported to the schools.[65] School officials coordinated schedules with a local church council and the "small minority" of nonparticipating students remained in the classroom but did not continue their formal instruction.[66]

The court in *Smith* faced the uncomfortable task of applying two distinct and incompatible Supreme Court tests. The *Smith* court found that the Harrisonburg program failed under the second prong of the *Lemon* test, its primary effect being to advance religion.[67] The court conceded that, under the *Lemon* rationale, that failure was sufficient to invalidate the program, but it felt constrained also to measure the plan against *Zorach*.[68] Because the two programs were very similar, the court reasoned that the effect noted by the district court—endorsement of the religion classes by

---

60. *Id.* at 615.

61. The court noted that "the modern governmental programs have self-perpetuating and self-expanding propensities." *Id.* at 624.

62. *Id.* at 623.

63. *See* R. MILLER & R. FLOWERS, *supra* note 9, at 301.

64. 523 F.2d 121 (4th Cir. 1975), *cert. denied*, 423 U.S. 1073 (1976). *Smith* was the first, and *Lanner* the second, released-time case to reach the federal courts since *Zorach*. Two post-*Zorach* released-time cases have reached state supreme courts. Dilger v. School Dist. 24 CJ, 222 Or. 108, 352 P.2d 564 (1960) (sustaining a released-time statute against a void-for-vagueness challenge); State *ex rel.* Holt v. Thompson, 66 Wis. 2d 659, 255 N.W.2d 678 (1975) (sustaining a *Zorach*-like released-time program against first and fourteenth amendment attacks).

65. 523 F.2d at 122.

66. *Id.*

67. *Id.* at 124. The court found that the cooperation between church and school created "an impression of endorsement of the program" by public officials. *Id.*

68. *Id.* In Meek v. Pittenger, 421 U.S. 349 (1975), decided shortly before *Smith*, the Supreme Court reaffirmed *Zorach*, *id.* at 359, and implied that *Zorach* was consistent with the tripartite *Lemon* test. *See id.* The *Smith* court said: "Indeed, *Zorach* illuminates the test." 523 F.2d at 124.

school officials—was "indirect or incidental rather than principal or primary."[69] The court concluded that the *Lemon* primary-effect test could not be taken at face value; religion was not unconstitutionally promoted by the state if the program was similar to that in *Zorach*.[70]

II. *Lanner* v. *Wimmer*

Lanner v. Wimmer[71] involved both free exercise and establishment clause challenges to a released-time program in the Logan, Utah, school district.[72] Students in grades nine through twelve were released daily for an hour-long religion class in a church seminary adjacent to the school.[73] The church supplied religion class registration forms to parents at their request and the students returned the forms to seminary personnel who, in turn, forwarded them to the schools for a record of who was to be released.[74] Attendance at the seminary was taken on slips furnished by the schools and picked up by students from the school's office practice class.[75]

Under the Logan School District program, four kinds of credit were possible for seminary attendance. First, high school students could get two of the "elective credits"[76] required for graduation by attending "Bible history" classes.[77] Second, students received "eli-

---

69. 523 F.2d at 125.

70. The court found the *Lemon* and *Zorach* positions to be incompatible: "If we were to decide this case solely by direct application of the tripartite test . . . , we would be inclined to agree with the district court's overall conclusion that the Harrisonburg release-time program is invalid." *Id.* at 124.

The Wisconsin Supreme Court expressed similar frustration at the task mandated by the Supreme Court, quoting a prior Wisconsin decision which had also applied the religion clause tests: " 'We are bound by the results and interpretations given the First Amendment in these high court decisions. Ours not to reason why; ours but to review and apply . . . .' " State *ex rel.* Holt v. Thompson, 66 Wis. 2d 659, 225 N.W.2d 678, 681 n.5 (1975) (quoting State *ex. rel.* Warren v. Nusbaum, 55 Wis. 2d 316, 322, 198 N.W.2d 650, 653 (1972)).

71. 662 F.2d 1349 (10th Cir. 1981).

72. *See id.* at 1354-55. Any religion may participate, but "the overwhelming use of the program is by persons attending released-time classes offered by a seminary operated by the Church of Jesus Christ of Latter-Day Saints." *Id.* at 1354.

73. *Id.* at 1354.

74. *Id.*

75. *Id.* at 1355.

76. This is the circuit court's term for this type of credit. *Id.* at 1362-63. Students were required to take seven and one-half credits of their choice. This article uses the court's terms for ease in distinguishing among the various types of credit.

77. *Id.* at 1355, 1360. No seminary classes counted towards the eight and one-half credits in specified subjects required for graduation. *Id.* at 1360. No credit was given for classes offered to junior high school students. *Id.* at 1355.

gibility credit"[78] for seminary attendance, helping them qualify for extracurricular activities, honor roll and school elections.[79] Third, seminary attendance helped satisfy the state's compulsory education law[80] and the local school board's requirements for minimum daily attendance,[81] filling "custodial credit"[82] requirements. Fourth, on occasion the school used seminary attendance to satisfy "funding credit"[83] requirements—daily attendance minimums the school must meet to receive state funds.[84] Additionally, the school and the seminary were connected by a school-to-seminary intercom,[85] and a notice box for the seminary was located in the school.[86]

The trial court found that the released-time program was not per se unconstitutional.[87] However, certain aspects of the Logan program—the school's provision and collection of attendance slips[88] and the granting of elective, custodial, and funding credit for seminary attendance—were found unconstitutional.[89] Those contacts implied an impermissible entanglement between church and school.[90]

The Tenth Circuit, citing *Zorach*, agreed that there was no per se establishment clause violation in released-time programs held off school premises.[91] Applying the *Lemon* test as illuminated by *Zorach*,[92] the Court of Appeals found that the Logan program passed both the "secular purpose" and "primary effect" prongs,[93]

---

78. *Id.* at 1362.
79. *Id.* at 1356.
80. UTAH CODE ANN. § 53-24-1 (1978).
81. 662 F.2d at 1356.
82. *Id.* at 1362. The district court did not specifically refer to this type of credit.
83. *Id.* at 1362-63.
84. *Id.* at 1356. A student must attend four classes a day to be counted in the funding formula. *Id.*
85. *Id.* at 1355.
86. *Id.*
87. 463 F. Supp. 867, 878 (D. Utah 1978).
88. *Id.* at 883.
89. *Id.*
90. *Id.*
91. 662 F.2d at 1357. The court acknowledged the continuing scholarly debate on the issue. *Id.*
92. *Id.* at 1357-58. The court quoted the Fourth Circuit conclusion in *Smith*: " '*Zorach* is not inconsistent with the tripartite test. Indeed, *Zorach* illuminates the test. Therefore, it is our duty to follow *Zorach* and to understand the modern test in the light of *Zorach*'s continuing viability.' " *Id.* at 1358 (quoting Smith v. Smith, 523 F.2d 121, 124 (1975), *cert. denied*, 423 U.S. 1073 (1976)).
93. 662 F.2d at 1358. The school board's desire to accommodate the public's spiritual needs was an adequate secular purpose. Citing *Zorach* and *Smith*, the court found that

but found that the school's collection of attendance reports and the awarding of graduation credit for seminary did not pass the "entanglement" test and were therefore unconstitutional.[94]

The court agreed with the trial court that the attendance gathering procedure was not the least entangling alternative;[95] however, the court found the school's provision of uniform attendance slips a matter of efficiency and not unconstitutionally entangling.[96] Like the district court, the Tenth Circuit found that the intercom and the notice boxes also passed the *Lemon* test.[97] The court also upheld the district court's finding that there was no free exercise violation, noting that although the plaintiffs complained of social pressure to participate in the released time program, they did not allege pressure from the school district.[98]

The Court of Appeals gave major emphasis in its opinion to the various types of credit awarded for seminary attendance.[99] Reasoning that the released-time program is like the parochial school in its relationship to the state,[100] the court found "custodial,"[101] "eligibility"[102] and "funding"[103] credit constitutional. It found "elective" (graduation) credit unconstitutional only because the state board of education's policy regarding that credit implicitly required state monitoring of seminary subject matter—an unconstitutional entanglement.[104]

---

merely releasing the students did not unconstitutionally advance or inhibit religion, satisfying the primary effect test. *Id.* at 1359, 1362.

94. *Id.* at 1361-62.

95. *Id.* at 1358-59.

96. *Id.* at 1358.

97. *Id.* at 1359-60.

98. *Id.* at 1360.

99. *See id.* at 1360-63.

100. *Id.* at 1361-62.

101. *Id.* at 1362. The court reasoned that granting custodial credit for seminary (i.e., allowing it to help satisfy the compulsory education law) was tantamount to permitting a released-time program, and released-time programs were constitutionally permissible according to the Supreme Court in *Zorach*. *Id.*

102. *Id.* The court, focusing only on the entanglement issue, simply asserted that awarding eligibility credit for seminary was constitutional because it "requires no state involvement not inherent in Zorach-approved released-time." *Id.*

103. *Id.* at 1362-63. The court found funding credit constitutional because (1) none of the funds go to the seminary, (2) the effect assists the schools but neither enhances nor inhibits religion and (3) a state may decide to allocate its funds on the basis of attendance in all lawful educational programs. The court did not mention *Zorach* here. *Id.* at 1363.

104. *Id.* at 1360-61. The board of education's policy statement provided:

> Credits for work taken in released-time classes should be recognized by public schools only upon an official transcript of credits from the institution conducting such classes, and upon evaluation on the same basis that similar credits from established

The court further asserted that graduation credit could be constitutionally awarded for any seminary class, even those purely denominational, as long as the seminary met certain "secular criteria"[105] and there was no state monitoring of content.[106] Using the entanglement yardstick, the court reasoned that if sectarian classes could count toward graduation in a private religious school, sectarian released-time classes could also be allowed toward graduation in the public school.[107] Thus the Tenth Circuit concluded that the extent of state involvement in church affairs, not class content, determined the constitutionality of the released-time classes.[108]

III. ANALYSIS

A. *Harmonizing* McCollum *and* Zorach

The Tenth Circuit's conclusions in *Lanner* must be evaluated in light of *McCollum* and *Zorach*, given the continuing viability of those cases and the similarity of their issues to those in *Lanner*. A major problem facing the Tenth Circuit in drafting *Lanner* was ascertaining the guidelines provided by the Supreme Court's previous released-time decisions. Although many commentators have attempted a synthesis of the *McCollum-Zorach* holdings, no

---

private high schools are evaluated. (The State board has authorized high schools to recognize not to exceed two units of Bible history for work taken in private seminaries or schools. Such credit may be counted as an "elective" but may not fill any requirements for required instruction in the fields of social studies, English or literature. *No credit is to be given to courses devoted mainly to denominational instruction.*)

Id. at 1360 (emphasis added).

105. *Id.* at 1361. These criteria, applied by the Supreme court to parochial school classes, are that specified subjects must be taught for a required minimum number of hours and that teachers must meet certain qualifications. *See* Board of Education v. Allen, 392 U.S. 236, 245-46 (1968).

106. The court stated:

If the extent of state supervision is only to insure, just as is permitted in the case of church-sponsored full-time private schools, that certain courses are taught for the requisite hours and that teachers meet minimum qualification standards, nothing in either the establishment or free exercise clauses would prohibit recognizing *all* released-time classes *or none*, whether religious in content or not, in satisfaction of graduation requirements. It is when, as here, the program is structured in such a way as to require state officials to monitor and judge what is religious and what is not religious in a private religious institution that the entanglement exceeds permissible accommodation and begins to offend the establishment clause. *See Lemon v. Kurtzman*, 403 U.S. 602 (1972); *see generally Cantwell v. Connecticut*, 310 U.S. 296 (1943); *Espinosa v. Rusk*, 634 F.2d 477 (10th Cir. 1980).

662 F.2d at 1361 (emphasis in original).

107. *Id.*

108. *Id.* at 1361-62; *see supra* note 107.

consensus has been achieved.[109] The significance of *McCollum* and the factors that determine a released-time program's constitutionality are still debated.[110] The individual Justices of the Supreme Court appear to have irreconcilable views on these issues.[111] The confusing decisions on church-state relations have been characterized as "a complicated settlement hammered out at the bargaining table . . . , rather than the product of rational and consistent adjudications."[112]

It seems clear that *McCollum* has not been overruled.[113] The unresolved question, then, is what *McCollum* and *Zorach* together mean. The two cases reflect the historical debate on the meaning of the establishment clause—strict separation versus even-handed support of religion.[114] Despite the use of the wall metaphor in *Everson* and *McCollum*, commentators generally agree that it has proven neither a useful nor an accurate description of the constitutional relationship between church and state.[115] The Court itself

---

109. *See* R. DRINAN, RELIGION, THE COURTS AND PUBLIC POLICY 84-85 (1963); *The "Released Time" Cases Revisited, supra* note 6, at 1202, 1228-34 (1974).

110. R. DRINAN, *supra* note 109, at 80, 85; *The "Released Time" Cases Revisited, supra* note 6, at 1225-26 & n.132.

111. In *McCollum*, for example, Justice Black's majority opinion seemed to preclude all released-time programs, 333 U.S. at 210-12 (impregnable wall between church and state), while Justice Frankfurter's concurrence specifically reserved judgment on other programs not before the Court, *id.* at 231 (Frankfurter, J., concurring). *See also* R. DRINAN, *supra* note 109, at 78-80 (contrasting the various *McCollum* opinions). Nor has the Court's subsequent use of *McCollum* and *Zorach* given a consistent, clear interpretation. *The "Released Time" Cases Revisited, supra* note 6, at 1230-33.

112. Oaks, *Introduction* to THE WALL BETWEEN CHURCH AND STATE 1, 7-8 (1963). At least one researcher has concluded from intra Court papers that complicated negotiations, resulting at times in intentional ambiguity, were a major cause of the inconsistencies. *The "Released Time" Cases Revisited, supra* note 6, at 1233-36.

113. For most commentators, *McCollum* is very much a factor in their analysis of church-school relations. *See, e.g.*, L. TRIBE, *supra* note 21, at 823-25. The Court has continued to cite *McCollum* as viable precedent. *See, e.g.*, Committee for Public Education v. Nyquist, 413 U.S. 756, 772 n.29 (1973); School Dist. of Abington Township v. Schempp, 374 U.S. 203, 216-17 (1963).

114. *See supra* notes 16-18, 29, 53 and accompanying text.

115. Professor Tribe has stated: "[T]he principle evoked by the image of a wall furnishes less guidance than metaphor." L. TRIBE, *supra* note 21, at 820. Another critic has found that "[t]he metaphor is not an aid to thought and it can be a positive barrier to communication." Oaks, *supra* note 113, at 3. Still another has observed: "The wall has done what walls usually do: it has obscured the view. It has lent a simplistic air to the discussion of a very complicated matter." Hutchins, *The Future of the Wall*, in THE WALL BETWEEN CHURCH AND STATE 17, 19 (1963). *But see* Fey, *supra* note 29, at 39-40: "It is only when the institutions, including the financial institutions, of church and state are kept scrupulously separated that civil as well as religious liberty is secure."

The word "separation" does not appear in the religion clauses. The wall metaphor originated in a letter by Jefferson, who did not participate in the drafting of the Bill of

seems to have abandoned the wall metaphor in favor of more flexible guidelines.[116] Since *Zorach*, commentators have not found absolute separation descriptive even of the *McCollum* Court's rationale. Rather than absolutely prohibiting *any* type of church-state involvement, *McCollum* is best interpreted as holding that only a particular kind of contact was per se unconstitutional.[117]

In addition to the "nature of the church-state contact" analysis established by *McCollum*, *Zorach* contributed a second tier to the test of constitutionality under the establishment clause. Justice Douglas' opinion in *Zorach* looked first for a *specific kind of contact*—the use of state funds and classrooms. Not finding that contact, Douglas examined the *quantity or degree* of involvement and found that there were not enough contacts to jeopardize free exercise rights.[118] Using this analysis as a paradigm, the constitutionality of released-time programs apparently depends first on the kind of church-state contacts involved in the program, and second on their overall effect.

In *McCollum* lies the key to understanding just what kind of entanglement is unconstitutional. The use of the public school classrooms in *McCollum* has emerged as a crucial factor,[119] though

---

Rights, since he was in France at the time. *See supra* note 18.

116. The wall was not mentioned in *Zorach*. In Committee for Public Education v. Regan, 444 U.S. 646 (1980), the Court noted its tendency "to avoid categorical imperatives and absolutist approaches at either end of the range of possible outcomes." *Id.* at 662 (dictum). In Tilton v. Richardson, 403 U.S. 672 (1971), the Court looked to "the cumulative criteria developed over many years" to apply the establishment clause. *Id.* at 678. In Lemon v. Kurtzman, 403 U.S. 602 (1971), citing *Zorach* but not *McCollum*, the Court declared: "Our prior holdings do not call for total separation between church and state; total separation is not possible in an absolute sense." *Id.* at 614. School Dist. of Abington Township v. Schempp, 374 U.S. 203, 215, 226 (1973), presents a neutral, no-favoritism approach to religion. *See also* L. TRIBE, *supra* note 21, at 820-21 (the "wall" theory was simply the beginning of the Court's attempts to reconcile both religion clauses).

117. For a collection of "key" factors commentators have pointed to to explain *McCollum* and *Zorach*, see *The "Released Time" Cases Revisited*, *supra* note 6, at 1225 n.133. The list includes: use of public school classrooms, use of compulsory attendance laws, segregation of participants and nonparticipants, close cooperation between church and school. *Id.*

118. Justice Douglas said that the first amendment "studiously defines the manner, the specific ways" in which church and state are separated. *Zorach*, 343 U.S. at 312. But courts and commentators have contemplated the amendment in vain for specific, studious definitions. Justice Douglas himself did not point to any in his opinion. The religion guarantees contain, rather, two general values—free exercise and nonestablishment—which may conflict, as they do in the released-time cases. What actually emerged from the Douglas opinion is a recognition of both of those values: free exercise, *id.* at 313-14, and nonestablishment, *id.* at 312, 314. Justice Douglas then looked at the facts of the program to see if one value suffered at the hands of the other and found that it did not. *Id.* at 314-15.

119. *See* 374 U.S. at 261-62 (Brennan, J., concurring). *See also* R. MILLER & R. FLOWERS, *supra* note 9, at 300 (naming the on-premises factor as the distinguishing element in

for different reasons than the Court gave.[120] Religion teachers in public school classrooms are a tangible locus for attitudes about religion—both the attitudes of the state and those of the students. They become symbols of state authority, serving the same function in the same place as the state teachers, and thus commanding the same respect.[121] Other state contacts with religion do not represent state power in the same way. The release of students in *Zorach*, for example, involved no tangible locus for religious attitudes and no symbol of state authority. The other contacts in *Zorach* also did not transfer state authority to religion the way the use of the classrooms did in *McCollum*. Neither weekly attendance reports made by the churches to the schools[122] nor written requests to the schools from the parents[123] coerce the respect of all students the same way or with the same frequency as a teacher in the hall or classroom. The difference is one of both kind (the state in *Zorach* neither sponsored religion nor invested it with governmental authority) and extent of contact (the state's role in *Zorach* was passive and minimal). This harmonization, though not explicit in *McCollum* or *Zorach*, seems to consistently explain both opinions and can be used in applying them.

The analysis of church-state involvement suggested by *McCollum* and *Zorach* is also consistent with the tone of *Lemon v. Kurtzman*.[124] Although the facts of *Lemon* suggest that constitutional entanglement may be determined by looking only at the

---

McCollum and *Zorach*); L. TRIBE, *supra* note 21, at 823-25 (whether classes are on or off school premises is significant because of public school's prominent place in American life). In *Zorach* Justice Black expressly denied that the expenditures of public funds to support religion were critical; the crucial factor, he said, was the use of compulsory education laws to support religion. *Zorach*, 343 U.S. at 316-19 (Black, J., dissenting).

120. Justice Douglas' specific reasons for validating the program in *Zorach* are not totally clear, but they seemed to center on the fact that no public funds or classrooms were involved. *Zorach*, 343 U.S. at 308-09. Justice Douglas expressly discounted any coercion issue in the absence of either overt force or the pressure of on-campus classes. *Id.* at 311.

121. Justice Brennan suggested something like this in his *Schempp* concurrence. Use of public funds did not per se invalidate the program, but what those funds represented may. Brennan felt that placing the religion teacher in exactly the same position as the public school teacher (through the use of the classrooms) lent to religion the power, respect and authority of the state. At stake was not free exercise but establishment. By that standard, according to Justice Brennan, the *McCollum* program gave much more state power to the church than did the *Zorach* program. *Schempp*, 374 U.S. at 261-63. With the *Zorach* clarification, Justice Black's dissent in *Zorach* is remarkably similar to Justice Brennan's concurrence in *Schempp*. *See supra* note 121.

122. *Zorach*, 343 U.S. at 308.

123. *See id.* at 308 n.1. Registration cards in *McCollum* were obtained from school authorities. 333 U.S. at 207 n.2.

124. 403 U.S. 602 (1971).

form of church-state interaction, the Court actually required a deeper analysis. *Lemon*'s three-pronged test was developed and applied in the context of monetary aid to religion.[125] The impact of monetary aid is probably easier to measure, under the *Lemon* test, than the less tangible kind of support embodied in released-time. At least the "excessive government entanglement"[126] prong of *Lemon* implies more of a quantitative measurement of constitutional interaction, under the *Lemon* facts, than provided by the *McCollum* and *Zorach* guidelines. However, the *Lemon* Court specifically warned against applying the test mechanically:

> This is not to suggest . . . that we are to engage in a legalistic minuet in which precise rules and forms must govern. A true minuet is a matter of pure form and style, the observance of which is itself the substantive end. Here we examine the form of the relationship for the light that it casts on the substance.[127]

The Court found excessive entanglement in *Lemon* only after examining "the cumulative impact of the entire relationship" between the schools and the state.[128] The objective of the Court's analaysis was "to prevent, as far as possible, the intrusion of either into the precincts of the other."[129]

B. Problems with the Lanner *Approach*

In *Lanner v. Wimmer,* the Tenth Circuit faced a combination of facts unlike those in the earlier released-time cases considered by the federal courts.[130] More church-state contacts were at issue in *Lanner,*[131] and the credit question clearly broke new ground. The court's analysis was essentially that used by the Fourth Circuit in *Smith v. Smith*[132] and so resulted in some of the same

---

125. *See supra* notes 58-59.

126. *Lemon,* 403 U.S. at 613.

127. *Id.* at 614. The opinion, though reflecting some of the historical confusion of establishment clause application, seems to suggest that a harmony of *McCollum* and *Zorach* values is intended by the *Lemon* test.

128. *Id.*

129. *Id.*

130. Zorach v. Clauson, 343 U.S. 306 (1952); McCollum v. Board of Education, 333 U.S. 203 (1948); Smith v. Smith, 523 F.2d 121 (4th Cir. 1975). There was no on-premises *McCollum* issue here, and the court easily hurdled the *Zorach-Smith* question of the basic program's constitutionality. *Lanner,* 662 F.2d at 1357.

131. The Logan program released students daily, not weekly as in *Zorach* and *Smith,* and involved the additional contacts of the intercom, the notice box and the four kinds of credit.

132. 523 F.2d 121 (4th Cir. 1975). *See Lanner,* 662 F.2d at 1358.

problems,[133] though the problems are not recognized in *Lanner* as they were in *Smith*. The court either discarded or overlooked other possible approaches, some of them with roots deeper in the Supreme Court's released-time cases.[134]

Lanner is forthright in its accommodation stance, following *Zorach*.[135] But it is an accommodation made facile by the *Lemon* test and untroubled by any reference to *McCollum*. Although the court used the facts in *Zorach* to soften the effect of the *Lemon* test, as the Fourth Circuit did in *Smith*, the concerns of *McCollum* with a specific kind of contact, involving symbolic transfer of authority from state to church, were not addressed at all in *Lanner*. Therefore, the Tenth Circuit's application of the entanglement prong of the *Lemon* test is a formalistic application, concerned mainly with the extent of the contacts, not their effect. That is the very danger the *Lemon* Court warned against.[136]

The *Lanner* court also avoided the deeper probing of *Zorach*'s accommodation mandated both by *McCollum*'s apparent viability[137] and the many questions raised by commentators on the two cases.[138] While the district court was somewhat sensitive to the symbolic impact of church-state contacts, which concerned the Supreme Court in *McCollum* and *School District of Abington Township v. Schempp*,[139] the Court of Appeals applied a more literal reading of *Zorach* and made only one discounting reference to symbolic identification.[140]

The court considered each aspect of the program in turn[141] and concluded that, except for attendance gathering and the existing system of awarding graduation credit, neither the individual contacts nor their cumulative effect violated the establishment clause.[142] Because the court did not acknowledge any transfer-of-

---

133. *See supra* note 101. The court in *Lanner*, like the court in *Smith*, applied the three-pronged *Lemon* test as illuminated by *Zorach*. *See supra* note 122.

134. Several authorities list the various theories the Supreme Court has used in applying the religion clauses. *See* L. TRIBE, *supra* note 21, at 819-23; Oaks, *supra* note 113, at 3-4; Clark, *Comments on Some Policies Underlying the Constitutional Law of Religious Freedom*, 64 MINN. L. REV. 453, 455-57 (1980); Pepper, *Reynolds, Yoder, and Beyond: Alternatives for the Free Exercise Clause*, 1981 UTAH L. REV. 309, 345-76.

135. *Lanner*, 662 F.2d at 1352-54.

136. *See supra* note 128 and accompanying text.

137. *See supra* note 114.

138. *See supra* notes 110-13 and accompanying text.

139. *See* Lanner v. Wimmer, 463 F. Supp. 867, 872-74 (D. Utah 1978).

140. *Lanner*, 662 F.2d at 1360.

141. *Id.* at 1357-63.

142. *Id.* at 1359.

authority issue, however, its assessment of the cumulative effects of the Logan released-time program was an assessment of form, not substance. The intercom, the notice box, the daily release of students, the eligibility, custodial and funding credit, taken together, suggest the danger that "[t]he operation of the state's compulsory education system . . . [unconstitutionally] assists and is integrated with the program of religious instruction."[143]

The court's failure to recognize establishment violations inherent in symbolic identification also brought it to the fallacious analogy of parochial schools and released time.[144] The fact that parochial school classes count towards graduation, given certain minimum standards, does not mean that released-time religion classes can fill public school requirements, given the same minimum standards. The equation ignores the differences between the parochial school and the released-time class. The parochial school system is separate from the state and therefore its program is not subject to *Lemon*-type scrutiny, as is the public school program. Released-time students are public school students, meeting requirements set by representatives of the public, and released-time programs are simply "accommodated" by those representatives. Parochial school students are by choice apart, in a religiously homogeneous environment. Public school students, by contrast, are a microcosm of the larger society and are guaranteed an environment in which a single religion is not favored by the state. The private status of the parochial school insulates it from the dangers of church-state entanglement flagged in *Lemon*.[145] On the other hand, granting public school credit for denominational released-time classes imbues the classes with public school status—the kind of symbolic identification prohibited by *McCollum*.

After accepting the analogy between parochial schools and released-time, the *Lanner* court assumed that the *Lemon* entanglement test could be applied mechanistically in the same way to each.[146] The court reasoned that there is no danger of excessive entanglement in awarding credit for parochial school classes; therefore, there must be none for the same award to released-time classes.[147] But *Lemon* does not allow a blanket application of the test;

---

143. McCollum v. Board of Education, 333 U.S. 203, 209 (1948).

144. *See supra* notes 101 and 108 and accompanying text.

145. For example, the potential for political divisiveness and the propensity of any government intrusion to enlarge itself. *See supra* notes 62-63 and accompanying text.

146. *See Lanner*, 662 F.2d at 1361.

147. *Id.*

the rules do not work like a precision "minuet."[148] Unconstitutional entanglement under *Lemon* depends on several factors, including the character of the institution benefited, the nature of the state aid, and "the resulting relationship between government and religious authority."[149] These factors, which correspond well to the symbolic transfer prohibitions of *McCollum*, should have tempered the Tenth Circuit's application of the *Lemon* test in *Lanner*.

The Tenth Circuit's conclusion that awarding credit for released-time classes does not threaten excessive entanglement seems simply wrong. Granting public school credit for purely denominational classes suggests public school sponsorship of a religious activity. It equates the religion class with the public school class. The state transfer of authority to church is inevitably coercive because unlike the parochial school student, the nonparticipating public school student is subjected to a religious identification and norm not of his choosing.

C. *The* McCollum-Zorach *Alternative*

In applying the entanglement test, the *Lanner* court looked at the number and complexity of the contacts between church and state that a particular facet of the program involved. The court did not specifically consider what those contacts might connote. Harmonizing *Zorach* and *McCollum* would seem to require this type of analysis.[150] If excessive entanglements were measured in part by whether or not a contact imbues church with state power or authority, more reliable limits to accommodation could be set. Awarding graduation and eligibility credit for seminary is clearly a new kind of accommodation, exceeding the schedule adjustment accommodations allowed in *Zorach* and *Smith*. Efficiency and convenience do not justify the awarding of credit; the dangers to free exercise and nonestablishment should positively prohibit it. Even with no state monitoring of the classes, the credit awards are unconstitutional under the harmonized *McCollum-Zorach* holdings.

IV. Conclusion

McCollum, like the religion clauses, is a historical fact. Combined with *Zorach*, it is a source of puzzlement for many courts that choose to ignore it as did the Tenth Circuit in *Lanner v.*

148. *Lemon*, 403 U.S. at 614.
149. *Id.* at 615.
150. *See supra* notes 118-24 and accompanying text.

Wimmer. The Supreme Court in *Zorach*, however, acknowledged the lesson of *McCollum* that a certain kind of church-state involvement violates the establishment clause. In *Lemon* the Court underscored the message that impermissible contacts are a matter of substance, not form. If the substance—the symbolic effect of the contact—is not recognized, a purely formalistic assessment of entanglement, like that applied in *Lanner*, seems all that is required. *Lanner* suggests that permissible accommodation can be expanded to include any church-state relationship in which the state is a nonintrusive party. This approach is clearly a new direction in establishment clause construction, one not supported by a close reading of the Supreme Court's released-time cases.

ELIZABETH A. SHAW

Equitable Reformation of Long-Term Contracts—The "New Spirit" of *ALCOA*

Contracts supply a measure of predictability in an uncertain world. By entering into contracts, parties ensure predictable future sources of supply and secure buyers for their products. The very uncertainty that promotes the formation of contracts, however, may cause the obligations undertaken to become disastrously burdensome. When that occurs, courts have used the well-established doctrines of impracticability and frustration to rescue a promisor from the consequences of his promise. Because contractual predictability is so basic to our society, however, courts have been reluctant to release parties from their contractual obligations or to rewrite contractual terms. Thus, they have used these escape doctrines as narrow exceptions, allowing restricted remedies.

The recent case of *Aluminum Company of America v. Essex Group, Inc. ("ALCOA")*,[1] rejected the traditional reticence of courts to alter contractual obligations. Noting that the usual remedies of rescission and restitution may not be responsive to problems created by long-term contracts,[2] the *ALCOA* court asserted that a "new spirit" of equity should allow courts to adjust the terms of long-term contracts to meet changing circumstances.[3] Additionally, *ALCOA* suggested that the doctrine of mistake may be applied to relieve a party to a long-term contract from unexpectedly burdensome obligations.

I. Background

A. *The Escape Doctrines*

Courts have long recognized that parties may be released from contractual obligations when their expectations are upset by some unknown or unexpected condition. When the event causing the dislocation arises after the contract was made, the courts have generally granted relief under the doctrines of frustration and imprac-

1. 499 F. Supp. 53 (W.D. Pa. 1980).
2. *Id.* at 78-79.
3. *Id.* at 89.

ticability.[4] When the dislocation is caused by the parties' misapprehension of some critical fact[5] existing at the time they made the contract,[6] the courts have applied the doctrine of mistake.

The doctrine of frustration focuses on the contract's loss of value to one of the parties.[7] In keeping with common law notions of the sanctity of contractual terms,[8] frustration is applied only when the principal purpose for entering into the contract has been "substantially frustrated."[9] Most courts have interpreted the sub-

---

4. *See, e.g.*, Hass v. Pittsburgh Nat'l Bank, 495 F. Supp. 815 (W.D. Pa. 1980) (holding supervening rise in interest rates insufficient to constitute frustration of contract); MA v. Community Bank, 494 F. Supp. 252 (E.D. Wis. 1980) (requiring the defendant to show that the principal purpose of the party was substantially frustrated after the contract was made in order to prove commercial frustration); Associated Grocers of Iowa Co-op, Inc. v. West, 297 N.W.3d 103 (Iowa 1980) (impracticability of performance "refers to extraordinary circumstances which could not have been anticipated"); Missouri Pub. Serv. Co. v. Peabody Coal Co., 583 S.W.2d 721 (Mo. 1979) (commercial impracticability is "applicable upon occurrence of supervening, unforeseen event"); *see also* J. CALAMARI & J. PERILLO, HANDBOOK OF THE LAW OF CONTRACTS 498-99 (1977).

5. Reid v. Graybeal, 437 F. Supp. 24 (W.D. Okla. 1977) (refusing to enforce automobile accident settlement because mistake concerning limits of insurance coverage went to essence of the agreement); *see* Beachcomber Coins, Inc. v. Boskett, 166 N.J. Super. 442, 400 A.2d 78, 79 (1979) (genuineness of collectable coin held basis of transaction).

6. *See, e.g.*, Raphael v. Booth Memorial Hosp., 67 A.D.2d 702, 412 N.Y.S. 2d 409, 411 (App. Div. 1979) (denying plaintiff's request to cancel settlement agreement because of agreement's adverse effect on another action on grounds that mistake did not concern a fact existing at the time the contract was entered); Turderville v. Upper Valley Farms, Inc., 616 S.W.2d 677, 678 (Tex. Civ. App. 1981) (requirement of mutual mistake that the parties believe "in the present existence of a thing . . . that does not exist" satisfied by mistakes concerning ownership of land); J. CALAMARI & J. PERILLO, *supra* note 4, at 499.

7. *See* Guthrie v. Times-Mirror Co., 51 Cal. App. 3d 879, 124 Cal. Rptr. 577, 583 n.7 (1975) (anti-trust action against buyer of publishing company requiring divestiture after purchase did not destroy the value of consideration given to seller); Twin Harbors Lumber Co. v. Carrico, 92 Idaho 373, 442 P.2d 753 (1968) (frustration doctrine held not applicable where promised consideration is money and promisor becomes insolvent, because value of money to promissee is not diminished); Krell v. Henry, [1903] 2 K.B. 740, 747-55 (C.A.) (purpose of contract for rental of window frustrated when coronation cancelled). *See generally* RESTATEMENT OF CONTRACTS § 288 (1932); 6 A. CORBIN, CORBIN ON CONTRACTS § 1361-63 (rev. ed. 1962).

8. Under the original common law doctrine of impossibility, relief was unavailable unless performance was objectively impossible. *See* Taylor v. Caldwell, 122 Eng. Rep. 309, 312-15 (K.B. 1863) (allowing rescission of contract to rent music hall after destruction by fire). In *Taylor*, the court stated the general rule that "where there is a positive contract to do a thing . . . the contractor must perform it or pay damages for not doing it although in consequence of unforeseen accidents, the performance of his contract has become unexpectedly burthensome [sic] or even impossible." *Id.* at 312; *see also* Lord v. Wheeler, 67 Mass. (1 Gray) 282, 283 (1854) (releasing contractor from promise to repair roof of hotel that had been burned to the ground on basis of impossibility).

9. *See* Hass v. Pittsburgh Nat'l Bank, 495 F. Supp. 815, 819 (W.D. Pa. 1980) (dramatic rise in other interest rates did not substantially frustrate agreement to pay passbook rate of interest on settlement funds of class action); MA v. Community Bank, 494 F. Supp. 252, 257 (E.D. Wis. 1980) (plaintiff's move from Wisconsin to California and then to New York did

stantiality requirement as requiring "total or nearly total frustration,"[10] and have generally not recognized lack of profitability as a proper basis for the doctrine's application.[11]

The concept of impracticability, now codified in the Uniform Commercial Code[12] with respect to sales contracts, focuses on an excessive burden to a party in performing his promise.[13] Generally, the courts have required that the performance have become so unreasonably difficult that the agreement is "commercially senseless."[14] As with the frustration doctrine, the courts have not awarded relief simply on the basis that the contract has become unprofitable to one of the parties.[15]

Courts have applied the doctrine of mistake when a misapprehension of the quantity[16] or character[17] of the subject of the con-

---

not substantially frustrate purpose of "savings certificate of deposit" agreement by making positive identification of payee difficult); RESTATEMENT (SECOND) OF CONTRACTS § 285 (Tent. Draft No. 9, 1974).

10. *See, e.g.*, Lloyd v. Murphy, 25 Cal. 2d 48, 153 P.2d 47, 50 (1944) (government ban on sale of automobiles did not destroy value of performance in lease for car dealership); Castagno v. Church, 552 P.2d 1282, 1283 (Utah 1976) (inability of seller to provide water rights did not amount to "total or nearly total destruction of the purpose" of real estate contract). *But see* West Los Angeles Inst. for Cancer Research v. Mayer, 366 F.2d 220, 255 (9th Cir. 1966) (purpose of stock transaction frustrated because of adverse tax ruling on stock sale).

11. *See* Acme Markets Inc. v. Dawson Enters., 253 Md. 76, 251 A.2d 839, 846 (1969); Perry v. Champlain Oil Co., 101 N.H. 97, 98, 134 A.2d 65, 75 (1957) (refusing to allow landlord to rescind lease where rent based on lessee's decreased sales stating: "courts have been careful not to find commercial frustration if it would only result in allowing party to withdraw from a poor bargain").

12. U.C.C. § 2-615 (1972).

13. *See* Mineral Park Land Co. v. Howard, 172 Cal. 289, 145 P. 458, 460 (1916) (obligation to quarry all the gravel in a pit became impracticable when water table was reached, increasing cost of recovery ten-fold); *cf.* F. J. Busse, Inc. v. Department of Gen. Servs., 47 Pa. 539, 408 A.2d 478, 581 (1979) (added expense of removing silt and mud left by hurricane did not make contractor's performance impracticable).

14. *See* Transatlantic Fin. Corp. v. United States, 363 F.2d 312, 315 (D.C. Cir. 1966) (denying merchant ship owner relief where voyage costs increased by closing of Suez Canal, stating that courts must decide when the "community's interest in having contracts enforced . . . is outweighed by the commercial senselessness of requiring performance"); Eastern Air Lines, Inc. v. Gulf Oil Corp., 415 F. Supp. 429, 438 (S.D. Fla. 1975) (requiring delievery of jet fuel at price drastically below market, stating that before impracticability will apply, "it must be positively unjust to hold the parties bound"); International Paper Co. v. Rockefeller, 161 A.D. 180, 146 N.Y.S. 371 (1914) (holding land owner obligated to sell timber for $5.50 per cord although fire destroyed all timber except that which could be harvested for $20 per cord).

15. *See supra* notes 11 and 14; *see also* U.C.C. § 2-615 comment 4 (1972) ("A rise or collapse of the market [is not] in itself a justification [for non-performance], for that is exactly the type of risk which business contracts made at fixed prices are intended to cover").

16. *See, e.g.*, Schwaderer v. Huron-Clinton Metropolitan Auth., 329 Mich. 258, 45

tract results in an unjust enrichment of one of the parties.[18] Generally the mistake must be mutual[19] and concern a basic assumption of the contract about a fact existing at the time the contract was made.[20] The requirement of mutuality has sometimes been relaxed, however, to achieve equitable results.[21]

B. Allocation of Risk

Because anyone who finds himself on the wrong side of a burdensome contract can truthfully assert that he made a mistake, his purpose is frustrated and the contract is, from his point of view, commercially senseless; liberal application of the escape doctrines would destroy the utility of contracts. The courts, therefore, have not relieved parties if the risk of the unforeseen occurrence or mistake was allocated to or assumed by the party seeking relief.[22]

Under the doctrines of frustration and impracticability, if the risk involved was not explicitly allocated by the contract, the fore-

---

N.W.2d 279 (1951) (mistake as to acreage of brush to be cleared would result in unjust enrichment to landowner if contractor not paid extra expenses); RESTATEMENT (SECOND) OF CONTRACTS § 300 illustration 1 (Tent. Draft No. 10, 1975) (when parties make a contract for sale of 100 acre tract, at $1,000 per acre; $100,000 total price, court may imply a term allowing purchaser to pay only $90,000, where tract only contains 90 acres).

17. *See* Beachcomber Coins, Inc. v. Boskett, 166 N.J. Super. 442, 400 A.2d 78 (1979) (counterfeit coin sold as genuine collectible); Sherwood v. Walker, 66 Mich. 568, 33 N.W. 919 (1887) (mistake as to whether prize cow was barren or with calf); Wood v. Boynton, 64 Wis. 265, 25 N.W. 42 (1885) (diamond sold as topaz).

18. *See* G. PALMER, MISTAKE AND UNJUST ENRICHMENT 53 (1962): "There is no simple formula for testing relievable mistake. For the most part relief is given only if the mistake produced an unexpected gain to one party, whose claim to retain it finds less than full support in the protection of his contract expectations."

19. Cohen v. Merrill, 95 Idaho 99, 503 P.2d 299, 304 (1972) (rescission of real estate contract conveying more property than seller expected unavailable unless mistake was mutual). *Accord* RESTATEMENT OF CONTRACTS § 503 (1932); *cf.* RESTATEMENT (SECOND) OF CONTRACTS §§ 153-54 (1979) (providing relief for unilateral mistake if risk of mistake was not allocated to the party and the mistake would make enforcement unconscionable).

20. *See supra* notes 5-6. *See generally* J. CALAMARI & J. PERILLO, *supra* note 4, at 300 (mistake must be as to a "basic assumption upon which both parties acted").

21. With respect to the requirement of mutuality of the mistake, see Franklin Nat'l Bank v. Austin, 99 N.H. 59, 104 A.2d 742 (1954) (sole bidder at foreclosure auction released from $1,400 bid where the actual value of the truck was only $200). *See also* Rabin, *A Proposed Black-Letter Rule Concerning Mistaken Assumptions in Bargain Transactions*, 45 TEX. L. REV. 1273, 1277-79 (1967) (arguing that the distinction between unilateral and mutual mistakes should be abolished).

22. *See* Publicker Indus. v. Union Carbide Corp., 177 U.C.C. Rep. Serv. 989, 992 (E.D. Pa. 1975) (finding existence of ceiling price in contract indication of intentional allocation of risk of extraordinary price increases); Trans-State Investments, Inc. v. Deive, 262 A.2d 119, 121 (D.C. 1970) (patron of health spa assumed risk of inability to use facilities by agreeing to pay membership fee whether facilities were used or not); RESTATEMENT (SECOND) OF CONTRACTS, *supra* note 9, § 285 comment a, § 286 comment b.

seeability of the intervening event is a prime indicator of an implicit risk allocation.[23] Courts reason that the disadvantaged party might have protected himself from the consequences of a particular contingency if it was foreseeable.[24]

Under the doctrine of mistake, the risk of the mistake is generally allocated to the party on which it falls if the parties made their agreement "consciously ignorant" of the facts.[25] Relief is also denied when the parties' agreement was based on an inaccurate prediction of future events.[26] Furthermore, the escape doctrines generally have not been applied when the party seeking relief is responsible for his misfortune.[27]

C. *Remedies Under the Escape Doctrines*

The remedy usually invoked under the escape doctrines is re-

---

23. *See, e.g.*, Eastern Air Lines, Inc. v. Gulf Oil Corp., 415 F. Supp. 429, 441 (S.D. Fla. 1975) (increased oil costs foreseeable; relief denied under U.C.C. § 2-615 doctrine of impracticability); Robberson Steel, Inc. v. J. D. Abrams, Inc., 582 S.W.2d 558, 564 (Tex. Civ. App. 1979) (failure to provide explicitly for foreseeable delays in steel supply contract tacitly assigned risk of delay to supplier). *See generally* J. CALAMARI & J. PERILLO, *supra* note 4, at 500.

24. *See supra* note 23; *cf.* Transatlantic Finance Corp. v. United States, 363 F.2d 313 (D.C. Cir. 1966). In *Transatlantic*, the court stated that "foreseeability or even recognition of a risk does not necessarily prove its allocation." *Id.* at 318. Nevertheless the court found that the foreseeability of political unrest in the Middle East was indicative of the shipper's willingness to assume the abnormal risk of the closing of the Suez Canal. *Id.* at 319. *But see* West Los Angeles Inst. for Cancer Research v. Mayer, 366 F.2d 200, 225 (9th Cir. 1966) (finding that even though an adverse tax ruling on a contract for the sale of stock was foreseeable, it did not prevent rescission of the agreement on grounds of frustration).

25. 3 A. CORBIN, CORBIN ON CONTRACTS §§ 598, 605 (rev. ed. 1962). Conscious ignorance means that both parties understood from the beginning that the facts were not known to them. *See* Backus v. Maclaury, 278 A.D. 504, 106 N.Y.S.2d 401 (1951) (sale of sterile bull calf not voidable where both parties knew that the bull was too young to test for fertility); Harvey v. Robey, 211 Va. 234, 176 S.W.2d 673, 675 (1970) (settlement with insurance company for known injury not voidable on grounds that injury was more serious than plaintiff thought); Wood v. Boynton, 64 Wis. 265, 25 N.W. 42 (1885) (erroneous guess that stone was topaz, not diamond, not ground to void contract for mistake).

26. Shear v. National Rifle Ass'n, 606 F.2d 1251 (D.C. Cir. 1979) (erroneous belief of broker and seller that proposed sale would be approved by management not mistake justifying relief); Japhe v. A-T-O Inc., 481 F.2d 366, 370 (5th Cir. 1973) (erroneous estimate of business prospects of purchased company not mistake justifying relief); Leasco Corp. v. Taussing, 473 F.2d 777, 781 (2d Cir. 1972) (erroneous prediction of future earnings not mistake justifying relief); Moffat Tunnel Improvement Dist. v. Denver & S.L. Ry., 45 F.2d 715 (10th Cir. 1930) ("inaccurate prophesies" of cost and completion time of railroad tunnel not mistake justifying relief); RESTATEMENT (SECOND) OF CONTRACTS, *supra* note 16, § 293 comment a ("A party's prediction or judgment as to events to occur in the future, even if erroneous, is not a 'mistake' as that word [is] defined here").

27. *See* Lyons v. Keith, 316 S.W.2d 785 (Tex. Civ. App. 1958); RESTATEMENT (SECOND) OF CONTRACTS, *supra* note 9, § 285 comment b, § 286.

scission,[28] sometimes coupled with restitution to avoid unjust enrichment.[29] Reformation is an alternative remedy to rescission, but courts have largely restricted the reformation of contracts to situations where the party seeking relief can prove by clear and convincing evidence that a written contract does not conform to the actual agreement.[30] Courts usually have not reformed contracts to relieve parties burdened by changing circumstances because basic contract principles prohibit binding parties to obligations to which they have not assented.[31] There have been instances, however, where reformation has been used outside of these traditional boundaries in order to reach just results.[32] Those instances include the use of reformation to make the contract conform to what the parties would have done had they known all the facts.[33] In this

---

28. *See* Butterfield v. Byron, 153 Mass. 517, 27 N.E. 667 (1891) (contractor released from obligation to build house where house destroyed by fire after partial construction); *see also* RESTATEMENT (SECOND) OF CONTRACTS, *supra* note 9, §§ 281, 285 (discharge by supervening impracticability or frustration); RESTATEMENT (SECOND) OF CONTRACTS, *supra* note 16, §§ 294-95 (mistake of fact makes contract voidable). *But see* U.C.C. § 2-615 (1972) (requiring allocation of production where only part of seller's performance rendered impracticable).

29. Auger v. Champdelaine, 106 N.H. 242, 209 A.2d 710, 711 (1965) (where cities' refusal made performance of real estate vendor's promise to supply city water impossible, vendor must restore consideration received for promise); Libman v. Levenson, 236 Mass. 221, 128 N.E. 13 (1920) (buyer who contracted to purchase building damaged before delivery awarded restitution of downpayment); Panto v. Kentucky Distillers & Warehouse Co., 215 A.D. 511, 214 N.Y.S. 19 (1926) (buyer released from obligation to buy whisky awarded return of downpayment). *See generally* J. CALAMARI & J. PERILLO, *supra* note 4, at 508-509. Comment, *Apportioning Loss After Discharge of A Burdensome Contract: A Statutory Solution*, 69 YALE L.J. 1055 (1960).

30. *See* Day v. Fireman's Friend Ins. Co., 67 F.2d 257, 258 (5th Cir. 1933) (evidence of parties' understanding held sufficiently clear and convincing to warrant reformation of insurance contract to include mortgage clause); Oliver-Mercer Elec. Coop. v. Fisher, 146 N.W.2d 346, 355 (N.D. 1966) (evidence of parties' intent to terminate contract for electricity when electricity no longer needed by defendants not sufficiently clear to allow reformation); RESTATEMENT (SECOND) OF CONTRACTS, *supra* note 16, § 297.

31. Tallackson Potato Co. v. MTK Potato Co., 278 N.W.2d 417, 424 (N.D. 1979) (refusing to reapportion damages award between the parties, stating that courts will not make new contracts by reforming them in a manner never considered by the parties). *But cf.* U.C.C. §§ 2-204, -207(3), -208 (1972) (directing court to fill in the blanks of price, timing, manner of payment, etc., where the parties have manifested an intent to form a contract but have left gaps in their agreement). *See generally* J. CALAMARI & J. PERILLO, *supra* note 4, at 311-313 (overview of reformation for mistake).

32. *See* Thomas v. Satfield Co., 363 Mich. 111, 108 N.W.2d 907, 910 (1961) (reforming executory lease because of breach of fiduciary duty by lessor but refusing to find actual fraud); Schwaderer v. Huron-Clinton Metropolitan Auth., 329 Mich. 258, 45 N.W.2d 279 (1951) (reforming contract to avoid unjust enrichment); 3 G. PALMER, THE LAW OF RESTITUTION § 13.9 (1978).

33. *See* National Presto Indus. v. United States, 338 F.2d 99, 110 (Ct. Cl. 1964) (reforming contract to require that parties share costs of additional equipment required for

Comment, reformation of the agreement will be called "equitable reformation" in order to avoid confusion with the more traditional remedy of reformation of the writing.

D. Long-Term Contracts

The restricted application of the escape doctrines and the traditional limitations on the reformation remedy have been criticized as failing to address unique problems inherent in long-term contracts.[34] Long-term[35] obligations tend to be riskier than short-term obligations[36] because the longer the term of the contract, the more likely it is that unforeseen events will intervene to make the promised performance burdensome. The length of the contract also increases the promisor's exposure to severe continuing loss once an adverse unforeseen event occurs.

Because of those problems, long-term contracts frequently contain an adjustment clause to create flexibility.[37] The clause may call for arbitration, termination or renegotiation or contain a price escalation formula.[38] The contract may specify that the clause is to be activated by the occurrence of various contingencies or by other criteria such as specified time intervals.[39] Despite elaborate planning, unanticipated events may occur, causing one party to claim that the contract is not as flexible as intended or that the contract

---

performance of contract).

34. *See* 6 A. CORBIN, *supra* note 7, § 1360 (severely imbalanced relationship in long-term contracts should be adjusted); Macauley, *Elegant Models, Empirical Pictures, and the Complexities of Contract*, 11 L. & SOC'Y REV. 407 *passim* (1977) (advocating extrajudical means of dealing with contract disputes); MacNeil, *Contracts: Adjustment of Long-Term Economic Relations Under Classical, Neoclassical and Relational Contract Law*, 72 NW. U.L. REV. 854 *passim* (1978) (concluding that traditional contract law is inadequate to deal with complex contracts of long duration because of the flexibility they require); Speidel, *Excusable Nonperformance in Sales Contracts: Some Thoughts about Risk Management*, 32 S.C.L. REV. 241, 274-75, 277-78 (1980) (arguing that long-term supply contracts are a class of contract requiring judicial adjustment).

35. The distinction between long-term contracts and other term contracts cannot be precisely defined because it obviously involves a continuum.

36. MacNeil, *supra* note 34, at 859-61. *See* Iowa Elec. Light & Power Co. v. Atlas Corp., 467 F. Supp. 129, 140 (N.D. Iowa 1978) (price of uranium rose approximately 58% over life of 4-year contract causing supplier loss of $3,097,312); *see also* Wall St. J., Nov. 10, 1972, at 28, col. 1 (describing Westinghouse's potential liability of over 2 billion dollars on uranium contract, where price rose from $9.50 a pound in 1975, to $40 a pound in 1977).

37. MacNeil, *supra* note 34, at 865.

38. For a discussion of those approaches to contract flexibility and how they relate to the present legal philosophy of contract interpretation, see *id.* at 866-73.

39. MacNeil states that the adjustment clauses of long-term contracts often use a standard that is not controlled by either party. *Id.* at 866-68.

has become unduly burdensome.[40]

If the court allows the contract to stand, without adjustment, in the face of circumstances that have drastically changed the cost or value of performance, a commercially imbalanced exchange will occur. A party may be forced to buy or sell products at a price completely unrelated to prevailing economic conditions. Indeed, strict enforcement of the contract may award one party an undeserved windfall gain and result in the bankruptcy of the other party.[41] On the other hand, if the court allows a party to escape his promise, the other party may be deprived of the benefit of the bargain and rescission may do nothing more than shift the windfall gain or loss from one party to the other.[42] This problem has caused at least one judge to plead with litigants to negotiate and settle their dispute rather than make the court decide the case under traditional concepts of contract law.[43]

The risks that are inherent in long-term contracts have been particularly acute in the rapidly changing economic climate of recent years. Governmental regulation and increased costs of energy and other commodities have combined to drastically upset elaborately structured long-term contracts for the supply of coal,[44] jet fuel[45] and uranium.[46] Problems created in a long-term service con-

---

40. *See, e.g.*, Eastern Air Lines, Inc. v. Gulf Oil Corp., 415 F. Supp. 429, 439 (E.D. Fla. 1976) (price escalation clause linked to price of "West Texas Sour" oil rendered ineffective by government regulation).

41. *See infra* notes 89-97 and accompanying text.

42. *Id.*

43. Judge I. Martin Wekselman of the Allegheny County Common Pleas Court stated:

> The fiscal well-being, possibly the survival, of one of the world's corporate giants is in jeopardy. Likewise, the future of thousands of jobs.
>
> Any decision I hand down will hurt somebody and because of that potential damage, I want to make it clear that it will happen only because certain captains of industry could not together work out their problems so that the hurt might have been held to a minimum.

N.Y. Times, Feb. 11, 1977, at D-1, D-10.

44. *See* Georgia Power Co. v. Cimmaron Coal Corp., 526 F.2d 101 (6th Cir. 1975). In *Georgia Power*, the market price of coal rose to $32 per ton while the price to the power company under the contract escalation formula was only $10 per ton. The coal supplier demanded that the selling price be increased to $27.50 per ton or that the dispute be arbitrated relying on arbitration clauses found in the contract. *Id.* at 103-04. The power company argued that the agreement to arbitrate did not include that situation. The court found for the coal supplier, citing the "strong federal policy in favor of arbitration." *Id.* at 106.

45. *See supra* note 40.

46. *See* Wall St. J., Sept. 14, 1977, at 23, col. 1. The potential liability of Westinghouse Corp. was $2.6 billion in multiple suits for breach of uranium supply contracts. Westinghouse had contracted to deliver approximately 80 million pounds of uranium to power companies in the United States and Europe. Deliveries were to be made in periodic installments until 1993 at a fixed price of about $10 per pound. Uranium prices rose to over $26 per

tract by the dramatic rise in energy costs were presented in the recent case of *Aluminum Company of America v. Essex Group, Inc.*[47]

II. *ALCOA v. Essex Group, Inc.*

In *ALCOA*, the Federal District Court for the Western District of Pennsylvania[48] rewrote the price escalation formula in a long-term aluminum smelting contract to require that the buyer and seller share the burden of an unexpected increase in production costs due to rising energy expenses. ALCOA and Essex had entered a twenty-year contract[49] that required Essex to deliver aluminum ore to ALCOA,[50] which ALCOA would smelt and hold in a molten state for Essex to pick up. Essex would then make wire out of the molten aluminum. The contract contained an elaborate price escalation clause that set the price of molten aluminum to Essex according to three variable factors: a construction cost index,[51] ALCOA's average hourly labor cost, and the wholesale price index for industrial commodities ("WPI-IC").[52] The third factor was chosen as an objective measure of ALCOA's nonlabor production costs after both parties satisfied themselves that the WPI-IC's historical performance matched the actual cost increases incurred by ALCOA within reasonable limits.[53] The parties expected that ALCOA's profit would vary between one and seven cents per pound.[54]

Approximately eight years after the signing of the contract, the OPEC oil embargo and subsequent energy crisis dramatically increased the cost of electricity—ALCOA's principal nonlabor ex-

---

pound in 1975 and went as high as $40 per pound in 1976. Because Westinghouse found it necessary to resort to the open market to meet its obligations, it notified the power companies in September, 1975, that it would not perform. *Id.* At least 17 law suits ensued, all of which were settled by April, 1981. *See* Wall St. J., Apr. 16, 1981, at 16, col. 3.

47. 499 F. Supp. 53 (W.D. Pa. 1980).

48. Jurisdiction of the court was based on diversity of citizenship and the court applied Indiana law in deciding the case. *Id.* at 55, 78.

49. *Id.* at 57.

50. *Id.* at 58. The aluminum ore was supplied to Essex through an ALCOA subsidiary. *Id.* at 83-84.

51. This index is published by the Engineering News Record and is called the Engineering News Record Construction Cost—20 Cities Average Index. *Id.* at 58.

52. The Wholesale Price Index—Industrial Commodities ("WPI-IC") is published by the United States Bureau of Labor Statistics. *Id.*

53. *Id.* at 69. ALCOA employed the eminent economist Alan Greenspan to aid in the development of the escalation formula. *Id.* at 58.

54. *Id.* at 58.

pense in aluminum conversion—and ALCOA found itself losing money on the contract at an increasing rate.[55] In 1978, ALCOA lost approximately $9,000,000 with nine years remaining on the contract.[56] ALCOA sued Essex, seeking reformation or equitable adjustment of the contract under the doctrines of mutual mistake of fact, frustration of purpose and commercial impracticability.[57]

The court found that the parties' mistaken belief that the WPI-IC index was suitable as an objective measure of ALCOA's nonlabor production costs was a mutual mistake that justified relief.[58] The court found that the parties recognized the need for flexibility in their agreement to make the long-term contract workable[59] and that the parties' reliance on the WPI-IC as a reliable indicator of ALCOA's nonlabor production costs constituted a mistake of existing fact.[60] Although the court conceded that the "mistake [in this case] is not wholly isolated from prediction of the future,"[61] it characterized the parties' mistake as "essentially a present actuarial error" and not a "naked prediction" of uncertain future events.[62] According to the court, the difference between a mistake of existing fact and a mistake with respect to a prediction of future events was not a distinction of law but of fact that is to be determined by the use of common sense.[63]

Continuing its analysis, the court held that the parties' mistake concerned a basic assumption underlying the contract and that the severe imbalance in the agreed exchange satisfied the basic requirements for relief under the doctrine of mistake.[64] Rejecting Essex's argument that the mistake was unilateral on ALCOA's part, the court stated flatly that both parties understood that the purpose of the agreement was to protect ALCOA from price fluctuations.[65] The court also minimized the necessity of find-

---

55. *Id.* at 58-59 (Table 1).

56. *Id.*

57. *Id.* at 57.

58. *Id.* at 64, 70.

59. *Id.* at 64 ("Here the practical necessities of the very long-term service contract demonstrated an agreed risk limiting device. Both parties understood this and adopted one").

60. *Id.* at 61-63. The court said that the WPI-IC's "capacity to work as the parties expected it to work was a matter of fact, existing at the time they made the contract." *Id.* at 64.

61. *Id.* at 63.

62. *Id.*

63. *Id.* at 61-62.

64. *Id.* at 64-65.

65. *Id.*

ing an "existing fact" in order to apply the mistake doctrine on the strength of the deletion of the word "existing" from the *Restatement (Second) of Contracts* between the tentative draft and the "approved final form."[66] In rejecting "existing fact" as an absolute prerequisite to relief, the court also relied on the case of *National Presto Industries, Inc. v. United States*.[67] In that case, the plaintiff had contracted to produce artillery shells through an experimental process on the erroneous assumption that grinding equipment would not be necessary for production. The *Presto* court did not find the existing fact requirement a barrier and granted relief to the contractor under the doctrine of mistake.[68]

Turning to the question of whether the agreement implicitly allocated the risk of the increased cost to ALCOA, the court noted that it had to allocate the risk "in some reasonable way."[69] The court rejected Essex' arguments that the escalation clause must be construed against ALCOA as its drafter and that the nonexistence of a profit floor provision compelled the conclusion that ALCOA had impliedly assumed the risks of increased costs. In doing so, the court relied on the great care taken and expense incurred by ALCOA to limit its risks.[70] The court also dismissed Essex' contention that this was a case of "conscious ignorance" since both parties knew that the escalation formula would not protect ALCOA in all circumstances. The court stated: "Both [parties] consciously undertook a closely calculated risk rather than a limitless one. Their mistake concerning its calculation is thus fundamentally unlike the limitless conscious undertaking of an unknown risk . . . ."[71] The court conceded, however, that Essex's arguments were correct within limits and stated that "within those limits they affect the relief ALCOA may receive."[72]

---

66. *Id.* at 71. *Compare* RESTATEMENT (SECOND) OF CONTRACTS, *supra* note 16, § 293 ("A mistake is a belief that is not in accord with existing facts"), *with* RESTATEMENT (SECOND) OF CONTRACTS, *supra* note 19, § 151 ("A mistake is a belief that is not in accord with the facts"). *See also* National Presto Indus. v. United States, 338 F.2d 99, 107 (Ct. Cl. 1964) (erroneous assumption that additional equipment would not be necessary characterized as mistake of fact rather than prediction). *But see* Cook v. Kelly, 352 Mass. 628, 227 N.E.2d 330, 333 (1967) (in action to rescind sale of newspaper business on ground of mistake, existing fact required and defined as something that can be ascertained at the time of the making of the contract).

67. 338 F.2d 99 (Ct. Cl. 1964).

68. *Id.* at 111.

69. 499 F. Supp. at 76.

70. *Id.* at 68.

71. *Id.* at 70.

72. *Id.* at 68; *see infra* notes 89-101 and accompanying text.

Noting that the likelihood of appeal of its decision was high, the court proceeded to analyze the case under the doctrines of impracticability and frustration as well as mistake.[73] The court found that the changes in circumstances had made performance of the contract impracticable[74] and distinguished the cases relied on by Essex as differing "from the present case in the absolute extent of the loss and in the proportion of loss involved."[75] The court held that although the predicted loss of $60,000,000 over the life of the contract was only an estimate, it showed a burden of performance beyond that contemplated by the parties, constituting impracticability.[76]

In analyzing Essex' contention that the variation in the WPI-IC index and ALCOA's costs were foreseeable and therefore implicitly allocated to ALCOA, the court criticized Judge King's holding in *Eastern Air Lines, Inc. v. Gulf Oil Corp.*,[77] that relief was only available under the impracticability doctrine, as codified by U.C.C. § 2-615, for "an unforeseeable failure of a pre-supposed condition."[78] The court argued that "foreseeability or even recognition of a risk does not necessarily prove its allocation."[79] The

---

73. *Id.* at 70.

74. *Id.*

75. *Id.* at 74. Essex relied on: Transatlantic Financing Corp. v. United States, 363 F.2d 312 (D.C. Cir. 1966) (extra expense of $44,000 required for performance of shipping contract of $306,000 not sufficient to show impracticability); Iowa Elec. Light & Power Co. v. Atlas Corp., 467 F. Supp. 129, 134-40 (N.D. Iowa 1978), *rev'd on other grounds*, 603 F.2d 1301 (8th Cir. 1979) (projected loss of approx. $2,500,000.00 on uranium supply contract attributable to production cost increases of approximately 60% not sufficient to show performance of contract impracticable); Publiker Indus. Inc. v. Union Carbide Corp., 17 U.C.C. Rep. Serv. 989 (E.D. Pa. 1975) (performance not impracticable even though seller's cost of gasoline rose from 21.1¢ per gallon to 37.2¢ per gallon within 18 months); Eastern Air Lines, Inc. v. Gulf Oil Corp., 415 F. Supp. 429 (S.D. Fla. 1975) (oil price increase from $5 per barrel in September, 1974, to $11 per barrel in January, 1975, did not make performance of requirements contract for jet fuel impracticable); Maple Farms, Inc. v. City School Dist., 76 Misc. 2d 1080, 352 N.Y.S.2d 784 (N.Y. Sup. Ct. 1974) (milk supply contract not impractical despite 23% cost increase).

76. 499 F. Supp. at 73.

77. 415 F. Supp. 429, 438 (S.D. Fla. 1975).

78. 499 F. Supp. at 75-76.

79. *Id.* at 76 (quoting Transatlantic Fin. Corp. v. United States, 363 F.2d 312, 318 (D.C. Cir. 1966)). It is not clear what degree if any of unforeseeability is required for relief under the U.C.C. Comment 4 to section 2-615 requires "some unforeseen contingency which alters the essential nature of the performance." U.C.C. § 2-615 comment 4 (1972). Comment 8 requires that the contingency causing the increased burden be sufficiently foreseeable "to be included among the business risks that are fairly to be regarded as part of the dickered terms" of the agreement before relief will be denied. *Id.* comment 8. Under comment 8, it might still be true that the parties may even contemplate and expressly provide for a type of risk that in fact occurs, but the magnitude or duration of the risk is so great that it

court, however, stopped short of dispensing with unforeseeability altogether, stating that the wide variation between the WPI-IC and ALCOA's costs "was unforeseeable in a commercial sense."[80]

The court applied the doctrine of frustration by holding that one of ALOCA's principal purposes[81] was to make a profit and that this purpose had been frustrated.[82] The court rejected the view that profitability was not an appropriate basis for application of the frustration doctrine but cited no supporting authority.[83] The court also argued that loss avoidance, another of ALOCA's principal purposes, had been totally frustrated.[84] The court relied on Professor Corbin's denunciation of the 1926 English case of *Anderson v. Equitable Life Assurance Society*[85] as "demonstrating that at times courts should treat loss avoidance as a principal purpose of a party."[86] The court further asserted that cases decided during the revolutionary and civil wars granting relief for problems created by serious inflation would be explained today in terms of frustration of purpose.[87] The court argued that neither the exact explanation of the decisions found in those cases nor the exact character of the relief granted was important because the cases predated the evolution of the doctrine of frustration.[88]

---

cannot fairly be said to have been part of the dickered terms of the contract. *Cf.* Glenn R. Sewell Sheet Metal, Inc. v. Loverde, 70 Cal. 2d 666, 451 P.2d 721, 728 n.13, 75 Cal. Rptr. 889, 896 n.13 (1969) (discussing effect of foreseeability on duty to perform under land lease).

80. 499 F. Supp. at 76.

81. The court's analysis focused on the frustration of a principle purpose of a party in entering into the contract rather than the purpose of the contract itself. *See infra* note 138 and accompanying text.

82. 499 F. Supp. at 76-77.

83. *Id.* The court questioned the *Restatement (Second) of Contracts* characterization of the principle purpose of a lease agreement for the operation of a gas station as the operation of the station rather than the profit from operation. *Id.* at 77 (discussing RESTATEMENT (SECOND) OF CONTRACTS, *supra* note 9, § 285 illustration 7). The court concluded that the *Restatement*'s preclusion of profits from being a principal purpose of a contract extended only to lease agreements because of their "particular circumstances." *Id.*

84. 499 F. Supp. at 76-78.

85. 134 L.T.R. 557 (1926). *See* 6 A. CORBIN, *supra* note 7, § 1560, at 486-88. In *Anderson*, the plaintiff was the beneficiary of a life insurance policy with benefits payable in German marks. As the policy premiums were paid, they amounted to 2,377 English pounds, but, because of German hyperinflation, the benefits under the policy came to less than one English penny. The insurance company thus asserted that it owed nothing on the policy. The *Anderson* court held for the insurance company. Of that case, Corbin states, "if the facts show that the promisee is left suffering a heavy economic loss, while the promisor reaps a corresponding great profit, the established rules of our legal system do not require such a decision." *Id.*

86. 499 F. Supp. at 77.

87. *Id.* at 78.

88. *Id.* The doctrine of frustration is generally believed to have originated in Krell v.

Considering the question of what remedy would be appropriate, the court acknowledged the traditional restriction of the reformation remedy to stenographic errors in the written contract but stated that this case fell "within the more general rules of equitable restitution" which emphasized the prevention of unjust enrichment.[89] The court cited cases dealing with real property in which the parties miscalculated the actual acreage involved and justifiable reliance had made rescission inequitable.[90] The court noted that, in such cases, the remedy invoked is often called "'reformation' in the loose sense of 'modification' "[91] and that they involved fully executed contracts, but concluded that equity could be achieved in a "long-term executory contract by a similar remedy."[92] The court indicated that rescission of the contract would grant ALCOA a windfall gain in the current aluminum market and at the same time deprive Essex of the assured long-term aluminum supply it had bargained for.[93] The court concluded, therefore, that equitable reformation was "essential to avoid injustice"[94] and modified the price escalation formula to allow ALCOA the lesser of one penny profit per pound of aluminum smelted (the minimum profit contemplated by the parties) or the contract ceiling price.[95] This modification was to provide Essex the benefit of its bargain and at the same time contain ALCOA's losses within the limits anticipated by the parties.[96]

The court asserted that the equitable reformation remedy

Henry, [1903] 2 K.B. 740, 747-55 (C.A.). J. CALAMARI & J. PERILLO, *supra* note 4, at 495.

89. 499 F. Supp. at 78.

90. *Id.* at 79. The court cited: Schwaderer v. Huron-Clinton Metropolitan Auth., 329 Mich. 358, 45 N.W.2d 279 (1951) (where contract called for brush to be cleared from tract of land containing 239 acres, court reformed contract to compensate plaintiff for extra work because actual acreage was 545 acres); McMahan v. Terkhorn, 67 Ind. App. 501, 116 N.E. 327 (1917) (where buyer purchased land for less than fair value on basis of erroneous land survey and then resold the land before discovery of the mistake, the court reformed the contract to require the purchaser to pay for the extra land conveyed); 3 G. PALMER, *supra* note 32, § 13.9 ("mistake in expression or integration"). The single case cited by the court as supporting equitable reformation that did not deal with mistake was Parev Products Co. v. Rokeach and Sons, Inc., 124 F.2d 147 (2d Cir. 1941) (awarding damages for breach of implied covenant not to compete). It might be inferred from the authorities cited that the court felt more comfortable equitably reforming the contract under the mistake doctrine than under the doctrines of frustration or impracticability. *See supra* notes 56-60, 73-76 and accompanying text.

91. 499 F. Supp. at 79.

92. *Id.*

93. *Id.*

94. *Id.*

95. *Id.* at 80.

96. *Id.*

would be appropriate under any of the three doctrines discussed if injustice could not be otherwise avoided.[97] In conclusion, the court likened its use of equitable reformation to "gap filling" under the Uniform Commercial Code when parties have omitted a necessary term from their contract.[98] The court suggested four factors courts should consider in deciding whether to equitably reform contracts: "(1) the parties' prevision of the problems which eventually upset the balance of the agreements and their allocation of the associated risks; (2) the parties' attempts at risk limitation; (3) the existence of severe out of pocket losses and (4) the customs and expectations of the particular business community."[99] The court expressed its conviction that "limitation of judicial relief to cases where the parties evidence a desire to limit their risks, where one party suffers severe out of pocket losses not adequately foreseen . . . seems adequate to prevent a general disruption of commercial life by inflation."[100] Judge Teitlebaum concluded "at stake in this suit is the future of a commercially important device—the long-term contract . . . . If the law refused a remedy when a prudently drafted contract goes badly awry, the risks attending such contracts will increase. Prudent business people would avoid using this sensible business tool."[101]

III. Analysis

A. *The Remedy of Equitable Reformation*

Long-term contracts play a large and essential role in modern society[102] because they allow businessmen to plan for the future and to specialize their productive efforts.[103] Judicial restriction of remedies available to the parties when those contracts become imbalanced would seem to promote harsh and unreasonable re-

97. *Id.* at 78-80. *See supra* note 90.

98. 499 F. Supp. at 91 (citing U.C.C. §§ 2-204, -305, -308, -310 (1972)).

99. *Id.* at 92.

100. *Id.* at 89.

101. *Id.*

102. *See* MacNeil, *supra* note 34. The United States Bureau of Labor Statistics reports that, in 1976, more than one-half of 600 responses to 20,000 inquiries of users of federal price data responded that they used price escalation clauses in contracts representing $100 billion in transactions. United States Bureau of Labor Statistics, Rept. No. 509, BLS Industries Price Program: A Survey of U.S. Users (1977).

103. *See* MacNeil, *supra* note 34, at 857 (advanced societies need specialization of effort and thus require advance planning).

sults.[104] *Eastern Air Lines, Inc. v. Gulf Oil Corp.*[105] vividly demonstrates the inequitable results of rigidly applying traditional contract law to long-term contracts after the contract has become imbalanced by unforeseen events. In that case, the court refused to mitigate Gulf's burden of performance under a requirements contract for approximately 100 million gallons of jet fuel despite dramatically increased oil costs. The price of Gulf jet fuel to Eastern was determined by an escalation clause based on domestic oil prices. The escalation clause failed to maintain the equivalence of the parties' exchange when OPEC oil prices rose rapidly and domestic price increases were stifled by government regulation.[106] In refusing to grant relief under the doctrine of impracticability, the *Eastern Airlines* court held that Gulf had failed to prove the amount of its out of pocket losses,[107] and that the rise in OPEC oil prices was foreseeable at the time the contract was signed—approximately one year before the embargo.[108] Although the court's findings are supported by the reported facts, the carefully structured escalation clause suggests that it is highly unlikely that Gulf ever expected to sell or Eastern ever expected to buy jet fuel at a price drastically below that prevailing in the market. Under the reasoning of *ALCOA*, the *Eastern Airlines* court could therefore have justified awarding relief to Gulf.

Assuming some relief is appropriate in circumstances like those in *ALCOA* and *Eastern Airlines*, the traditional remedy of rescission[109] is not always satisfactory. In the context of long-term contracts the remedy will sometimes simply shift windfall profits and unexpected losses incident to the unforeseen event from one party to the other.[110] For example, if a supply contract is rescinded because of large increases in costs, the seller will gain the

---

104. *See supra* notes 42-44.

105. 415 F. Supp. 429 (S.D. Fla. 1975).

106. *Id.* at 433-34.

107. *Id.* at 440, 442.

108. *Id.* at 432-33. The court reasoned that a party disadvantaged by a foreseeable contingency is not entitled to judicial relief as he could have protected himself through the contract. *Id.*

109. *See supra* note 28 and accompanying text.

110. The *ALCOA* court stated:

> To decree rescission in this case would be to grant ALCOA a windfall gain in the current aluminum market. It would at the same time deprive Essex of the assured long-term aluminum supply which it obtained under the contract and of the gains it legitimately may enforce within the scope of the risk ALCOA bears under the contract.

499 F. Supp. at 79.

opportunity to sell his products at the current market price and the buyer will be forced to buy at that price. This effectively shifts the total burden of the unforeseen event from the seller to the buyer. The traditional remedy of rescission allows the court only to decide which of two innocent parties must bear the total burden of an unexpected event. Rather than forcing the loss on either of the parties, courts have tended to leave it on the party on which it has fallen. This result seems harsh, however, when the parties have attempted to share unforeseen risks through an escalation clause. In these instances, the use of equitable reformation seems appropriate.[111]

Equitable reformation has been used by several courts to avoid harsh and unjust results. For example, in *National Presto Industries, Inc. v. United States*,[112] the United States Court of Claims divided unanticipated extra costs incident to a munitions supply contract between the parties, stating simply, "though the particular result here may be unprecedented that is, of course, the way of the common law."[113] Similarly, in *Parev Products Co. v. Rokeach & Sons*,[114] the court reformed the contract to require the licensee of a secret formula to pay royalties to the licensor on competing products manufactured and marketed by the licensee, stating that the "intention of parties is a good formula by which to square doctrine with results."[115]

The court in *ALCOA* likened the remedy of equitable reformation to equitable restitution, which is used to remedy unjust enrichment that has already occurred.[116] Equitable reformation is a similar concept, but it is used to prevent prospective unjust enrichment. The *Restatement (Second) of Contracts* seems to approve of equitable reformation when a more traditional remedy would not avoid injustice.[117] The *Restatement* provides:

> [I]n any case governed by the rules stated in [these] chapter[s] [dealing with impracticability of performance and mistake], if those rules together with the rules stated in Chapter 16 [Remedies] will not avoid injustice, the court may grant relief on terms as justice

111. *See generally* 3 G. PALMER, *supra* note 32, § 13.9.
112. 338 F.2d 99 (Ct. Cl. 1964).
113. *Id.* at 111.
114. 124 F.2d 147 (2d Cir. 1941).
115. *Id.* at 149.
116. 499 F. Supp. at 78-79.
117. RESTATEMENT (SECOND) OF CONTRACTS, *supra* note 19, §§ 125(2), 272(2).

requires including protection of the parties' reliance interests.[118]

Although the commentary to the *Restatement* does not address the appropriateness of equitable reformation as used by the court in *ALCOA*, it also does not place any conceptual restriction on the courts' fashioning of an equitable remedy.[119]

The *ALCOA* court's analogy of equitable reformation to "gap filling" under the Uniform Commercial Code[120] is interesting, yet inapposite. There seems to be an obvious difference in magnitude, if not in kind, between filling in missing terms of a contract and rewriting a complex escalation clause that has failed to work as anticipated. The two remedies are similar, however, in that they both relate to what the parties would have done had they known all the facts.[121]

The competence of courts to adjust the parties' agreement in a truly equitable manner seems questionable. The adjustment of an allegedly disfunctional escalation clause in a modern long-term contract may require courts to make difficult and complex economic decisions.[122] A court may also find itself responsible for the supervision and perhaps the readjustment of the contract throughout the contract term.[123] The potential burden on the judiciary is undeniably great;[124] however, in circumstances where the economic issues are not too complex and where the court perceives that the intention of the parties was to share risks rather than merely allocate them, equitable reformation is a remedy with much appeal.[125]

---

118. *Id.*

119. *Id.*

120. *See supra* note 98 and accompanying text.

121. With respect to the process of supplying an omitted term, RESTATEMENT (SECOND) OF CONTRACTS, *supra* note 19, § 240 comment (d), states: "Both the meaning of the words used and the probability that a particular term would have been used if the question had been raised may be factors in determining what term is reasonable in the circumstances." As the RESTATEMENT suggests, there are also other factors, such as fairness, that courts may consider in supplying a missing term. *See supra* notes 92-99 and accompanying text.

122. At a minimum, such judicial determinations may require great expenditures of time and money for lengthy trials and expensive advocacy.

123. Brief of Appellant at 66, *ALCOA*.

124. Equitable reformation, however, may not so much invite litigation as it does compromise. Although the *ALCOA* decision was appealed, the parties renegotiated their contract and settled before judgment. Telephone interview with Richard W. Gladstone II, counsel for ALCOA (Nov. 14, 1982).

125. *See* Speidel, *supra* note 34, at 279 (arguing that, when dealing with burdens in long-term contracts created by unforeseen circumstances, "the appropriate response for a court is to press aggressively for an agreed modification or to impose an adjustment as a condition to equitable relief"). Section 2-719(2) of the Uniform Commercial Code states that where the court grants specific performance, "the decree . . . may include such terms and

Conceptual barriers[126] to equitable reformation put parties seeking enforcement of long-term contracts in an undeservedly advantageous bargaining position. Potential litigants undoubtedly understand that if courts feel that the only alternatives are enforcement or recission, they are apt to leave the burden of an unforeseen event where it has fallen.[127] The advantageous negotiating position of the party seeking enforcement seems particularly undeserved when the parties understood from the outset that flexibility was an integral part of their agreement.[128] This completely fortuitous advantage may discourage negotiation altogether.[129] Judicial acceptance of equitable reformation may decrease the strength of the bargaining position of the party seeking enforcement[130] and thereby encourage the resolution of problems caused by unforeseen events through compromise rather than litigation.

B. The Court's Use of the Escape Doctrines

The *ALCOA* court found the doctrines of impracticability, frustration and mistake to be alternative grounds for finding relief appropriate and invoked the remedy of equitable reformation.[131] The court followed the traditional approach in analyzing the case under the impracticability doctrine, focusing on the hardship of ALCOA's performance and the unforeseeability of the event causing that hardship.[132] ALCOA's great burden of performance was

---

conditions as to payment of the price, damages, or other relief as the court may deem just." U.C.C. § 2-719(2) (1972). Spiedel argues that this clause recognizes the need to leave courts free to fashion equitable remedies on a case by case basis where traditional remedies will not avoid injustice. Spiedel, *supra* note 34, at 377. *But see* Iowa Elec. Light & Power Co. v. Atlas Corp., 467 F. Supp. 129, 138 (N.D. Iowa 1978) (U.C.C. § 2-719 does not allow courts to equitably adjust terms of contract), *rev'd on other grounds*, 603 F.2d 1301 (8th Cir. 1979).

126. The *ALCOA* court refers to "half-remembered truths and remembered half-truths" from the first year course in contract law and "the hoary maxim that the courts will not make a contract for the parties." 499 F. Supp. at 91.

127. *See* Jennings, *Commercial Impracticability—Does it Really Exist?*, 2 WHITTIER L. REV. 241, 256 (1980).

128. *See* MacNeil, *supra* note 34, at 865 (referring to techniques used by contract planners to create flexibility).

129. *See* Hurst, *Drafting Contracts in an Inflationary Era*, 28 U. FLA. L. REV. 879, 882 n.24 (1976) (parties to contracts will not negotiate in the face of changed circumstances if they feel that litigation may provide them with what will amount to a windfall gain).

130. *See* Speidel, *supra* note 34, at 271 ("Excuse under [U.C.C.] section 2-615 is a game with a predictable outcome—the buyer wins"). Speidel's view is that the party seeking enforcement of a long-term supply contract is in a very strong bargaining position indeed. *See id.*; *see also* Macauly, *supra* note 34, *passim* (arguing the economic benefits of avoiding litigation through negotiation).

131. 499 F. Supp. at 78.

132. *Id.* at 72-73, 75-78; *see supra* notes 4-21 and accompanying text.

evident in the $60,000,000 projected loss ALCOA would have suffered had it been required to supply Essex with aluminum over the life of the unadjusted contract.[133] Not only was the magnitude of the loss great, but ALCOA's nonlabor production costs rose over 200% faster than the WPI-IC index chosen by the parties to monitor them.[134] The court's conclusion[135] that the rapid rise in energy costs beginning in 1973 was not foreseeable by the parties in 1968, when the contract was signed seems quite reasonable.

The court's analysis of the case under the doctrine of frustration is more troublesome. Profitability has not been the kind of contractual purpose that courts have recognized under the doctrine of frustration,[136] perhaps because the mere nonprofitability of a contract has been regarded as a risk clearly assumed by the parties.[137] The requirement adopted by some courts that the frustrated purpose be a purpose common to both parties also would seem to exclude nonprofitability as an appropriate basis for the use of the frustration doctrine.[138] The court's alternative rationale that one of the parties' purposes was to avoid severe loss and that this purpose had been frustrated seems somewhat more acceptable in light of the contract's escalation clause. That rationale might be justified under reasoning similar to the analysis of the court's use of the doctrine of mistake.[139]

The court's use of the doctrine of mistake departs from traditional formulations.[140] Finding that the parties' belief that the

---

133. 499 F. Supp. at 73.

134. *Id.* at 59. ALCOA's actual nonlabor production costs increased approximately 200% faster between 1974 and 1978 than the WPI-IC index.

135. *Id.* at 76.

136. *See supra* note 11 and accompanying text.

137. *See supra* note 11.

138. *See* Edward v. Leopoldi, 20 N.J. Super. 43, 89 A.2d 264, 271 (N.J. Super. Ct. App. Div. 1952) (both parties' purpose for entering into agreement to send union funds to national headquarters on disbandment of local chapter not frustrated by severance of alliance between the union and the CIO); North Am. Capital Corp. v. McCants, 510 S.W.2d 901, 905 (Tenn. 1974) (refusal of Federal Home Loan Bank Board to approve site for savings and loan association did not destroy both parties' purpose for entering into lease agreement); *see also* 18 S. WILLISTON & W. JAEGER, WILLISTON ON CONTRACTS § 1954, at 133 (1979) ("fortuitous circumstances [must] nearly or quite completely destroy the purpose both parties to the bargain had in mind"). Of course, at some point the unprofitability of the contract may make performance so burdensome as to make the doctrine of impracticability applicable. *See supra* notes 132-135 and accompanying text.

139. The court's reliance on Corbin as authority for its "purpose to avoid loss" rationale, 499 F. Supp. at 77-78, is questionable, however. Corbin's discussion is based on problems created by inflation so severe that the currency is rendered virtually worthless. *See supra* notes 84-86 and accompanying text.

140. *See supra* notes 16-21 and accompanying text.

WPI-IC index would accurately reflect nonlabor production costs was a mistake of existing fact[141] amounts to a finding that a mistaken prediction of future events is a mistake of existing fact. That view is contrary to the traditional rule[142] and virtually eliminates the distinction between mistake and the doctrines of frustration and impracticability.[143] There seems to be nothing inherently undesirable in that result,[144] however, because all three doctrines are closely related and sometimes have been applied to similar situations.[145]

Although analysis under the concept of impracticability would have been more defensible from a traditional contract law perspective, the *ALCOA* court's reliance on the doctrine of mistake clearly helped to justify the use of reformation to allow ALCOA a minimum profit. Where courts have used equitable reformation to make a contract conform to what the parties would have done had they known all the facts, they have generally justified its use under the rubric of mistake.[146] Because the doctrine of mistake is

---

141. *See supra* notes 58-63 and accompanying text.

142. *See supra* notes 6, 26 and accompanying text.

143. *See supra* notes 4-6 and accompanying text.

144. Professor Farnsworth, the reporter for the RESTATEMENT (SECOND) OF CONTRACTS, discussing comments received at the annual meeting of the American Law Institute stated:

> It would be fair to say that there were probably as many reasons for dropping [the existing fact requirement in § 293] given to me as there were people who had advanced the opinion. [A]t the end at least a dozen people had said they didn't like "existing" and nobody had defended it.

499 F. Supp. at 62. *See supra* note 66.

145. The following three cases provide a good example of how the doctrines of mistake, impracticability and frustration overlap. In each case, the parties were mistaken as to the amount of mineral available. The land owners sued the mining company for minimum royalties due on the contract and the court ruled in favor of the company on one or more of the three doctrines. In Carr v. Whitebreast Fuel Co., 88 Iowa 136, 55 N.W. 205 (1893), the defendant coal company was released from its obligation to pay royalities on the ground that the contract was based on a mutual mistake of fact. 55 N.W. at 209-11. In Virginia Iron, Coal & Coke Co. v. Graham, 124 Va. 692, 93 S.E. 649 (1918), the defendant iron company was released on the grounds of impossibility of performance because the subject matter of the contract had been destroyed when the iron ore ran out. 93 S.W. at 662-63. The court also granted relief on the alternate ground of mutual mistake of fact. *Id.* at 665. In Fritzler v. Robinson, 70 Iowa 500, 31 N.W. 61 (1886), the court released the defendant coal company on the grounds that the absence of coal on the premises constituted a failure of the consideration "arising out of mutual mistake." 31 N.W. at 63. The latter case could just as easily have been decided on a theory of frustration in light of the court's analysis of a failure of consideration. *See also* United States v. General Douglas MacArthur Senior Village, Inc., 508 F.2d 377, 383 (2d Cir. 1974) (majority refuses to grant relief under doctrines of impossibility and frustration, dissent argues mutual mistake of fact).

146. *See* National Presto Indus. v. United States, 338 F.2d 99, 110 (Ct. Cl. 1964) (allocating unexpected expenses under mistake doctrine where the parties "if made aware of the true facts, 'would have agreed at the outset to change now sought.' "); Schwaderer v. Huron-

grounded on notions of unjust enrichment,[147] the court was able to rely on the fact that ALCOA's projected out-of-pocket losses were mirrored by windfall gains to Essex in deciding whether relief was appropriate.[148] Under the doctrines of frustration or impracticability, the courts have not looked to the relative gains and losses of the parties in deciding whether to grant relief.[149] Thus, the *ALCOA* court's initial emphasis on unjust enrichment rather than on burden of performance may make it somewhat easier for a party to obtain relief under the doctrine of mistake than under frustration or impracticability when the contract has become imbalanced. It is not clear that this would always hold true, however, because the imbalance must be "severe" before relief will be granted[150] and "severe" imbalance may be tantamount to an impracticable performance. The *ALCOA* court did not address that question.

The court's determination that the contract did not allocate to ALCOA the risk of variation between the WPI-IC index and ALCOA's actual costs seems just. To avoid that kind of subjective determination, courts traditionally have held that an adjustment clause expressly defines the limits of the parties' efforts to maintain an equivalent exchange.[151] In *ALCOA*, however, risks accompanying the lengthy contract and the complexity of the escalation formula suggest that the parties' objective was to share rather than allocate risks.

Critics of the *ALCOA* court's subjective analysis might argue that if the parties intended to share rather than allocate the risks of unforeseen contingencies, they could have included an express force majeure,[152] arbitration or renegotiation clause in their agree-

---

Clinton Metropolitan Auth., 329 Mich. 258, 45 N.W.2d 279 (1951) (reforming low contract bid to allow reasonable compensation for extra work done clearing brush where land contained more acres than supposed by parties).

147. G. PALMER, *supra* note 18, at 36-38, 53.

148. *See* 499 F. Supp. at 64.

149. *See* Note, *U.C.C. § 2-615: Excusing the Impracticable*, 60 B.U.L. REV. 575, 581 (1980) (indicating that the only relevant factor in determining whether a party has met the hardship requirement of the impracticability doctrine is "the extent by which the actual cost of performance exceeds the anticipated cost of performance").

150. 499 F. Supp. at 64.

151. *See, e.g.*, Eastern Air Lines v. Gulf Oil Corp., 415 F. Supp. 429, 439 (S.D. Fla. 1975) ("The short and dipositive answer to Gulf's . . . argument . . . that the price escalation indicator . . . no longer reflects the intent of the parties by reason of the [government regulation], is that the language of the contract is clear and unambiguous").

152. *Force majeure* means "superior or irresistable force." BLACK'S LAW DICTIONARY 774 (4th ed. 1968). In contracts, the term refers to a clause designed to protect the promisor from the occurrence of unforeseen contingencies beyond his control. Hurst, *supra* note 139, at 900.

ment.[153] However, the parties may not have included such clauses and still have intended to share the risks of future events. Those kinds of clauses are not necessarily easier to negotiate than escalation clauses,[154] and the contracting parties may see little advantage in such clauses over an escalation clause that promises to maintain the equivalency of their exchange. Furthermore, some research indicates that businessmen tend to view "too detailed planning or intervention by third parties" as inconsistent with a long-term commitment to a working relationship[155] and may, therefore, prefer an escalation clause to detailed contingency planning.

The *ALCOA* court's analysis of mistake might be viewed as a threat to the stability of contractual relationships, giving unhappy promisors one more escape doctrine to use against the promisee. A party seeking relief from a long-term contract under *ALCOA*'s standard, however, still has the difficult burden of proving that: the parties contemplated a flexible agreement to maintain the basic equality of their exchange,[156] the magnitude of his loss falls outside of the risks allocated to him by a reasonable interpretation of the agreement[157] and his loss is mirrored by an unbargained for gain to the other party.[158] In short, he must show that the reasonable expectations of the party seeking enforcement will not be disappointed by equitable adjustment. That burden seems adequately protective of contractual relations.

CONCLUSION

ALCOA departed from a long line of cases that have declined to award relief from contracts that have become burdensome because of increasing costs.[159] In holding that the failure of the contract's escalation clause to accurately predict dramatically increased energy costs was a mutual mistake of fact, the court blurred the distinction between mistake and the doctrines of frustration of purpose and impracticability. That blurring, however, is justified in light of the confusion that has often attended the appli-

153. For a discussion of contract adjustment clauses, see MacNeil, *supra* note 34, at 865-73.

154. *See* Hurst, *supra* note 129, at 900-03 (discussing difficulties of negotiating adjustment clauses).

155. Macauley, *supra* note 34, at 508.

156. 499 F. Supp. at 65 (discussing escalation clause).

157. *Id.* at 69-70; *see supra* notes 58-59 and accompanying text.

158. 499 F. Supp. at 64; *see supra* note 64 and accompanying text.

159. 499 F. Supp. at 63-64.

cation of all three doctrines and is less important than the remedy prescribed by the court.

The *ALCOA* court rejected the traditional alternatives of enforcement of the contract or rescission and instead rewrote the price escalation clause to permit ALCOA a profit. The court thus went further than other courts that have equitably adjusted the prices of fully performed contracts. Although this departure from tradition may be criticized as making a contract for the parties, equitable reformation provides a middle ground of compromise that avoids placing the burden entirely on one of the parties. The "new spirit" of *ALCOA* provides a needed addition to the alternatives available to a court faced with the injustice that results under the traditional approach.

MICHAEL N. ZUNDEL

INDEX TO VOLUME 1982

ARTICLES AND ESSAYS—AUTHORS

ARTICLES AND ESSAYS—TITLES

NOTES & COMMENTS

CASES NOTED

LIST OF SUBJECTS INDEXED

SUBJECT INDEX

Reasons to own USCA

■ Complete—United States Code Annotated offers you every word, every section and every title of the United States code. It also includes Executive Orders, Proclamations and Reorganization Plans affecting the laws contained therein.

■ Annotated—United States Code Annotated includes annotations to meaningful court decisions construing the federal statutory law.

■ Indexed—individual title indexes and an eight volume Master Index put the pertinent code sections at your fingertips in just seconds.

■ Cross-Referenced—United States Code Annotated points you from one section and title to other sections and titles containing relevant material.

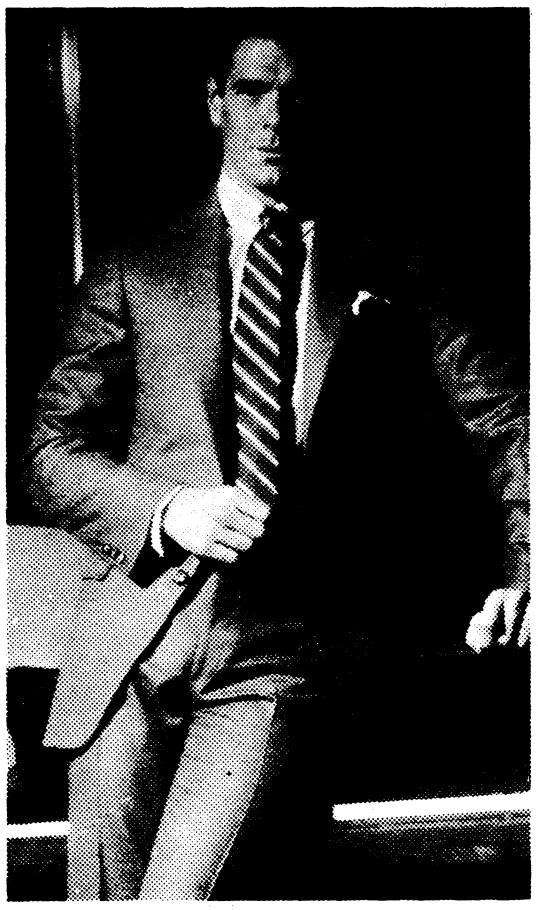
■ Library-Referenced—United States Code Annotated links you directly with the Key Number Digests to lead you to the full study of American case law and also to Corpus Juris Secundum to provide you with a solid full text development of the substantive and procedural aspects of your question.

■ References to the Code of Federal Regulations and West's Federal Forms, the Federal Rules of Criminal Procedure, Civil Procedure, Evidence, and Appellate Procedure, the rules of the U.S. Supreme Court, Court of Claims, Court of Appeals, Court of International Trade, and other courts, Historical Notes, rules of practice before federal agencies—these are just a glimpse at the many additional reasons you have to own and use United States Code Annotated.

■ And right now, special discounts, low down payments, extended upkeep service at no extra charge, and level payment plans give you ample reason to buy United States Code Annotated. Call or write today for complete details.

WEST PUBLISHING COMPANY

SOUTHWICK PRESENTS VARIATIONS ON THE THEME OF FLANNEL.



Southwick's new collection of wool flannel suits: Soft, relaxed and quietly assertive. The colors include not only greys, but blues, mauves, browns . . . in fact a nearly infinite variety of subtle shades. And all impeccably tailored in the Southwick natural-shoulder tradition.

We're proud to be one of a relatively small number of stores to offer these classics.

Hepworth's

Fine Traditional Clothiers

S.L.C., Utah