HARVARD JOURNAL
on
LEGISLATION

VOLUME 27, NUMBER 1 WINTER 1990

ARTICLES

STATUTE

NOTES

HARVARD JOURNAL ON LEGISLATION
Harvard Law School
Cambridge, MA 02138

Phone: (617) 495-4400

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

Publication Policy

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

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

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

Subscriptions

The JOURNAL's subscription rate for the 1990 volume (Volume 27) is $20.00 domestic rate; $23.00 foreign surface rate; $35.00 foreign air mail. All correspondence concerning subscriptions should be addressed to the Subscriptions Manager, Publications Center, Harvard Law School, Cambridge, MA 02138.

Unless otherwise requested by the subscriber, the JOURNAL is mailed at bulk rate. For some overseas subscribers, delivery via air mail may be more reliable than standard surface rate service. Claims for missing issues of Volume 27 will be honored until January 1991.

Individual Issues

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

Copies of issues from all past volumes of the JOURNAL are available from Fred B. Rothman & Co., 10368 West Centennial Road, Littleton, CO 80123, (303) 979-5657. Single isues are $8.50 per copy. Contact Fred B. Rothman & Co. for information on prices for bound volumes.

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

Indexes

Starting with Volume 23, the JOURNAL discontinued publication of volume-by-volume and five-year cumulative indexes. The JOURNAL is indexed in the *Index to Legal Periodicals*.

ARTICLE

NEW JUDGMENT LIENS ON PERSONAL PROPERTY: DOES "EFFICIENT" MEAN "BETTER"?

WILLIAM J. WOODWARD, JR.*

Recently, state legislatures have undertaken procedural reform of the judgment collection system. In this Article, Professor Woodward reviews these changes, noting that despite an apparent improvement in the efficiency of judgment collection, the statutes have many potential side effects. In light of the impact the statutes may have on state taxpayers, small creditors, and the federal bankruptcy system, Woodward cautions state legislatures to move slowly in adopting the reforms.

I. INTRODUCTION

Not many would argue that yesterday's cumbersome rules of procedure are intrinsically superior to the lean, efficient procedural rules of today. Just as the flashy styles of nineteenth-century Victorian architecture have made room for the unembellished twentieth-century styles, the nineteenth century's intricate, tangled procedural law has been displaced by the twentieth century's clean, unadorned, simplified approach. Indeed, students of the law have come to accept as gospel the idea that procedural reform means less complexity—that the most efficient procedure in settling disputed rights is the best.[1] In procedure as in architecture, we have come to think that clean and trim is simply better.

One edifice of legal procedure has largely escaped the twentieth-century wrecking ball: the procedural law that greets the plaintiff who, upon recovering a civil money judgment, must attempt to collect it from a defendant who will not or cannot

* Professor of Law, Temple University. B.A., University of Pennsylvania, 1968; J.D., Rutgers University, 1975. Special thanks to Amy Boss, Lissa Broome, David Papke, Elizabeth Warren, Bill Whitford, Harold Weinberg and Richard Woodward for their helpful comments on earlier drafts and to Donna Byrne for her research and editorial assistance.

[1] Elihu Root, Secretary of State under Theodore Roosevelt, made the point as follows: "Everybody knows that the vast network of highly technical rules of evidence and procedure which prevails in this country serves to tangle justice in the name of form. It is a disgrace to our profession." 15 J. AM. JUD. SOC. 119 (1931). *See also* Miller, *The Proposed Federal Procedure Rules*, 11 TUL. L. REV. 425 (1937) (finding the objective of simplifying civil procedure to be commendable).

pay. The procedure is a legal eyesore. It is crowded with writs, sheriffs, obscure actions,[2] and traps for the uninitiated; the procedure differs in every state and its statutory foundations are often scattered throughout a state's code. Surely many a victor has emerged from exhausting litigation only to learn from her lawyer that collecting the judgment will cost more than the judgment is worth.[3] If ever a system demanded less complexity and more simplicity and efficiency, it is the judgment collection system.

Three states have responded with significant procedural reforms that promise to make the collection process more efficient by allowing judgment creditors to gain nonpossessory liens on their debtors' personal property. As a long-overdue innovation, the legislation deserves description, scrutiny, and analysis in its own right. But the increased efficiency these changes bring to the collection process raises a host of issues that might not be apparent to the observer steeped in the "efficient is better" fashion of twentieth-century procedural reform.

Using these procedural reforms as a specific focus, this Article examines the implications of increased efficiency in the judicial collection process within the debtor-creditor system. The analysis proceeds at two levels of generality. The first level examines the statutes and their impact on the execution process. The statutes are viewed narrowly, ignoring the effects they might have in the broader debtor-creditor system. Part II begins with the background needed to understand the new provisions and Part III shows how they operate. The analysis in Parts II and III demonstrates how the new procedures will make judgment collection a less expensive, easier process for creditors and therefore shows that the statutes can be called more "efficient."[4]

---

[2] "Amercement," for example, is an action that asserts that the sheriff failed within the execution process to discharge his duties. *See generally* Wyatt, *Amercement of Sheriffs*, 10 WAKE FOREST L. REV. 237 (1974); Meyers, *In League with the League*, 65 COM. L. J. 238 (1960); 9 DEBTOR-CREDITOR LAW ¶ 37A.12[B] (1989).

[3] *See generally* Leff, *Injury, Ignorance, and Spite—The Dynamics of Coercive Collection*, 80 YALE L.J. 1 (1970).

[4] In Parts II and III, "efficient" means that it takes less time, effort, or money for a party to achieve the same result with the new procedures than with the old. In this narrow sense, the provisions apparently save economic resources. As will be suggested in Part IV, however, widespread use of these less expensive procedures might yield reactions in other parts of the debtor-creditor system that cost more than the new procedures save. If that is the case, the new procedures will not save resources but will consume them.

Yet if enthusiasm for the new legislation is generated by Parts II and III, Part IV will dampen it. In Part IV, the Article moves to the second level of analysis and takes a critical look at the statutes in the much broader context of the debtor-creditor system. Considered in the context in which the statutes will actually operate, these new provisions, which make collecting judgments easier and less complicated, become troubling.

Among the problems with the legislation are the following: First, the new statutes may make bankruptcy a less viable option for those whose interests it now serves. Second, they may provoke increased use of the legal system as creditors without judgments scramble to secure judgments that give them access to the new procedures. Third, and related to the second point, the new statutes may give a priority advantage to those relatively few large creditors who are expert at getting judgments. Fourth, they may shift some of the costs of collection from debtors and creditors to taxpayers both in and outside of bankruptcy. Finally, the statutes seem likely to yield a redistribution of leverage away from the debtor class and toward the creditor class, a political issue that needs to be addressed by legislators considering the new provisions. Ultimately, the analysis in Part IV suggests that when they operate in context, the statutes may turn out not to be truly "efficient" at all; rather, they may yield greater overall system losses than the archaic system currently in place.

The two-level analysis pursued here thus serves additional purposes. It demonstrates that one's normative judgments—even when "efficiency" is the standard for judging—can depend on the level of generality within which one looks at legal change.[5] The analysis further indicates that in the extraordinarily interconnected debtor-creditor system, a very broad context for viewing legal change is more likely to produce valid conclusions. Finally, the analysis may even prompt some to wonder whether our attraction to "efficiency" in legal procedure, like our embrace of clean lines in architecture, may be more the product of fashion and popular culture than we might otherwise think.

---

[5] This point was made in the context of tort law in Balkin, *Too Good to be True: The Positive Economic Theory of Law*, 87 COLUM. L. REV. 1447, 1477–78 (1987).

II. REFERENCE

A. *The Immediate Context of the New Statutes*

California, Connecticut, and Maine have each recently developed statutes that enable a judgment creditor to get a nonpossessory lien on a debtor's personal property by filing a simple document[6] in an appropriate state office.[7] In other states a judgment holder must usually proceed with expensive execution proceedings in order to get a lien on the debtor's personalty. If the judgment holder does not have a lien, the debtor can jeopardize the judgment holder's interests by paying other creditors, frittering away assets, or entering bankruptcy with the law's blessing.[8]

The only way for a judgment creditor to get a lien on personal property as distinguished from real estate is to commence execution proceedings and levy on it.[9] This costly procedure generally requires that the creditor locate leviable property, file numerous documents, motivate and direct the sheriff, and hope that the execution does not precipitate bankruptcy.

If the debtor owns unencumbered real estate, the situation is dramatically different. In most places, a simple filing by the judgment creditor in an appropriate office fixes a nonpossessory

---

[6] Maine requires a judgment creditor to file a court-issued execution in order to obtain a judgment lien. While it requires the issuance and filing of an execution, obtaining a lien by this method is considerably less complex than taking the writ to the sheriff, having the sheriff execute it, and hoping that the sheriff can seize something of value.

[7] *See infra* notes 26–32 and accompanying text.

[8] The law of fraudulent conveyances, of course, places some restraints on the defendant's freedom to dispose of assets while a judgment is in force. *See generally* UNIF. FRAUD. CONV. ACT OF 1918; UNIF. FRAUD. TRANSFER ACT OF 1984. The classic treatise in the area is G. GLENN, FRAUDULENT CONVEYANCES AND PREFERENCES (rev. ed. 1940).

[9] *But see* ALA. CODE § 6-9-211 (1975); GA. CODE ANN. § 9-12-80 (1982); MISS. CODE ANN. § 11-7-191 (1972). In these three states, a judgment creditor has long obtained priority in all of a debtor's personal property without actually executing on it. Illustrative is Georgia's statute which provides:

> All judgments obtained in the superior courts, justice of the peace courts, or other courts of this state shall be of equal dignity and shall bind all the property of the defendant in judgment, both real and personal, from the date of such judgments except as otherwise provided in this Code.

GA. CODE ANN. § 9-12-80 (1982). A later provision, however, makes the lien so obtained invalid as to "third parties acting in good faith and without notice who have acquired a transfer or lien binding the property of the defendant in judgment." GA. CODE ANN. § 9-12-81 (1982).

For several reasons, other states shunned the approach of these states in favor of a process requiring actual execution for creation of a lien. *See infra* notes 19–20 and accompanying text.

lien to the real estate and thereby preserves the creditor's position inexpensively and without wresting possession from the debtor.[10] This alternative may be all that is needed to get paid.[11]

Consider, then, the dilemma facing the traditional unpaid judgment creditor whose debtor has no unencumbered real estate. If the creditor does nothing, she has no priority at all and risks losing the value of the judgment if the debtor's situation deteriorates further. If, on the other hand, she levies on, for example, the debtor's inventory, the debtor may be prompted to respond with a bankruptcy petition to halt and avoid the execution and to preserve the business. In short, the creditor can do nothing and risk deterioration of the hard-won value of the judgment, or she can spend the funds to execute on it and risk throwing more good money after bad. High procedural cost, risk of further losses, and uncertainty of result are the hallmarks of the execution system generally and, more specifically, the process for obtaining priority in personal property. As models of inefficient procedure, execution systems may have no equals.

B. *The Larger Context*

Execution statutes and judgment lien statutes are only small pieces of the much larger debtor-creditor system. The combination of discrete elements of that system working with and against one another is what establishes the relationships between debtors and creditors as groups. The most important thing to recognize here is that because the disparate parts of the system have strong ties to one another, one cannot assess change in one corner of the system without considering the consequences of the change throughout the system. Exemption statutes,[12] the federal Bankruptcy Code[13] and its state analogues,

[10] Kentucky, Massachusetts, Michigan, New Hampshire, and Rhode Island have no such provisions. *See* S. RIESENFELD, CREDITORS' REMEDIES AND DEBTORS' PROTECTION 89 (4th ed. 1987).

[11] When the lien appears with the real estate records, potential purchasers and lenders will discover the judgment lien because record searches typically accompany real estate transfers. *See, e.g.*, MO. REV. STAT. § 511.500 (1986); N.Y. CIV. PRAC. L. & R. § 5203(a) (McKinney 1978). When they discover it, potential purchasers will recognize that the lien will be superior to their interests and will modify their own behavior when dealing with the debtor. Indeed, to avoid their own involvement with the judgment creditor, prospective lenders and buyers may require the debtor to satisfy the lien as a condition to their secured loan or purchase of the property.

[12] Exemption provisions are discussed in sections IV, B and IV, C, 1 of this Article.

[13] 11 U.S.C. §§ 101–329 (1982 & Supp. V 1987).

lending regulations,[14] priority rules in secured financing[15] and other state laws creating liens, and extralegal methods used by debtors to avoid repaying debt and by creditors to collect debt are all part of the larger picture.

Considering legal change in this large, extraordinarily complex context makes analysis difficult. Each piece of the system is connected to the others in various ways, and movement in one part of the system yields some reaction in other parts.

Lawyers for creditors and debtors understand the interrelationship of these parts even if others do not. These scores of rules, practices, freedoms, and inefficiencies come together every day in an informal collection process, within which creditors and debtors settle with one another without direct recourse to the legal system.[16] Legal change in any corner of the larger system undoubtedly affects the everyday negotiations between creditors and their debtors. Most debts are not collected coercively; debtors and creditors both know that the formal legal system is often far too expensive to be of much use in debt collection.[17] Thus, in addition to assessing the impact of change on the operation of other formal rules within the larger system, it is important to address the impact that a legal innovation might have on the informal process of negotiation and settlement.

Improving the efficiency of the collection process—making the formal legal system less expensive to use—seems likely to increase creditors' recourse to lawyers and courts in collecting debts. This raises a multitude of questions. For example, what impact might increased use of formal processes for debt collection have on our already-burdened legal system? Is it desirable to reduce the strong incentives debtors and creditors now have to arrive at consensual settlements? If creditors will more read-

---

[14] Usury laws are perhaps the oldest examples of lending regulations. More recently, attention has focused on wage assignments, which came under scrutiny in the early 1930's and eventually were prohibited. 16 C.F.R. § 444.2(a)(3) (1989) (unfair credit practice to accept assignment of future wages unless revocable at will by debtor or preauthorized payroll deduction plan). The Federal Trade Commission rule prohibiting a debtor from creating a nonpossessory, non-purchase money security interest in substantially all her personal property is a lending regulation of more recent origin. 16 C.F.R. § 444.2(a)(4) (1989).

[15] The primary source of these priority rules today is Article 9 of the Uniform Commercial Code (U.C.C.).

[16] The legal system, of course, operates here most importantly as the context or backdrop for negotiation. *See* Whitford, *A Critique of the Consumer Credit Collection System*, 1979 WIS. L. REV. 1047, 1048–49, 1057–58 (1979).

[17] *Id.* at 1054.

ily compete in the courts rather than informally for the debtor's assets, are some creditors naturally better suited to the competition than others? After considering the specifics of the new statutes and their effects on discrete legal regimes elsewhere within the debtor-creditor system, this Article will try to assess the impact these statutes might have on this important informal process.

III. THE STATUTES AND THEIR IMPLICATIONS[18]

A. *Introduction*

California, Connecticut, and Maine have each made it inexpensive and easy for a judgment creditor to get a judgment lien on personalty. In each state, a judgment creditor can, by filing an appropriate document in the correct office, stake a claim to much of a debtor's personal property. The lien created by this procedure will secure the judgment creditor's priority in the personalty against many later claimants. Obtaining a judgment lien in these states is possible without using the sheriff, without removing the property from the debtor's control, and without much of the risk and cost one must sustain in other states to get a similar priority advantage. Given the long history of judgment liens on real property, one might preliminarily consider why such innovation took so long to arrive.

The absence of a dependable, centralized, routinely utilized system for recording title to personal property is surely a first reason. The lack of such a system no doubt raised the fear that buyers or lenders would advance money on personal property without any real chance of learning about a nonpossessory judgment lien.[19] With real estate, prospective buyers or mortgagees typically examine the public record as part of the transactions creating their interests. As long as the judgment lien is recorded where they look, they learn about it and adjust their assessment

[18] The author developed early ideas for this Part in 9 DEBTOR-CREDITOR LAW ¶ 37A.03[B][3] (1989).

[19] Policy makers have been increasingly concerned with ensuring that parties potentially affected get notice of the execution process. *See* 9 DEBTOR-CREDITOR LAW ¶ 37A.03[B][1][a] (1989). The legislative concern with the bona fide purchaser has, however, more recently been questioned in another context. *See, e.g.*, Gilmore, *The Good Faith Purchase Idea and the Uniform Commercial Code: Confessions of a Repentant Draftsman*, 15 GA. L. REV. 605 (1981).

of the transaction accordingly. By contrast, a procedure for filing judgment liens against personalty would have been unlikely to alert those who might purchase encumbered chattels from the debtor. Until the appearance and assimilation of Article 9 of the Uniform Commercial Code (U.C.C.), secured lending consisted of a hodgepodge of legal devices with separate files for record keeping. As it was, buyers and lenders had difficulty determining which files to examine. It is not obvious that another set of files to record judgment liens on personal property would have offered realistic prospects for satisfying a felt need for real notice.

In addition, personal property has been economically inconsequential for most of our history. Indeed, early execution procedures themselves did not even extend to intangible assets,[20] now a main category of personal property. This lack of importance probably eased whatever pressure there otherwise might have been to extend judgment liens to personal property.

Two twentieth-century developments have contributed to the political feasibility of extending judgment liens to personal property. The first is the extraordinary rise in importance of intangible wealth and the legal system's increasing sophistication in dealing with it.[21] The second is the arrival of Article 9 of the U.C.C. with its simplified filing systems.[22]

Article 9 of the U.C.C. permits one to take a security interest in all the debtor's personal property—tangible and intangible—inexpensively and easily. A very simple filing makes the security interest thereby created good against most competing claimants including, in most cases, the trustee in bankruptcy.[23] Article 9 satisfies a craving for notice through its accessible recording system and through complex priority provisions that extend special protection to many who might not be expected to check personal property files.

---

[20] Loyd, *Executions at Common Law*, 62 U. PA. L. REV. 354, 363 (1914).

[21] *See* C. BERGER, LAND OWNERSHIP AND USE 60–63 (3d ed. 1983); Dolzer, *Welfare Benefits as Property Interests: A Constitutional Right to a Hearing in Judicial Review*, 29 ADMIN. L. REV. 525 (1977); Reich, *The New Wealth*, 73 YALE L.J. 733 (1964); Weinberg, *Tort Claims as Intangible Property*, 64 KY. L.J. 49 (1975).

[22] The U.C.C. is simplified only in relation to what preceded it. A national filing system, though possible today with modern technology, has not yet arrived. Currently, secured lenders must cope with a central file in each state and with local county files within each of those jurisdictions. This has resulted in very complex provisions designed to steer filing and searching creditors to the correct file. *See* U.C.C. §§ 9-103, 401 (1978).

[23] One exception is when the debtor gives a security interest to secure preexisting debt. A security interest thus secured can be successfully attacked as a preference in some cases. *See* 11 U.S.C. § 547(b) (Supp. IV 1986).

As will be detailed below, Connecticut, Maine, and California have taken the natural next step by building on these recent commercial law developments. Their procedures for getting a nonpossessory judgment lien are generally to file a simple form within the U.C.C. filing system.[24] Given Article 9, its filing system, and the commercial practice that has developed in its wake, we might now expect such a public filing to alert some potential claimants that the judgment creditor has a claim to the debtor's property. In enacting these provisions, these states have sensibly determined that today's creditors need not seize the debtor's personalty in order to put all competitors on notice of their interests.[25]

B. *The New Provisions*

1. The Statutes Themselves

Connecticut's new provisions are the least complicated of the three. Section 52-355a specifies in part:

> (a) Except in the case of a consumer judgment, a judgment lien . . . may be placed on any nonexempt personal property in which, by a filing in the office of the secretary of the state,

---

[24] In addition, Iowa has a provision which uses the U.C.C. files within the levy process to create a nonpossessory lien in favor of the judgment creditor. IOWA R. CIV. P. 260(b) reads:

> If the creditor or his agent first so requests in writing, the officer may view the property, inventory its exact description at length, and append such inventory to the execution . . . ; and, if the property is consumer goods or if the judgment debtor is not a resident of this state, file with the County Recorder of the county where the property is located his certified transcript of such inventory and statement; and, in all other cases, file with the Secretary of State his certified transcript of such inventory and statement. Such filing shall be accepted by the County Recorder or the Secretary of State as a financing statement . . . and shall be constructive notice of the levy to all persons. . . . The fees normally charged by the County Recorder or Secretary of State for the filing of a financing statement and the filing of a termination statement shall be paid by the officer and shall be taxed by him as a part of his costs of the levy.

Iowa's procedure uses the U.C.C. files to give the judgment creditor an additional way to execute on the judgment. While the new levy does not separate the judgment debtor from his property, the procedure involves the sheriff and execution procedures generally. In that respect they resemble execution procedures more than they do the judgment lien procedures.

Similarly, Minnesota allows a U.C.C. filing to function as a levy "when personal property, by reason of its bulk or other cause, cannot be immediately removed." MINN. STAT. ANN. § 550.13 (Supp. 1987).

[25] The new procedures differ fundamentally from those long existing in Georgia, Alabama, and Mississippi by enlisting accessible and frequently-used U.C.C. files to supply notice to those who might later assert interests in personalty. *See supra* note 9.

> a security interest could be perfected under title 42a. The judgment lien shall be created by filing a judgment lien certificate in the office of the secretary of the state.
>
> . . .
>
> (c) Any such judgment lien shall be effective, in the same manner and to the same extent as a similar security interest under the provisions of title 42a . . . .[26]

By explicitly tying the lien to the interest created under Article 9 of the U.C.C., Connecticut permits, in its title 42a, the judgment creditor to secure her judgment with the equivalent of an Article 9 security interest.

Maine's provision is comparable. It reads in part:

> § 4651-A. EXECUTION LIENS
>
> . . .
>
> 2. LIEN ON PERSONAL PROPERTY. The filing of an execution duly issued by any court of this State or an attested copy thereof with the proper place or places for perfecting a security interest in personal property pursuant to Title 11, section 9-401, subsections (1) and (5) within one year after issuance of the execution shall create a lien in favor of each judgment creditor upon the right, title and interest of each judgment debtor in personal property which is not exempt from attachment and execution and which is of a type against which a security interest could be perfected by filing pursuant to Title 11, section 9-401.[27]

Unlike the Connecticut statute, Maine's provision is ambiguous on questions of priority. As will be developed below, the result is that some battles between judgment lienholders and other claimants will have less than certain results.

California has been the most explicit in defining and refining[28] its new provisions. In their broad compass, California's provisions are similar to the others. The judgment creditor may file a "notice of judgment lien on personal property"[29] with the Secretary of State[30] and obtain priority in the debtor's business property[31] against other claimants largely in accordance with the priority scheme in Article 9 of the U.C.C.

---

[26] CONN. GEN. STAT. ANN. § 52-355a (West 1986).

[27] ME. REV. STAT. ANN. tit. 14, § 4651-A (Supp. 1988).

[28] The California provisions, first enacted in 1982, have gone through at least one major revision since that time.

[29] CAL. CIV. PROC. CODE § 697.550 (West 1987 & Supp. 1989).

[30] CAL. CIV. PROC. CODE § 697.570 (West 1987 & Supp. 1989).

[31] CAL. CIV. PROC. CODE § 697.530 (West 1987 & Supp. 1989).

Both California and Connecticut have tried to exclude consumer debtors from the reach of the new procedures.[32] Maine has no comparable exclusion.

2. Cost and Risk Considerations

Compared with the alternatives available in other jurisdictions, these new procedures are extraordinarily "efficient,"[33] because they reduce the costs and risks of converting a mere judgment into a specific claim to assets. Before considering the priority in specific assets that these new systems supply, it is worth comparing generally the costs of getting something more than an unsecured judgment under these new systems with the costs a creditor must sustain under more traditional systems.

If the creditor's objective is merely to establish priority, the new systems save one the costs of executing on a judgment in order to get priority. Execution in many places involves meticulous document preparation,[34] involvement with the sheriff's office, and, to be successful, personal attention from the lawyer during this extended process. Under the new procedures, the costs of getting priority are those of completing a financing statement form and filing it. In many cases, the cost savings possible under the new schemes are sizeable.

Under the old systems, the general need to direct the sheriff to assets and the relatively high costs of execution make it ill-advised for a creditor to attempt execution without knowing the

---

[32] Connecticut directly excludes judgments against consumers. CONN. GEN. STAT. ANN. § 52-355a (West 1986). Although Connecticut's language, "except in the case of a consumer judgment" could be read as "except in the case of a judgment held by a consumer," it seems more likely that Connecticut attempted to protect consumers rather than disadvantage them by the exclusion. As will be seen *infra* text accompanying notes 127–131, Connecticut might not have delivered all the consumer protection it may have intended.

California's legislation has a similar effect by extending the lien only to U.C.C. categories of property less likely to be held by consumers than businesses, *e.g.*, accounts, chattel paper, equipment, farm products, inventory, and negotiable documents of title. CAL. CIV. PROC. CODE § 697.530 (West 1987 & Supp. 1989).

[33] *See supra* note 4 for a definition of "efficient" as used in this context.

[34] In Pennsylvania, a plaintiff must arrive at the sheriff's office with (1) the original and the correct number of copies of a writ of execution (secured from a different office), P.A.R.C.P. 3108(b); (2) envelopes addressed to all those who have to be served with the writs with postage on them, *id.*; and (3) a "Writ of Execution Notice" which contains advice to the defendant about exemptions and a form through which to claim them. P.A.R.C.P. 3252(a). To have the clerk's office issue the Writ of Execution, the plaintiff must also prepare a "Precipe for Writ of Execution." P.A.R.C.P. 3103, 3251. Each of the forms must be properly completed by the plaintiff or his attorney; the execution will not go forward without a complete and correct package.

whereabouts and character of the debtor's personal property. Acquiring that knowledge is expensive, requiring either investigation or discovery, both of which entail nonrecoverable expenses. Under the new statutes there is no need to learn of the existence and whereabouts of the judgment debtor's personal property. None of the three new systems requires detailed specification of the personal property to be encumbered. Thus, under these systems a judgment creditor might, without any investigation, file a notice broadly describing the types of property to be subjected to the lien and hope that the lien will stick to something of value.

In addition, under the new procedures, one may often avoid the litigation inevitable under traditional procedures. Often the valuable forms of personal property are intangibles such as patent or royalty rights, contract rights, judgments, and claims not reduced to judgment. The law has long been confusing, at best, as to whether such rights may be reached at all and, if so, how one should proceed to execute on them. Thus, a judgment creditor attempting to establish a claim to such property in most states must begin by uncovering answers to these often-indeterminate legal questions. One needs answers at the beginning, because the process of directing the sheriff may include persuading the sheriff that such property can be levied on and instructing the sheriff on how the law specifies that levy be done.

Moreover, once the sheriff acts, the legal questions may well arise in litigation with the judgment debtor about the propriety of the levy. The judgment debtor will be provoked by the lien-creating process of execution to litigate because that process will have directly interfered with the debtor's possession and enjoyment of the property. Thus, one can expect the debtor in many cases to dispute uncertain legal questions if he can afford litigation or to file a bankruptcy petition if he cannot. Either way, unreimbursable legal costs escalate and dilute the value of the judgment. Yet the judgment creditor's alternative—doing nothing—is equally unattractive.

By contrast, in these three states, the judgment creditor's position, *whatever it will later turn out to be*, can be preserved just by filing the notice. The major question that arises at the outset is in which office to file, a far less complex legal question under Article 9 of the U.C.C. than the question how to levy under state law.

In addition, while the liens created under the new statutes are probably as strong as execution liens both in and outside of bankruptcy,[35] the notice itself is not so likely as actual execution to provoke an immediate battle with the debtor.[36] The new liens are nonpossessory, leaving the judgment debtor in control of his property. While the lien may eventually have a serious impact on the judgment debtor's ability to finance his business,[37] the initial provocation the debtor receives with a judgment lien filing is far less than with actual execution.[38]

For the same reasons, the risk that the debtor will immediately respond to a judgment lien notice with a bankruptcy petition seems far less than the risk of such a debtor response to actual execution. If this is true, it follows that the danger of losing priority through a preference attack is lower under the new systems than it was under the old: unless someone files a bankruptcy petition within ninety days of the fixing of the lien, the priority will be largely immune to preference law.[39] Although empirical study is needed to assess the interaction of these new provisions with the bankruptcy system, one would expect far more of these liens to survive bankruptcy than survive under the present system.

In "efficiency" terms then, these new systems warrant high praise. If one believes that the law should enhance what it means to have a judgment, the new liens—even if they were weak and subordinate to many other interests—surely would advance that end. As the discussion will now show, the new state provisions

---

[35] The priorities of the new liens against various competing claimants are discussed in section III, B, 3 of this Article.

[36] A judgment lien filed under the new provisions will be avoidable as a preference if a bankruptcy petition is filed within 90 days of the lien filing. The judgment debtor may not, however, be as directly concerned with who has priority to certain assets under the new systems as he would have been under the old if he maintained possession and use of the assets. Other creditors, of course, might be concerned and provoked by the judgment lien filing to bring an involuntary bankruptcy proceeding against the judgment debtor.

[37] The debtor's ability to finance his business will be affected by priority rules for Article 9 secured creditors. These priority rules are discussed in section III, B, 3 of this Article.

[38] Similarly, execution carries with it the risk that seizure of the debtor's property is not legally warranted; there is thus always a possibility that execution will give a debtor a later claim for wrongful execution if the creditor proceeded to seize property without legal authority. A judgment lien, however, does not deprive the debtor of use of property to the same extent as seizure; therefore, the risk that large damages will accrue following wrongful use of procedure seems lower under the new procedures than under the old.

[39] The bankruptcy law provides an extended period, however, for preferences to "insiders" as defined in the Bankruptcy Code. *See* 11 U.S.C. § 547(b)(4)(B) (Supp. V 1987). The provision is quoted *infra* note 111.

breed strong liens that give judgment creditors substantial priority over competing claimants.

3. Priority Implications: Contests with Those Competing for Debtor Assets

A lien is primarily important to a creditor because it fixes, as of a point in time, the creditor's claim to specific assets against possible competing claims. These new statutes are significant because they allow a judgment creditor to obtain a lien cheaply and with fewer risks of debtor retaliation. But liens vary in quality. For example, some are good against all competitors, with or without actual notice,[40] others fail in various contests with buyers,[41] and still others are specifically excluded from protection under the Bankruptcy Code.[42]

This section will first consider the protection these new state statutes afford those who buy the encumbered property from the debtor after the lien has been filed, and then will examine comparable contests with various Article 9 secured creditors. Part IV will consider contests with a bankruptcy trustee representing unsecured creditors.

a. *Priority Contests with Buyers*

Unlike transactions in land, sales of personal property have not typically featured record searches. The absence of trust-

---

[40] For example, a perfected security interest in industrial equipment is good against buyers, later secured creditors, and, unless preferential, the trustee in bankruptcy. *See* U.C.C. §§ 9-307, -312(5), -301(1)(b).

[41] An example is New York's execution lien which arises at delivery of the writ of execution to the sheriff but is not good against a buyer until the levy is actually made:

> Where a judgment creditor has delivered an execution to a sheriff, the judgment creditor's rights . . . are superior to the extent of the amount of the execution to the rights of any transferee of the debt or property, except:
>
> 1. a transferee who acquired the debt or property for fair consideration before it was levied upon; or
>
> 2. a transferee who acquired a debt or personal property not capable of delivery for fair consideration after it was levied upon without knowledge of the levy.

N.Y. CIV. PRAC. L. & R. 5202(a) (McKinney 1978). A security interest under the U.C.C. will generally be subordinate to rights of a buyer in the ordinary course of business. U.C.C. § 9-307(1).

[42] Landlords' liens, for example, are avoidable by the trustee as "statutory liens." 11 U.S.C. § 545(3), (4) (1982).

worthy records to accommodate movable personal property[43] no doubt contributed to the commercial practice and general legal presumption that possession constitutes the most reliable indication of ownership.[44] Somewhat related to this central place occupied by possession is the law's historic concern for the bona fide purchaser, the innocent buyer who advances money and takes real or personal property without knowledge that it is subject to a competing claim.[45]

Under traditional execution systems, there is little need to worry about persons who might buy without notice of the lien created in the execution process. Many of these systems require a seizure of the property from the judgment debtor to create a lien,[46] and the seizure itself puts any reasonable prospective purchaser on actual notice.[47] In those places where the execution lien can arise before the debtor's property is actually seized,[48] states sometimes protect those who buy without notice of the lien.[49] In any event, the law requires seizure to follow soon after delivery of a writ to the sheriff.[50]

---

[43] The technology is probably available to record reliably the status of each individual's assets, real and personal. While privacy concerns probably will not constrain private industry, these concerns will likely impede efforts to construct in the near future a public file with such comprehensive information.

[44] *See generally* Baird, *Notice Filing and the Problem of Ostensible Ownership*, 12 J. LEGAL STUD. 53 (1983); *cf.* Helman, *Ostensible Ownership and the Uniform Commercial Code*, 83 COM. L.J. 25 (1978); Mooney, *The Mystery and Myth of "Ostensible Ownership" and Article 9 Filing: A Critique of Proposals to Extend Filing Requirements to Leases*, 39 ALA. L. REV. 683 (1988).

[45] *See, e.g.*, Gilmore, *supra* note 19; Murray, *Execution Lien Creditors Versus Bona Fide Purchasers, Lenders and Other Execution Lien Creditors: Charles II and the Uniform Commercial Code*, 85 COM. L.J. 485 (1980).

[46] *E.g.*, ALASKA STAT. § 09.35.110 (1983); CAL. CIV. PROC. CODE § 697.710 (West 1987); IDAHO CODE § 11-201 (1979); MO. R. CIV. P. 76.07; N.C. GEN. STAT. § 1-313 (1983 & Supp. 1988); S.D. CODIFIED LAWS ANN. § 15-18-30 (1984); VA. CODE ANN. § 8.01-478 (1984); WASH. REV. CODE ANN. § 4.56.190 (West 1987).

[47] Where the property cannot be carried away, the law has developed forms of "constructive" seizure, such as immobilizing the property or tagging it, which, similarly, can be expected to put third parties on notice. *See, e.g.*, MD. R. CIV. P. 3-641, -642 (district court) & 2-641, -642 (circuit court); IOWA R. CIV. P. 260(b) (U.C.C. filing); MINN. STAT. ANN. § 550.13 (West 1987) (same).

[48] *E.g.*, ARK. STAT. ANN. § 16-66-112 (1987); COLO. REV. STAT. § 13-52-111 (1973); DEL. CODE ANN. tit. 10, § 5081 (1974); D.C. CODE ANN. § 15-307 (1981); HAW. REV. STAT. § 651-41 (1985); ILL. REV. STAT. ch. 110, para. 12-111 (Smith-Hurd 1985); IND. CODE ANN. § 34-1-34-9 (West 1983); KY. REV. STAT. ANN. § 426.120 (Michie/Bobbs-Merrill 1972); N.H. REV. STAT. ANN. § 528:4 (Supp. 1986); N.J. REV. STAT. § 2A:17-12 (Supp. 1987); N.Y. CIV. PRAC. L. & R. 5234(b) (McKinney 1978); PA. R. CIV. P. 3137 (1988); W. VA. CODE § 38-4-8 (1985).

[49] *E.g.*, N.Y. CIV. PRAC. L. & R. 5202(a) (McKinney 1978) quoted *supra* in note 41.

[50] Many states specify a "return date" of 60 days after which the writ of execution becomes void. States vary on when the 60-day period begins to run. *See* HAW. REV. STAT. § 651-34 (1988) (from issuance); IDAHO CODE § 11-103 (Supp. 1987) (from "receipt"); KAN. STAT. ANN. § 60-2401(c) (1983) (from issuance); MINN. STAT. ANN.

The new statutes could undercut the policy reflected in the older systems. Since the statutes create non-possessory liens in the judgment debtor's personal property and because many prospective buyers would not search public records before purchasing the debtor's property,[51] there is some chance that a court will resolve any dispute in favor of an innocent buyer.[52] Consequently, any legislative reform should resolve contests that may arise between new judgment lien holders and later buyers of the encumbered property.

Besides the importance of clearly resolving such contests, there is the policy question of how to settle the priority issue. On one side of the equation is the desire to strengthen the hand of the judgment creditor by making the debtor's sale of the encumbered property ineffective. On the other side is the desire not to impede free transfers of personal property by requiring buyers to check public files each time they buy personal property. California and Connecticut have articulated their balancing of these policies; Maine has not.

---

§ 550.05 (West 1988) (from "receipt"); MONT. CODE ANN. § 25-13-404 (1989) (from "receipt"); NEV. REV. STAT. § 21-040 (1985) (from "receipt"); N.M. STAT. ANN. § 39-4-9 (1978) (from "delivery"); N.Y. CIV. PRAC. L. & R. 5230(c) (McKinney 1978) (from "issuance"); N.D. CENT. CODE § 28-21-07 (Supp. 1987) (from "receipt"); OHIO REV. CODE ANN. § 2329.53 (Anderson 1981) (from "date"); OKLA. STAT. tit. 12, § 802 (1960) ("from the date thereof"); OR. REV. STAT. § 23.060 (1985) (from "receipt"); S.D. CODIFIED LAWS ANN. § 15-18-41 (1984) (from "receipt" or levy); UTAH R. CIV. P. 69(c) (from "receipt"); VT. STAT. ANN. tit. 12, § 2681 (Supp. 1989) (from "date"); WIS. STAT. § 815.06 (1977) (from "receipt"); WYO. STAT. § 1-17-339 (1977) (from "date"); V.I. CODE ANN. tit. 5, § 474 (Supp. 1986) (from "receipt"). Rhode Island has the longest period of one year. R.I. GEN. LAWS § 9-25-20, -21 (1985).

[51] Because many buyers without actual notice will take property subject to prior perfected Article 9 security interests, *see* U.C.C. § 9-307, at least some can be expected to check the U.C.C. files prior to buying if the property is valuable enough and if they have heard of Article 9.

[52] There is a strong tradition in this direction. For example, the court in Lanterman v. Luby, 114 A. 325, 327 (N.J. 1921), decreed that even where a relevant statute provided that loss of the garage keeper's "control" did not result in loss of its lien, if the legislation had "expressly included subsequent innocent purchasers for value without notice within those against whom the right of seizure [upon loss of "control"] . . . would exist (which it did not), the act would be unconstitutional as a deprivation of property without due process of law . . . ." In the process of reaching this result, the court articulated a strong policy of protection for bona fide purchasers:

> Secret liens upon chattels are an obstruction and a menace to trade, and as such are against the policy of the law. They attempt to contradict and to destroy the universally accepted and natural, as well as legal badge of ownership of chattels, which is possession. The law is most jealous in its protection of an innocent purchaser of a chattel for value without notice, who has relied upon possession as the badge of ownership.

Id. at 326. *Accord In re* Mission Marine Assoc., 633 F.2d 678 (3d Cir. 1980). *See also* Radcliff Finance Corp. v. City Motor Sales, 323 S.W.2d 591 (Tex. 1959).

California has been the most explicit in treating contests between buyers and holders of its new judgment lien and strikes a policy balance comparable to that struck by the drafters of Article 9 of the U.C.C. The legislation protects buyers in the ordinary course of business as defined in U.C.C. § 9-307(1); "[holders] to whom a negotiable document of title has been duly negotiated within the meaning of Section 7-501 of the Commercial Code"; and purchasers of chattel paper who give new value and take possession of the chattel paper in the ordinary course of business. Otherwise, the lien survives the sale of the property.[53]

Like California, Connecticut articulates the policy choice by specifying that its new judgment lien "shall be effective, in the same manner and to the same extent as a similar security interest under the provisions of [the Code]."[54] U.C.C. § 9-307(1) extends protection to buyers "in the ordinary course of business," generally buyers of the debtor's inventory.[55] Buyers of encumbered inventory from retailers are thus protected under both Connecticut and California's legislation. But those who buy encumbered property that the debtor does not sell "in the ordinary course of business" apparently are not protected, even if they buy without actual notice.[56]

In Maine, the outcome of a contest between a new judgment lien holder and a later buyer of the property is unclear, since Maine's legislation, unlike California's, is silent on the rights of buyers. Moreover, unlike Connecticut, Maine does not specify the nature of the lien created by the new filing procedures. This

---

[53] CAL. CIV. PROC. CODE § 697.610 (West 1987).

[54] CONN. GEN. STAT. ANN. § 52-355a (West 1989).

[55] "Buyer in the ordinary course of business" means a person who in good faith and without knowledge that the sale to him is in violation of the ownership rights or security interest of a third party in the goods buys in ordinary course from a person in the business of selling goods of that kind . . . .
U.C.C. § 1-201(9).

[56] U.C.C. § 9-307(2) has been interpreted to offer protection only in transactions in which a consumer is the seller and another consumer is the buyer. *See* J. WHITE & R. SUMMERS, UNIFORM COMMERCIAL CODE § 24.15 (3d ed. 1988). The Connecticut legislation excludes "consumer judgments." *See supra* note 32. U.C.C. § 9-307(3) addresses "future advances," which have no application in the judgment lien context because a judgment lien holder does not make new advances to the debtor.

A different result could follow from § 9-306(2) which allows a purchaser to take free of a security interest if disposition has been authorized by the secured party. While it seems unlikely, allowing a debtor to remain in possession in Connecticut might be seen as implied authorization to sell free of the lien.

defect in the legislation requires resolution by a court or the legislature.

If Maine's lien is interpreted as the equivalent of a security interest, the results of the buyer-judgment creditor contest will be the same as in the two other states. The later buyer will lose unless she was a "buyer in the ordinary course of business" falling within the protection of Code section 9-307(1). If, on the other hand, the judgment creditor holds a lien similar to an execution lien or other judicial lien, a court would have to decide whether to charge buyers with record notice or extend bona fide purchaser protection to all such buyers. Without Article 9 there is no easy way to treat separately the "buyer in the ordinary course of business" who clearly needs protection,[57] and thus a court interpreting the lien in this way probably would feel compelled to protect all buyers without actual notice, including those who buy the judgment debtor's inventory.

b. *Priority Contests with Article 9 Secured Parties*

(1) *Later secured parties.* One readily expects Article 9 secured parties to check the U.C.C. files before they make a secured loan to most debtors.[58] Checking the files is necessary under the Code because the statute grants priority to the first person to place a proper document—the financing statement—in the files.[59] If a judgment creditor places a judgment lien in the correct file, secured parties presumably will see it and take any necessary protective action. It thus offends no principles of notice to award a judgment lien holder priority over a later secured party. And, indeed, it is difficult to justify why a secured

---

[57] Without "buyer in the ordinary course of business" protection, the judgment lien holder would prevail against a buyer of an encumbered clothing store's shirt or of an encumbered appliance store's microwave oven. In the context of modern commercial law, such results would be extraordinary and unsound.

[58] One exception is the retailer who makes loans to consumers to finance their purchases. These seller-lenders take purchase money security interests in the goods they sell. The Code makes their security interests perfected without filing and gives them priority over competing security interests in the same collateral. *See* U.C.C. §§ 9-302(1)(d), -312(4). We can thus expect many such secured parties to operate largely without concern for the files. *But see infra* text accompanying notes 86–88.

[59] U.C.C. § 9-312(5).

lender with record notice of a judgment lien should have priority over the earlier party who already holds a judgment.[60]

As a broad proposition,[61] all three states adhere to the Article 9 approach that the first person to file a correct document or to otherwise perfect an interest covering the personal property will prevail over later parties who have acquired an interest in the same property.

The effect of this policy decision is significant: the holder of a judgment can, by a proper filing, keep the debtor from using the encumbered property for new financing because a new financier will not be able to acquire priority higher than the judgment lien holder's. The lien's ability to choke off the debtor's new financing might, in at least some cases, influence the debtor to pay the judgment without actual levy or garnishment.

As a baseline, Connecticut has incorporated a general first-to-file-or-perfect rule by treating the judgment lien as an Article 9 security interest[62] for priority purposes. California's statute is similar and somewhat more specific.[63] In Maine, regardless of whether the judgment lien is treated as a security interest or judicial lien, a judgment creditor who files before a subsequent secured party will prevail under the U.C.C., since a secured

---

[60] Under the statutes, a secured party who checks the files on day 1, loans money on day 2, and files her financing statement on day 4 will lose to a judgment lien holder who files on day 3 just as she would lose to a secured party who had filed on day 3. While the intervening judgment creditor and intervening secured party might be distinguished on the basis that the former acted without actual reliance on a clear record on day 3 (she would have filed anyway), the statutes appear to make no distinctions in this case. One might well argue that the prudent secured creditor, steeped in Article 9 practice, should be expected to check the files on day 4 before advancing the cash, thereby eliminating the risk of defeat by either an earlier secured party or a judgment lien holder. The U.C.C. permits some later purchase-money lenders to defeat earlier secured parties. The priorities of purchase-money lenders are discussed in the next section.

[61] There are several exceptions to the general first-to-file rule which these states embrace that will be developed *infra* at text accompanying note 65.

[62] CONN. GEN. STAT. ANN. § 52-355a(c) (West Supp. 1989) states that its judgment liens "shall be effective, in the same manner and to the same extent as a similar security interest under [Article 9 of Connecticut's Commercial Code]." U.C.C. § 9-312(5) provides that the first person to file or perfect a security interest will prevail.

[63] The California legislation provides, in part:

> [P]riority between a judgment lien on personal property and a conflicting security interest in the same property shall be determined according to this subdivision. Conflicting interests rank according to priority in time of filing or perfection. In the case of a judgment lien, priority dates from the time filing is first made covering the personal property. In the case of a security interest, priority dates from the time a filing is first made covering the personal property or the time the security interest is first perfected, whichever is earlier, provided that there is no period thereafter when there is neither filing nor perfection.

CAL. CIV. PROC. CODE § 697.590(b) (West 1987).

party loses to either an earlier lien creditor or to an earlier secured party.[64]

(2) *Earlier secured parties—priority as to future advances.* The first-in-time rule employed by all three states obviously means that a judgment lien holder will be subordinate to a secured party who already has filed a financing statement earlier. But Article 9 allows a party already properly secured to make additional later secured loans—future advances—without filing additional financing statements. This innovation forced U.C.C. policy makers to decide what priority the later advance should have over interests arising after the original financing statement but before the future advances.[65] The Code's resolution distinguishes among intervening buyers,[66] intervening lien creditors,[67] and intervening secured parties.[68]

Later secured parties make their loans after searching the files and uncovering the earlier secured party's financing statement; Code drafters determined that later parties should carry the risk that the earlier party will make a future advance. In this situation the Code provides that the future advance carries the same priority as the original advance.[69] While one might debate the wisdom of a scheme that supplies the first secured party with a monopoly on financing,[70] it is difficult to otherwise challenge the fairness of the Code towards the second secured party. Given the notice supplied by the files, the second party can choose not to take on the risk and either walk away from the transaction or seek a subordination agreement from the first secured party.

The situation is different when one obtains a judicial lien on property already subject to a security interest. Unlike the later secured party, a lien creditor cannot "walk away" from the debtor on discovering the earlier party's financing statement.

---

[64] U.C.C. §§ 9-301(1)(b), -312(5).

[65] Courts have struggled with the question. *See, e.g.*, Coin-O-Matic Service Co. v. Rhode Island Hosp. Trust Co., 3 U.C.C. Rep. 1112 (R.I. Super. Ct. 1966); James Talcott, Inc. v. Franklin Nat'l Bank, 194 N.W.2d 775 (Minn. 1972). U.C.C. drafters have addressed the question in at least three places in Article 9. *See infra* notes 66–68.

[66] U.C.C. § 9-307(3) gives the creditor who made a future advance priority over an intervening buyer unless the secured party knew of the sale or made the advance more than 45 days after the sale.

[67] U.C.C. § 9-301(4).

[68] U.C.C. § 9-312(7).

[69] *Id.*

[70] *See* Jackson & Kronman, *Secured Financing and Priorities Among Creditors*, 88 YALE L.J. 1143, 1179–80 (1979).

The lien is non-consensual, follows perhaps costly litigation, and is an attempt to secure an old debt. Once such a lien is asserted, it is much harder to justify a rule giving an earlier secured party her original priority in all future advances. Code drafters have recognized the fundamental difference between a later secured party and a later lien creditor by crafting a special rule giving future advances much more limited protection against a later judicial lien. Article 9 gives the future advance priority over the judicial lien only if the advance was made within forty-five days after the lien or the secured party had no notice of the lien.[71]

The three states seem to diverge on the question of what priority to award a secured party's advances made after the arrival of a new judgment lien. California has determined that the proper analogy is that of the judicial lien holder under Article 9. The other two states are silent on the specific question, but one can draw some tentative conclusions from their legislation.

California explicitly addresses the priority contest that will occur between a judgment lien holder and an earlier secured party who makes future advances. Section 697.590 provides in part:

> (f) A judgment lien that has attached to personal property and that is also subordinate . . . to a security interest in the same personal property is subordinate . . . only to the extent that the security interest secures advances made before the judgment lien attached or within 45 days thereafter or made without knowledge of the judgment lien or pursuant to a commitment entered into without knowledge of the judgment lien . . . . [A] secured party shall be deemed not to have knowledge of a judgment lien on personal property until (1) the judgment creditor serves a copy of the notice of judgment lien on the secured party personally or by mail and (2) the secured party has knowledge of the judgment lien on personal property, as "knowledge" is defined in Section 1201 of the Commercial Code.[72]

---

[71] The pertinent section provides:

> A person who becomes a lien creditor while a security interest is perfected takes subject to the security interest only to the extent that it secures advances made before he becomes a lien creditor or within 45 days thereafter or made without knowledge of the lien or pursuant to a commitment entered into without knowledge of the lien.

U.C.C. § 9-301(4).

[72] CAL. CIV. PROC. CODE § 697.590(f) (West 1987).

It should be easy to see the enormous effect a judgment lien in California might have on continued secured financing from the debtor's primary creditor. In revolving financing arrangements, the same secured party will make additional loans to the debtor as earlier loans are repaid or as the collateral changes. As discussed above, under Article 9 those future advances are protected against later secured parties to the same extent as the original advance. By making its new lien a judicial lien, California has given its judgment creditors the power to destroy these financing arrangements by simply filing the equivalent of a financing statement, serving notice, and waiting the requisite period.

If the lienholder gives the proper notice,[73] the original secured party may safely make advances only for forty-five days after the lien has been filed. After that, the advances will be subordinate to the judgment lien. As a practical matter, one would expect the secured party to refuse to make additional advances after that point and, perhaps, to begin to terminate the financing arrangements even sooner. This will obviously supply the judgment creditor with immense leverage to effect payment of the judgment without actual execution.[74]

Connecticut has treated the issue less explicity and, perhaps, differently. The Connecticut statute says only that its judgment lien is to be treated as a "similar security interest."[75] If the legislature intended the judgment lien to be treated as a security interest in this context, it made a policy choice very different from California's, because the Code does not give priority to a second secured party over the first secured party's later advances.[76] Therefore, if the Connecticut lien is treated as a security interest, the judgment holder will not be able to affect the judgment debtor's ongoing financing arrangements. Thus, the judgment holder in Connecticut will have a substantially weaker lien than a similar party in California. Such an interpretation would ignore fundamental differences between a judgment holder getting a lien to secure an old debt and a later secured party considering whether to make a new loan.[77]

---

[73] Given the California notice provision, quoted *supra* at text accompanying note 72, providing notice to secured parties of record should be part of a lawyer's standardized process of obtaining a judgment lien following entry of judgment.

[74] *See also infra* note 114.

[75] The Connecticut statute is quoted *supra* at text accompanying note 26.

[76] U.C.C. § 9-312(7). *See supra* text accompanying notes 69–70.

[77] *See supra* text accompanying note 70.

Yet, perhaps, Connecticut determined that it was simply bad policy to supply a party holding a judgment with this extraordinary leverage over the debtor's continuing financing. Current U.C.C. provisions were drafted against the backdrop of a traditional system which made a creditor's obtaining a lien on personalty a rather extraordinary event. The new statutes may make nonconsensual liens on personalty commonplace with unpredictable effects on a local economy. Connecticut may have decided that the stability of ongoing business financing was not outweighed by the need to secure the payment of judgments and may have intentionally resolved this policy question in favor of the pre-existing secured party.

As will be seen in Part IV, the leverage given a judgment holder is, ultimately, a deeply political question to which there are no simple answers. The policy choices imbedded within these new statutes inevitably require a delicate balancing of competing interests. Given the actual language of the Connecticut statute and the uncertainty of its legislature's policy choice, a strict reading of the statute to favor the pre-existing secured party is probably appropriate until the legislature clarifies its preference.

Maine's legislation is silent on how to treat the new lien in this context. By specifying that the proper procedure is to file an execution in the U.C.C. files and that the procedure "shall create a lien,"[78] the terminology suggests that the new lien should be considered a judicial lien and not a security interest. If that is the case, the results of this contest will be consistent with those in California.

(3) *Purchase money secured parties.* One of the U.C.C.'s innovations was to permit a party in a single transaction to acquire a security interest in property the debtor would acquire in the future, that is, in after-acquired property. Under the Code's priority provisions, priority goes to the first party who files a financing statement covering the collateral regardless of when the collateral was acquired.[79] To this general rule the U.C.C. makes an exception for purchase money secured parties—generally, parties that finance the debtor's purchase of the

---

[78] Maine's statute is quoted *supra* at text accompanying note 27.
[79] U.C.C. § 9-312(5).

collateral.[80] If a secured party meets various Code requirements, he can acquire priority in the new collateral over a pre-existing secured party despite that party's earlier filing.[81]

This strong policy of special priority for purchase money secured parties has a long history. Before the arrival of the Code, courts reached the Code result by manipulating property concepts: the earlier creditor's security interest never attached to the new collateral because the debtor had insufficient rights in that collateral to grant a security interest to the first party.[82] More recently, the policy has been justified on the basis that the purchase money secured party has added specific value to the debtor's assets, that this additional financing option is useful, and that it harms no one to give the new secured party priority as to that added value.[83]

When Article 9 was drafted, it was not necessary to give great thought to whether a purchase money secured party should take priority over a pre-existing lien creditor. This is because in most states actual execution on specified, existing assets was the sole method available to a judgment creditor for obtaining a lien on personal property.[84] The only after-acquired property that could be reached was property the debtor acquired between the time the judgment creditor delivered the writ of execution and the time the sheriff levied. U.C.C. § 9-301(2) addresses the problem in this way:

> If the secured party files with respect to a purchase money security interest before or within ten days after the debtor receives possession of the collateral, he takes priority over the rights of a transferee in bulk or of a lien creditor *which arise between the time the security interest attaches and the time of filing*.[85]

One might infer from the rule that if the lien arises either before or at the same time as attachment of the security interest, the lien creditor will win.

In the context of most state systems, the rule is consistent with the strong policy of facilitating purchase money lending

---

[80] U.C.C. § 9-107.

[81] U.C.C. §§ 9-312(3), -312(4).

[82] *See, e.g.*, United States v. New Orleans R.R., 79 U.S. (12 Wall.) 362, 364–65 (1871).

[83] *See generally* Jackson & Kronman, *supra* note 70, at 1164–78 (1979).

[84] *But see supra* note 9.

[85] U.C.C. § 9-301(2) (emphasis supplied).

that is found in the Code.[86] The risk that a purchase money loan will be defeated by an earlier lien creditor is very low as long as the secured party files within 10 days. Many states require actual levy in order to obtain a lien,[87] and levy cannot occur until the debtor acquires some rights in the collateral. In those states that permit an execution lien to arise on delivery of the writ to the sheriff,[88] the lien will typically be of short duration and may not reach after-acquired property.

The situation is different under these new systems, of course, because the liens are easier to get, last longer, and reach after-acquired property. Like security interests, the liens will arise when the debtor gets rights in the new collateral, that is, *at the same time* the purchase money security interest attaches.[89] Under a strict reading of current Code language that addresses the contest between the lien creditor and purchase money secured party,[90] the new liens might defeat many purchase money secured parties and thus could pose a substantial risk to purchase money lending.

On the other hand, if one wished to further strengthen the hand of the judgment creditor, one could give these liens priority over later purchase money lenders. The effect would be to force sellers and others who finance purchases to check the files before lending and to decline purchase money loans if the debtor's personal property is found to be encumbered with a judgment lien. Policy makers should weigh the probable economic costs of forcing purchase money lenders to check files prior to lending[91] against the benefits of further strengthening the judgment creditor's hand.

---

[86] U.C.C. §§ 9-312(4), -313(4)(a), and -314 are all illustrations of the Code policy that a person who finances the debtor's acquisition of new collateral ought to have priority *as to that collateral* over pre-existing secured parties or encumbrancers.

[87] *See supra* note 46.

[88] *See supra* note 48.

[89] A security interest can attach no earlier than the time the debtor acquires "rights in the collateral." U.C.C. § 9-203(1)(c).

[90] The provision is set out *supra* at text accompanying note 85.

[91] Where the contest is between a purchase money secured party and another secured party, Article 9 has resolved this policy issue in favor of the purchase money party. U.C.C. § 9-312(4).

Except in the case of financing a debtor's acquisition of inventory, a purchase money financier can be assured of priority without checking the U.C.C. files prior to lending. If the Code requirements are met, the later purchase money lender will have priority regardless of what she would have found in the files.

A policy maker might ask whether a judgment lienholder who has engaged in litigation to collect an old debt should be treated better than a secured party who, while not owed an old debt, is induced to lend money by the security her purchase money security interest will provide.

Because Article 9 was drafted in an era devoid of judgment liens on personal property, one can only speculate how its drafters would have treated the contest between a judgment lienholder and a later purchase money secured party. But the strong policy in the Code of protecting a purchase money lender by awarding priority over a pre-existing lienholder[92] suggests that the drafters probably would have protected the purchase money lender in this context as well.

California is the most explicit in determining the contest between a judgment lien holder and a later purchase money lender. Its legislature drafted a specific provision which favors the purchase money lender by providing:

> A purchase money security interest has priority over a conflicting judgment lien on the same personal property or its proceeds if the purchase money security interest is perfected at the time the judgment debtor receives possession of the personal property or within 10 days thereafter.[93]

A court in Connecticut, treating its lien as a similar security interest,[94] could reach a result similar to California's by ruling that the contest is determined by the U.C.C. provisions that resolve priority contests[95] between holders of security interests and later purchase money lenders.[96]

---

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

[93] CAL. CIV. PROC. CODE § 697.590(d) (West 1987).

[94] The Connecticut statute is quoted *supra* at text accompanying note 26.

[95] U.C.C. § 9-312(4) provides:

> A purchase money security interest in collateral other than inventory has priority over a conflicting security interest in the same collateral or its proceeds if the purchase money security interest is perfected at the time the debtor receives possession of the collateral or within ten days thereafter.

[96] U.C.C. § 9-312(3) establishes far more specific requirements that must be met before a purchase money lender on inventory can get priority over a pre-existing security interest in inventory. Among other things, the rule requires that the purchase money lender's security interest be perfected at the time the debtor receives possession of the inventory and that the purchase money lender notify the earlier secured party before the later lender files. U.C.C. §§ 9-312(3)(a), -312(3)(b). Since the Connecticut judgment lien reaches the judgment debtor's inventory, if U.C.C. § 9-312(3) is applied in the judgment lien context, the practical effect might well be to seriously impede a judgment debtor's acquisitions of new inventory through purchase money loans. On encountering the judgment lien within the files, the new lender might well reconsider making the loan. Or, on receiving the required notification from the purchase money lender, the judgment lien creditor might be prompted to begin execution proceedings.

While this result seems inoffensive as a policy matter, it differs from the result the California legislature mandated, because California made no distinctions between purchase money lenders on inventory and others. *See also* CAL. CIV. PROC. CODE § 697.590(e) (West 1987), which resolves a circular priority problem created by its judgment lien priority provision in the inventory financing setting.

By failing to define the nature of its lien or to specify priority rules,[97] the Maine legislature left the contest between a judgment lien holder and a later purchase money lender unsettled. If the lien is treated like a judicial lien, the U.C.C. rule[98] that resolves contests between holders of judicial liens and later purchase money secured parties will apply, and a strict reading of that rule could defeat a purchase money secured party who filed after the judgment lien was in place. Such a result would be at odds with the U.C.C. policy of protecting purchase money lenders.

Moreover, if a court ruled that a purchase money lender must check the files before lending to protect against earlier judgment liens, it would undermine the U.C.C. rule by forcing a pre-transaction file check in a situation where the Code drafters thought it unnecessary. Maine's legislation ought to be clarified on this point. By the same token, because a state has the capacity through these liens to undermine purchase money lending, U.C.C. policy makers might well consider a revision to the Code to accommodate this new kind of judicial lien.[99]

IV. Impact of the Provisions on the Larger System

A. *Contests Between Lien Holders and Unsecured Creditors: Implications for Bankruptcy*

It is a basic tenet of our debtor-creditor system that unsecured creditors have no claims to specific assets. They must themselves secure judgments on their claims and obtain liens before they are able to assert priority in specific assets. Consequently, there is no real contest between an ordinary unsecured creditor and any lien holder: within the state systems, a lien holder will defeat most creditors who do not have liens.

---

[97] The Maine legislation is quoted *supra* at text accompanying note 27.

[98] U.C.C. § 9-301(2) is quoted *supra* at text accompanying note 85.

[99] One final issue which policy makers might want to consider is the extent to which these new liens in a few jurisdictions affect desired uniformity of the U.C.C. The Code was drafted before such liens were possible and against a backdrop of the inefficient creditor enforcement system. As the text suggests, the new liens interact with secured lending in a different way than did execution liens under older systems. Given the major change these new statutes bring and given the state-to-state variation even in the new provisions, have we begun injecting further complexity into our basic commercial legislation whose strongest attribute is supposed uniformity?

Unsecured creditors are, however, represented by the trustee in bankruptcy and collectively get what is left of the debtor's assets after the claims of valid lien holders have been satisfied. This competition for the debtor's limited assets combined with the bankruptcy law's recognition of state-created liens[100] make bankruptcy, in many respects, a contest between secured and unsecured creditors. Because the debtor's finite assets must be divided among secured and unsecured claims, any state law that creates new liens that are enforceable in bankruptcy carries with it the potential of upsetting the current balance between secured and unsecured creditors. Although there may be nothing sacrosanct about the present balance, when evaluating the new provisions, policymakers may want to consider whether the new judgment lien provisions alter the current distribution. Some policymakers may consider an altered distribution in bankruptcy to be an undesirable side effect of the "efficiency" brought about by the new statutes.

The present balance of distribution amongst claimants within the bankruptcy process will remain undisturbed if the Bankruptcy Code enables trustees to avoid the new liens. It is, after all, federal bankruptcy law's deference to state law-created liens and security interests that ties the welfare of lien creditors to the claims of unsecured claimants, and Congress certainly has the power to make such liens ineffective in bankruptcy.[101] Trustees will probably be unable to avoid the new liens, however, because under current law, any attack on the liens is likely to fail.

The Bankruptcy Code's strong arm provision[102] gives the trustee the power, as of the date of the bankruptcy filing, of a "creditor on a simple contract [with a] . . . judicial lien."[103] The trustee's challenge under this strong arm power, however, will probably fail. In all three states, an earlier filed lien will defeat a later lien.[104] Therefore, a person (here read trustee) with a

---

[100] This recognition seems unlikely to be constitutionally based. *See* Rogers, *The Impairment of Secured Creditor's Rights in Reorganization: A Study of the Relationship Between the Fifth Amendment and the Bankruptcy Clause*, 96 HARV. L. REV. 973 (1983).

[101] Congress, for example, made liens "for rent" and "of distress for rent" avoidable as a class in bankruptcy. 11 U.S.C. § 545(3)-(4) (1982).

[102] 11 U.S.C. § 544(a) (Supp. V 1987).

[103] *Id.*

[104] The same rule that applies to contests between judgment lien holders and later secured parties will apply to later lien holders as well and thereby defeat the trustee. *See supra* text accompanying notes 59–61.

judicial lien arising after the judgment lien will be subordinate. The trustee's other strong arm powers would probably not fare any better.[105]

An attack on these liens as statutory liens under Code section 545(2)[106] also seems unlikely to succeed. Section 101(47) of the Code defines a statutory lien, in part, as a "lien arising solely by force of a statute on specified circumstances or conditions . . . , but does not include security interest or judicial lien."[107] The Code in turn defines judicial lien as a "lien obtained by judgment, levy, sequestration, or other legal or equitable process or proceeding."[108] Since the new judgment liens are "obtained by judgment," they appear to be judicial liens rather than statutory liens under the Bankruptcy Code's definitions. Moreover, even if one gets beyond these definitional problems, avoidable statutory liens must fail the bona fide purchaser[109] test found in section 545(2). Such failure seems unlikely, because judgment liens are valid against later secured parties—purchasers under the Bankruptcy Code's definitions.[110]

Like all other liens and security interests, judgment liens on personal property will be subject to attack as preferences, pro-

---

[105] Congress also gave the trustee the rights and powers of:

> (2) a creditor that extends credit to the debtor at the time of the commencement of the case, and obtains, at such time and with respect to such credit, an execution against the debtor that is returned unsatisfied at such time, whether or not such a creditor exists; or
>
> (3) a bona fide purchaser of real property, other than fixtures, from the debtor, against whom applicable law permits such transfer to be perfected, that obtains the status of a bona fide purchaser and has perfected such transfer at the time of the commencement of the case, whether or not such a purchaser exists.

11 U.S.C. § 544(a)(2)–(3) (Supp. V 1987).

[106] 11 U.S.C. § 545(2) permits the trustee to:

> avoid the fixing of a statutory lien on property of the debtor to the extent that such lien—
>
> . . .
>
> (2) is not perfected or enforceable at the time of the commencement of the case against a bona fide purchaser that purchases such property at the time of the commencement of the case, whether or not such a purchaser exists.

[107] 11 U.S.C. § 101(47) (Supp. V 1987).

[108] 11 U.S.C. § 101(32) (Supp. V 1987).

[109] See *supra* note 106 for the text of 11 U.S.C. § 545(2), which describes the circumstances under which a statutory lien may be avoided. "Purchaser" is defined in the Code as "transferee of a voluntary transfer," 11 U.S.C. § 101(37) (Supp. V 1987), which would include lenders who take security interests as well as buyers.

[110] 11 U.S.C. § 101(37) (Supp. V 1987). A "buyer in the ordinary course of business" may be able to defeat the liens, *see supra* text accompanying notes 53–57, but this is a far narrower class of buyers than "bona fide purchaser" as defined in the Bankruptcy Code.

vided they meet the requirements of section 547.[111] But as suggested earlier, since fixing these liens does not deprive the debtor of possession of personalty as does execution, a precipitous bankruptcy filing may be less likely as a matter of course. A debtor receiving notice of a judgment lien is unlikely to have the same reaction as, for example, a debtor whose property has just been seized by the sheriff in satisfaction of a judgment.

Moreover, Connecticut,[112] unlike Maine,[113] does not explicitly require the judgment creditor to notify the debtor that she has filed her judgment lien notice. In some cases this means that the debtor may not discover the lien soon enough to file a bankruptcy petition and capture the encumbrance within the statutory preference period. In short, it seems that these liens are less likely to be avoided as preferences than the execution liens which these judgment liens have partly displaced.[114]

Since it is likely that these liens will be enforceable in bankruptcy, the present distributional balance between secured and unsecured creditors may be significantly altered. Claimants in

---

[111] This section will apply only if there is a:

"transfer of an interest of the debtor in property—
(1) to or for the benefit of a creditor;
(2) for or on account of an antecedent debt owed by the debtor before such transfer was made;
(3) made while the debtor was insolvent;
(4) made—
(A) on or within 90 days before the date of the filing of the petition; or
(B) between ninety days and one year before the date of the filing of the petition, if such creditor at the time of such transfer was an insider; and
(5) that enables such creditor to receive more than such creditor would receive if—
(A) the case were a case under chapter 7 of this title;
(B) the transfer had not been made; and
(C) such creditor received payment of such debt to the extent provided by the provisions of this title.

11 U.S.C. § 547(b) (1982 & Supp. V 1987).

[112] CONN. GEN. STAT. ANN. § 52-355a (West Supp. 1989).

[113] Maine's provision specifies that:

[a] lien created by this section shall become void with respect to the right, title and interest of any particular judgment debtor, unless the judgment creditor notifies the judgment debtor by certified or registered mail sent to his last known address on or before 20 days after filing or recording of the existence of the lien.

ME. REV. STAT. ANN. tit. 14, § 4651 (Supp. 1988). CAL. CIV. PROC. CODE § 697.560 (West 1987) also requires service of the notice of the judgment lien on the judgment debtor "at any time of filing . . . or promptly thereafter." Yet the notice requirement may be negated by the fact that the Code also provides that "[t]he failure to comply with this requirement does not affect the validity of the judgment lien." *Id.*

[114] A debtor in a revolving financing arrangement, *see supra* text accompanying note 74, could avoid the lien if it interfered with the financing arrangements but, in this case, filing a bankruptcy petition to do so might be a cure worse than the disease.

these three states may well file complaints and obtain default judgments because a claim to specific assets can so easily follow such a process. Thus, as suggested earlier,[115] these cheaper and more efficient procedures for obtaining liens on personalty will probably result in more liens on personalty.

In addition, one can expect the new judgment liens to reach beyond narrow categories of property. Given the low expense and low risk of obtaining these judgment liens, judgment holders are unlikely to investigate a debtor's assets and assert interests only in the limited categories of personal property that they find. Rather, their lawyers are apt to mass produce judgment lien documents to assert priority in all personal property,[116] thereby making expensive individual treatment unnecessary.

All of this means that in these states more debtors who enter bankruptcy should enter with their property already encumbered by these new liens. States that have enacted these systems have, in short, set up a legal regime that may redistribute property from unsecured claimants to the new lien holders in bankruptcy.[117]

These distributional consequences may be worse for some unsecured creditors than for others. The impact of these new liens will be felt only by those who would have been paid something in bankruptcy had the liens not been in effect. The most likely classes of unsecured creditors to be affected in bankruptcy, therefore, are those near the top of the detailed federal priority scheme.[118] Employees, pension plans, taxing authorities, and other priority claimants, therefore, have more to fear than others from a reallocation of assets from unsecured to secured creditors in bankruptcy.

From the perspective of both policy makers and all unsecured bankruptcy claimants, perhaps the most important group of un-

---

[115] *See supra* text accompanying note 33.

[116] The lawyers' forms will likely assert claims to all the debtor's personal property, including after-acquired property, in words which meet the then-current U.C.C. test for specificity within a financing statement. Current cases hold descriptions such as "all assets" as inadequate and descriptions in U.C.C.-defined categories as adequate. *See* J. WHITE & R. SUMMERS, *supra* note 56, § 22-18, at 1040–44. It seems likely that the forms will simply list all U.C.C.-defined categories of collateral.

[117] Yet these results may not come to pass. The degree of reallocation from unsecured to judgment lien claimants will depend on the amount of unencumbered assets that would have been available for distribution absent the new provisions. If, for example, most California judgment debtors' assets are fully encumbered to begin with, the new judgment liens will have little impact on unsecured creditors in any event.

[118] 11 U.S.C. § 507 (1982 & Supp. V 1987).

secured creditors is the first priority class[119] of administrative claimants, including trustees who are paid by the estate to preserve the assets for the benefit of all unsecured creditors and to avoid liens, preferences, and other transfers. Since much of the bankruptcy system's operation is financed by administrative expenses, the new liens have the potential of draining money away from the bankruptcy system itself. If the new liens result in more cases where there are insufficient unencumbered assets to pay a private bankruptcy trustee, the system may have to rely on federal officials to oversee these no-asset bankruptcy cases that would have been asset cases under the traditional system.[120] The impact such a shift in responsibility would have on the federal budget and the taxpayer is uncertain but could be substantial.

Since the new judgment lien provisions are theoretically available to all unsecured creditors, it could be said that neither priority nor non-priority unsecured creditors have cause to complain. But for many creditors, the theoretical ability to use the new provisions is of little solace and the possibility of adjusting credit practices in light of the new provisions is limited. Employees, for example, will be unable to get judgments for their wages in time to avail themselves of the new provisions and will be unlikely to get security interests for their unpaid wages.[121] Tort creditors do not engage in consensual credit transactions and do not decide for themselves whether to extend credit. Buyers of goods who have already paid for them can be creditors on warranty claims, but claims for breach of warranty arise *after* the warranty has been extended.[122] Small or legally unsophisticated creditors may be unable to justify the expense of getting the judgment necessary to deploy the new procedures. And, administrative expense claimants only begin their work when most of the estate's assets are already spoken for.

---

[119] 11 U.S.C. § 507(a)(1) (1982).

[120] 28 U.S.C. § 586(a)(2) (1982).

[121] *See* 3A A. CORBIN, CONTRACTS § 676, at 209:

> Why must the employee give "credit" and the employer not? Why must the employee carry the risk of getting nothing for his labor, while the employer does not carry the risk of getting no labor for his money? The answer is that such is the almost universal custom of men.

Unsecured employee claims may not, however, be a great problem in fact. One commentator maintains that numbers of wage claims in bankruptcy are very low, because employers in financial trouble want least to precipitate trouble from their employees. *See* Buckley, *The Bankruptcy Priority Puzzle*, 72 VA. L. REV. 1393, 1407 (1986).

[122] *See* Leff, *supra* note 3, at 20. *But cf.* Buckley, *supra* note 121, at 1393, 1407.

Thus, while the new judgment lien provisions are technically available to everyone, they might actually tend to favor larger creditors and those in businesses that have reduced the unit costs of obtaining judgments. Put another way, the main beneficiaries of these provisions may be those that, as an empirical matter, get judgments most easily. The losers may be those who have the most difficulty obtaining judgments on their claims.

Ordinarily, questions of priority among various classes of creditors occupy a prime position in public policy debates in the debtor-creditor field. The set of priorities within the Bankruptcy Code[123] was itself the product of long deliberation. Similar priority issues in the new judgment lien statutes, however, have engendered little or no political or policy debate despite the fact that some creditors will be much better than others at obtaining judgments. Whether differences in judgment-getting potential stem from the nature of the underlying claims (*e.g.*, loan defaults versus personal injury), legal sophistication, access to legal resources, or other factors, in these states some creditors will enjoy priority over other state claimants and even over priority claimants in bankruptcy. Who the privileged creditors are is an empirical question; however, it seems nearly certain that they are *not* employees claiming wages, tort claimants, consumers with warranty claims, or persons without routine access to the resources required to get a judgment. This de facto subordination of whole classes of creditors not only entails a normative decision about which creditors are more deserving of recovery, it may also alter the system that supplies creditors with incentives to discover hidden debtor assets.

Under the traditional systems which award priority only to creditors who locate and seize property, claimants have a powerful incentive (*i.e.*, priority) to spend money in a search for a judgment debtor's hidden personal property. Such investigatory activity may be useful, since it improves the lot of creditors generally by increasing the total assets available for collection. Therefore, as a policy matter at the state and federal level, one might appropriately reward with priority those claimants who engage in the activity of uncovering debtors' hidden assets, regardless of the underlying nature of their claims.

By contrast, the new systems provoke no similar investigatory activity, because it is unnecessary to locate a judgment debtor's

[123] 11 U.S.C. § 507 (1982 & Supp. V 1987).

property in order to obtain priority over it. Getting a judgment and filing a document are all that are necessary to fix priority in the debtor's assets, whatever they happen to be. Once someone has filed for a judgment lien, remaining claimants have no incentive to engage in investigatory activity, since chances are they will lose to the earlier-filed judgment lien holder, even if they locate hidden assets and levy on them.[124] Policy makers should question whether the simple activity of getting a judgment and filing a document is an activity that justifies the award of priority regardless of the merits of the claim.

What is most troublesome is that these questions of priority may not have been raised when these new statutes were considered. State policy makers appraising the new provisions and federal policy makers contemplating changes to the bankruptcy laws should consider the significant changes that the judgment lien statutes may create within the preexisting priority system before advising adoption of these more "efficient" procedures.

The impact on the bankruptcy process may be yet more subtle and difficult to quantify. One can hypothesize, for example, that fewer business debtors will have the unencumbered assets that make the bankruptcy process worthwhile for trustees and unsecured claimants in the first place. Many such debtors may avoid bankruptcy altogether in these states and simply abandon their property to the lien holders. And while large numbers of the new liens may not make Chapter 11 reorganizations disappear, the liens could make it more difficult for a debtor to reorganize than it was before.[125]

If this is the case, is the diminished bankruptcy activity that will follow desirable? Does or should federal policy protect the bankruptcy process and its federal priority claimants from state procedures that, on their face, are available to all yet tend to favor some over others?[126] If the new judgment lien provisions

---

[124] Judgment liens on real property suffer from the same sorts of problems and might well be attacked on the same basis. They might be distinguishable from judgment liens on personal property, because real property may be harder for a debtor to conceal. In addition, there may well be less unencumbered real property in difficult cases than unencumbered personal property.

[125] A reorganizing debtor has far less latitude when dealing with holders of "secured claims" under the Bankruptcy Code than with "unsecured claims." *See, e.g.*, 11 U.S.C. § 1129(b)(2)(A) (Supp. V 1987). If enforceable in bankruptcy, a new judgment lien would qualify as a "secured claim." 11 U.S.C. §§ 506(a) (1982), 101(33) (1982 & Supp. V 1987).

[126] The question whether the bankruptcy law should, as a matter of policy, affect a distribution of assets different from that mandated by state law has received recent scholarly attention. *See* Warren, *Bankruptcy Policy*, 54 U. CHI. L. REV. 775 (1987);

are enacted by other states, these questions may become more pressing.

B. *Effects on Debtor Protection—Direct Impact on Exemptions*

Exemption provisions generally provide that some portion of the debtor's property is not available to creditors to satisfy their judgments. These provisions vary immensely from state to state but generally arise from a concern that the judgment debtor not be reduced to total destitution and dependency through the operation of the execution statutes. All exemption provisions reflect a general legislative judgment that preserving some amount or types of debtor property is more important, for one or more reasons, than allowing the collection of debts from that property. Execution statutes typically contain procedures through which a debtor can raise an exemption claim and through which, if the claim is sustained, the property cannot be reached by the execution.

The new liens will probably affect the protection afforded debtors by exemption statutes because they handle exemption rights differently from the old statutes. Connecticut has apparently[127] tried to keep these provisions from having an impact on consumers by excluding consumer judgments from the reach of these provisions.[128] Nonetheless, since the legislation treats the lien as a security interest,[129] if the lien were to reach the property of consumers it could have a devastating impact, because the law typically regards a debtor's exemption rights as subordinate to the rights of one possessing a security interest.[130]

Baird, *Loss Distribution, Forum Shopping, and Bankruptcy: A Reply to Warren*, 54 U. CHI. L. REV. 815 (1987).

[127] *See supra* note 32.

[128] The Connecticut legislation is quoted *supra* at text accompanying note 26.

[129] CONN. GEN. STAT. § 52-355a(c) (Supp. 1988).

[130] In most states, an exemption will be invalid against one holding a security interest in the exempt personal property. *See generally* Haines, *Security Interests in Exempt Personalty: Toward Safeguarding Basic Exempt Necessities*, 57 NOTRE DAME L. REV. 215 (1981).

This rule has not been lost on creditors. It has been an open secret for some time that creditors could simply and easily gain access to exempt property by getting the debtor to give a security interest in it. And this was so despite a near-universal rule that executory waivers of exemption rights were ineffective. *Id.*

The situation has been remedied somewhat by a provision in the Bankruptcy Code making many such security interests ineffective against exempt property, 11 U.S.C. § 522(f)(2) (1982), and by a similar rule enacted by the Federal Trade Commission to apply outside the bankruptcy context. 16 C.F.R. § 444.2(a)(2) (1984).

The Connecticut legislation is explicit in treating its lien as a security interest and, without substantial judicial counterspin, the lien could defeat a judgment debtor's exemption rights in the targeted property. Yet, in Connecticut one cannot obtain such a lien based on a consumer judgment. The question, then, is whether there are nonconsumer judgments that could be entered against those for whom exemption rights might be important. There may well be.[131]

Instead of excluding consumer judgments from the scope of its judgment lien legislation, California has limited its liens to certain categories of property. The legislation specifies that judgment liens can be acquired against accounts receivable, chattel paper, equipment (not including motor vehicles), farm products, inventory, and negotiable documents of title.[132] By its terms, the California legislation excludes virtually all types of property its exemption statutes protect. One exception, however, is the California exemption for "personal property used in trade, business, or profession," the traditional tools-of-the-trade exemption, which is limited in California to a relatively generous value of $2500.[133] Much of a small business debtor's business equipment fits this common exemption, thereby raising a question whether the judgment lien is subordinate to this exemption right.

The California statutes do not explicitly answer this question.[134] However, since the legislation tends to treat its new lien as a judgment lien in most respects and not as a security interest, a sensible resolution is that the lien is inferior to the exemption rights as are more traditional judgment liens elsewhere.

Unlike the statutes of Connecticut and California, Maine's provision is explicit on the status of exemption rights. The lien only covers those types of "[real and] personal property which

---

[131] Recently completed research confirms what many practitioners are already aware of: a substantial number of individual bankruptcies involve small businesses in which the business assets and the personal assets of the principal are hopelessly intertwined. *See generally* T. SULLIVAN, E. WARREN & J. WESTBROOK, AS WE FORGIVE OUR DEBTORS 108–27 (1989). Whether because debtors have given personal guarantees for the debts of their corporations or because their businesses lack a corporate form to shield their personal assets, creditors might seek potentially exempt personal assets for business-related debts.

[132] CAL. CIV. PROC. CODE § 697.530(a), (d)(1) (West 1987).

[133] CAL. CIV. PROC. CODE § 704.060 (West 1987).

[134] But see CAL. CIV. PROC. CODE § 703.010(a) (West 1987), which provides, in part, that "[t]he exemptions provided by this chapter . . . apply to all procedures for enforcement of a money judgment." The language suggests that the lien will be subordinate to the debtor's exemption rights.

are not exempt from attachment and execution."[135] The statute thus will have no direct effect on debtors' exemption rights.

C. *Other System-Wide Impact: "Efficient" Reconsidered*[136]

1. Indirect Impact on Exemptions

As suggested earlier, some debtors retain non-exempt property simply because for the creditor, the cost of attempting to collect is too great, and the possibility of success in collection efforts is too uncertain.[137] Any realistic assessment of the debtor defenses provided by a collection system ought to take account of this shielding of debtor assets, whether or not the legislature deliberately intended such debtor protection. Such protection, the direct product of collection system inefficiencies, may well be more important to debtors than exemption statutes.[138]

While it is highly unlikely that any legislature intended to bestow debtor protection through system inefficiencies, it does not follow that legislatures enacted substantive debtor protections within the collection system without regard to system inefficiencies. For instance, system inefficiency may have affected past legislative judgment on the appropriate substance of its exemption provisions. Similarly, an inefficient collection system may have made it less necessary to develop effective procedures for asserting exemption claims. Indeed, it appears that contemporary efforts to reduce system inefficiency may raise the same political question of wealth distribution that efforts to reduce the substance of exemption provisions themselves do.

Exemption statutes were generally enacted against the backdrop of local law, including typically inefficient collection sys-

---

[135] ME. REV. STAT. ANN. tit. 14, § 4651(A) (Supp. 1986).

[136] *See generally* Balkin, *supra* note 5, at 1477–78 (assailing a narrow "efficiency" analysis conducted by some proponents of law and economics as ultimately misleading and political in content).

[137] Voluntary bankruptcy and the accompanying power to avoid a levy as a preference inject a substantial risk of failure into any execution. As discussed earlier, the new judgment lien statutes will probably create liens more likely to survive bankruptcy. If so, the bankruptcy risk within the collection system will have been lowered by the new statutes.

[138] Surely this would be the effect in a state like Pennsylvania where the exemption statutes protect only $300 of personal property plus a few odd miscellaneous items such as sewing machines. *See* 42 PA. CONS. STAT. ANN. §§ 8123–8124 (Purdon 1982).

tems.[139] In enacting exemptions, legislatures were called on to assess the needs of their debtors and their assessments of these needs played a part in molding the exemption legislation. But surely legislatures could not have made these assessments in a vacuum. Evidence of debtor needs—indeed, the motivation to consider exemption reform at all—must come primarily from debtor groups within the political process. To some extent, at least, the preexisting collection system played a part in lawmakers' perception and assessment of debtor need; an exemption package designed for a procedurally inefficient system might look very different from one designed for an efficient one. In view of the foregoing considerations, the first question a legislature might consider in connection with new judgment lien legislation is whether preexisting exemption provisions are substantively adequate in light of a more efficient collection regime.

Second, in a related vein, traditional exemption statutes may have been premised on the existence of only two distinct groups of debtors: consumer debtors who needed exemptions, and business debtors who did not need them.[140] New data suggest that we might consider the small entrepreneurs who commingle business and personal assets and finances as a third group, one that has a disproportionately high rate of bankruptcy filings.[141] A more efficient collection regime directed primarily at business debtors could exacerbate the exemption-related problems this third group might have. A policy maker might well conclude that in light of these research findings, this third group should be getting more exemption protection at the state level,[142] and that making the collection system more efficient without addressing exemption protection will simply make matters worse.[143]

---

[139] California reconsidered its exemption provisions at the time it created its new judgment lien provisions. As the text makes clear, this comprehensive approach is desirable because of the interrelated nature of the collection process and exemption protection.

[140] *See* T. SULLIVAN, E. WARREN & J. WESTBROOK, *supra* note 131, at 119–20.

[141] *See id.* at 111–12.

[142] For example, given that this type of debtor uses her business (and its assets) to produce income and sustenance, might a state want to reconsider the breadth or size of its "tools-of-the-trade" exemption so that an executing creditor cannot through execution deprive the judgment debtor of her very livelihood by seizing nearly all the business assets?

The point here is not to advocate any particular resolution of the issue but rather to suggest that exemption protection may have been developed in a faulty conceptual environment.

[143] *Cf.* T. SULLIVAN, E. WARREN & J. WESTBROOK, *supra* note 131, at 121:

A second possible effect of more efficient procedure is an increased number of debtor demands on the exemption system. Indeed, exemption provisions and procedural inefficiency work in tandem in a given system to dispense protection to debtors. Reducing the costs of formal collection processes may well increase their use.[144] If that is the case, exemption statutes will be pressed into service more often than they were before. But will a state's procedures for claiming exemption protection still be suitable in a more aggressive collection environment? Even if the exemption procedures are theoretically adequate, will the state's judicial apparatus for determining exemption-related issues be sufficient to handle an increased volume of exemption litigation? Finally, if there are extra burdens imposed on the judicial system, who will pay for them—debtors and creditors, or taxpayers? These are all questions a legislature might also consider in advance of improving collection system efficiency.

Third, an increase in the efficiency of collection will result in the decline of debtor protection which inefficient collection procedures, much like exemption provisions themselves, provide. Whether one labels it "corrective legislation" or "efficient procedure," it is likely that the reduction of collection process inefficiency results in a transfer of wealth from debtors to creditors.[145] In other words, the promotion of efficiency has distri-

---

> [I]f we really are a capitalist country, committed to the notion that people should try to start their own businesses and nurture them into Apple Computers or Tandy Electronics, just how harshly should we treat the entrepreneurs in bankruptcy? The high-risk nature of entrepreneurship means, in effect, that bankruptcy policy is another part of small business policy. It seems to us that some systematic attention to the problem of small business should be placed on the agenda of bankruptcy policymakers.

One need merely to add that the issue also should be on the agenda of any state legislature considering a more "efficient" collection system.

[144] Whitford, *supra* note 16, at 1097–98. Professor Leff suggests that in a perfect system, reducing the costs of coercive collection merely shifts the settlement value of a claim in the direction of full payment of the amount of the claim and does not necessarily increase the use of formal procedures. Leff, *supra* note 3, at 38–40. In theory, what may happen under a less costly collection system is a transition period of greater use of the formal process until information about its availability is disseminated. *See generally id.* at 38–46. Without empirical work, we cannot really know how—or for how long—less costly collection procedures will affect the rate at which creditors use formal collection measures.

[145] While this is an empirical question, it seems likely that whether or not creditors will recover a larger portion of outstanding debt under a more efficient collection system depends on (among other things) whether (1) there are available assets *not* now being collected and whether (2) creditors will exert the same collection efforts under a more efficient system in order to get larger recoveries. Creditors could also respond to a more efficient collection system by exerting *less* effort to recover the same portion of outstanding debt; they would have to respond this way if they were currently collecting all available assets from defaulting debtors.

butional consequences. Thus, efficiency is not a politically neutral, but rather a politically loaded, proposition.

2. Impact on the Informal Collection System

It is widely known that debts are generally collected without resort to the legal process. With respect to the consumer system, for instance, it was stated in 1979:

> The single most important fact about the consumer credit collection system is that, of the delinquent debts that are ultimately paid, the vast majority are collected through "consensual" debtor payments made after some kind of bargaining between creditor and debtor, and on occasion between the debtor's various creditors as well. Only a small percentage of delinquent debts are ever paid as a direct result of coercive execution.[146]

One reason for preferring the informal system is that the costs of using the formal collection system are so great.[147]

a. *Distributional Impact*

(1) *Improved settlement value of judgments.* The expense creditors must sustain in coercive collection tends to affect the settlements that creditors and debtors negotiate within the informal system. The debtor's leverage within the informal process depends, in part, on the extra expense she can force the creditor to sustain through formal means.[148] Conversely, the

---

It seems doubtful that creditors will reduce their collection efforts in response to a more efficient collection regime; the text proceeds on the more likely assumption that creditors do not currently collect all available property from defaulting debtors because of system inefficiencies and that they will continue to invest the same resources in collection activities in a more efficient system as they did before.

Of course, an economist might argue that enhancing collection efficiency will produce a net gain for the economy, which will trickle down to everyone and, as a result, is a good that all should embrace, both debtors and creditors. Moreover, an economic analysis does not purport to comment on the distributional fairness of enhanced efficiency. The point here is not to dispute the analysis but merely to observe: (1) that it is nearly certain that the creditor class will be the initial beneficiaries of the newly achieved efficiency; (2) that it seems fairly certain that the debtor class will directly finance of some gains in efficiency that the new statutes yield (assuming that creditors maintain the same level of collection activity); and (3) that it is uncertain how and whether the "newly acquired" wealth will trickle down from the creditor class to the broader population.

[146] Whitford, *supra* note 16, at 1051.

[147] *See* Whitford, *supra* note 16, at 1053–55; Leff, *supra* note 3, at 5–18.

[148] *See* Leff, *supra* note 3, at 5–10.

creditor's attraction to an offer of less-than-full payment depends in part on the added unrecoverable expenses and risk of coercive collection. If the new judgment lien provisions shift the leverage of one group or the other within the process, that shift will have economic consequences for both groups. Two examples from the earlier discussion should suffice to show that negotiating leverage will probably be changed by the new provisions.

Consider the judgment debtor engaged in ongoing secured financing in California. Before the legislature enacted the new provisions, a judgment creditor would have had to deploy execution and either attempt to levy on the collateral or search for other unencumbered assets. Absent unencumbered property, the judgment creditor's main obstacle was the presence of the secured party who would assert prior rights to the collateral. The judgment creditor's relatively high risk of getting nothing or, worse, of litigation with a secured party, gave the judgment creditor (or the creditor with a claim considering whether to get a judgment) a powerful reason to settle with the debtor for less than the claim or judgment.

Now that California's new judgment lien provisions are in effect, that creditor's need to negotiate or settle drops markedly. For the price of a simple U.C.C. filing, that judgment creditor can destroy the debtor's secured financing and perhaps put her out of business. The process is cheap and involves almost no risk to the judgment creditor. The debtor must reorient priorities in the direction of the judgment creditor, risk loss of secured financing, or enter bankruptcy to avoid the judgment lien. It is likely that in many cases the path of least resistance is to settle with the judgment creditor. The settlement value of the judgment in this situation has soared.

The same dynamics hold when the debtor in California has unencumbered personal property. Before the advent of the new provisions, the creditor's collection option was an actual levy on the targeted personal property. The debtor's response could have been to avoid that execution as a preference by filing a bankruptcy petition within ninety days of the execution.[149] The debtor's power to thwart the creditor's expensive collection efforts (and the chances that a seizure of property would prompt

---

[149] See 11 U.S.C. § 547(b) (Supp. IV 1986).

the debtor to do so) gave the creditor a large incentive to work out a consensual resolution with the debtor.

The new provisions change that equilibrium. Because fixing a judgment lien is so cheap, the creditor is unlikely even to pause in response to the threat of the debtor's bankruptcy. Furthermore, once the lien is in place, the debtor does not have the same incentive to avoid it, because the lien has not deprived the debtor of possession. The threat of a bankruptcy filing does not carry the same leverage value for the debtor, because the creditor will not have much of a net loss (only filing fees and minimal effort) if the debtor files. Yet if the debtor does not file a petition within 90 days of the fixing of the lien, the lien will be unavoidable as a preference.[150] The leverage value bankruptcy has for debtors both before and after the fixing of a lien on personal property has been substantially reduced with these new provisions. Once again, the settlement value of the judgment has risen.

(2) *Distributional implications of more valuable judgments.* So what's the problem, the reader may well ask. Is not easier collection the underlying rationale for the provisions? Is not the whole point to redistribute wealth from nonpaying debtors to judgment creditors who fought hard for their judgments? Are we not promoting with these provisions a central value in the law, that judgments should be paid? Will not the greater collection system efficiency that comes with these statutes benefit everyone? Questions like these cut to the heart of the problems that come with oversimplified analysis of innovation in the debtor-creditor field. Simply understanding that there can be different answers to such questions is central in assessing the merit of the new provisions.

If the normative proposition is that these provisions that improve the settlement value of judgments are good simply *because* judgments should be paid, we are implicitly advancing the premise that full, complete payment of judgments is an unqualified good the law should seek to attain. But such a premise flies in the face of several hundred years of legal history: the law clearly does not do all it could do to secure the payment of its judgments. It does not execute defaulting judgment debtors; indeed, the legal system seldom imposes criminal sanctions

[150] To the extent that the lien might reach exempt property, the lien is probably avoidable in bankruptcy under 11 U.S.C. § 522(f) (Supp. IV 1986).

on debtors,[151] and offers debtors the escape hatch of voluntary bankruptcy. "Judgments should be paid" sounds absolute and is difficult to quarrel with in the abstract; however, when we consider the lengths to which the law actually goes to advance the proposition, we find it is merely one of several competing values in this field that the law might advance. If the law had not qualified this proposition with bankruptcy, exemptions, limits on the collection process, and system inefficiency, we would have had to develop a different lexicon to express the complexity of our normative conclusions on paying judgments.[152]

If "judgments should be paid" does not supply a strong normative grounding, perhaps "efficiency" will. If some wasted energy from the debtor-creditor system can be eliminated, all participants in that system might be the beneficiaries and those resources might be put to better economic use. The new provisions will probably increase the settlement value of judgments because they are cheaper to deploy. Is not everyone in business a creditor and a debtor at various times and is not everyone thus going to benefit from improved efficiency in the system?

Some important data bear on these points. Recent field studies suggest that there may be a definable debtor group and a definable creditor group. The work by Professors Sullivan, Warren, and Westbrook mentioned earlier shows that small business debtors account for a disproportionately high percentage of defaults resulting in business bankruptcy.[153] Small businesspeople default more often than large businesspeople, and one might therefore expect them to be on the debtor side more often than they will be on the creditor side. If this is actually the case, then to the extent that the new provisions increase the settlement value of a judgment, there will probably be a redistribution of wealth from small business debtors to their creditors.[154] The

---

[151] There are some circumstances in which jail may follow nonpayment of judgments. *See* Note, *Body Attachment and Body Execution: Forgotten But Not Gone*, 17 WM. & MARY L. REV. 543 (1976); Note, *Civil Arrest of Fraudulent Debtors: Toward Limiting the Capias Process*, 26 RUTGERS L. REV. 853 (1973). *See also* Moses, *Enforcement of Judgments Against Hidden Assets*, 1951 U. ILL. L.F. 73.

[152] Professor Warren makes this same point in connection with the lexicon of contract doctrine. Warren, *supra* note 126, at 779.

[153] In their study, the authors found that debtors with small businesses accounted for 10.4% of the bankruptcies but for only 7.3% of the general population. When they added debtors who formerly had small businesses, the total accounted for 20% of the total bankruptcy filings. T. SULLIVAN, E. WARREN & J. WESTBROOK, *supra* note 131, at 205.

[154] Creditors *could* respond to the new provisions by reducing their collection efforts to net the same proceeds, thereby realizing a cost savings without improving the settlement value of judgments. It seems unlikely that creditors would respond in this way to the new provisions and the text proceeds on the assumption that they will not. The question is ultimately an empirical one and a definitive answer awaits the evidence.

point that a wealth redistribution will come with these provisions was made earlier in connection with exemption protection but the political content of the provisions is worth emphasizing again in the business context.

Yet it remains difficult to quarrel with legislation that will reduce legal waste, even if the immediate distributional effects are politically delicate. After all, one might argue, reduced collection costs will ultimately result in cheaper credit for business debtors and a bounty of sorts for the broader economy. Once again, however, one must be sensitive to other parts of the system to see if there may be undesirable side effects of improved ability to collect judgments.

One potential side effect is the impact the provisions might have on the personal finances of business debtors. The finding that small business debtors commingle their personal and business finances suggests, once again, that exemptions may take on enhanced importance with the new statutes. As the settlement value of a judgment increases with the new provisions, judgment debtors will need to divert resources previously allocated elsewhere to settling accounts with judgment creditors. Diversion of resources from family support and nutrition, for example, might be more likely to occur under the new provisions[155] and might ultimately cost the broader economic system more than the new provisions save.[156]

b. *Implications of More Valuable Judgments for Creditors without Judgments*

In traditional jurisdictions, both ordinary creditors and judgment creditors can lay claim to the debtor's personal property only at considerable expense. The judgment creditor must de-

[155] *Cf.* T. SULLIVAN, E. WARREN & J. WESTBROOK, *supra* note 131, at 118–19.

[156] One can imagine many ways these statutes could wind up costing more than they save. Suppose the increased leverage supplied by these statutes and the tenacity of small business debtors to hang on resulted in nutritional or shelter deficiencies for the debtor and her family that the state ultimately had to remedy at high expense. Or suppose the provisions turned out to hasten the financial demise of small businesses which, in turn, resulted in lost jobs and unnecessary economic costs like moving expenses, job search fees, and unemployment. It is easy to imagine these latter items costing the overall economy more than the resources that are saved by the new statutes.

The counterpoint, of course, is that some debtors hang on longer than they "should" and that it is indeed better that losing enterprises fail sooner rather than later. We do not currently know where the optimal "failure point" is. The point here is that one's opinions on the subject no doubt depend in part on the number of potential effects one considers in assessing the facts.

ploy execution procedures to get a lien; the person with a mere claim must first get a judgment and then execute on it. Partly because of expense and partly because the debtor can undo execution with a bankruptcy petition, in many cases neither ordinary creditor nor judgment creditor has considerable leverage to force payment when the debtor has no real estate. In those situations, extra-legal leverage may well play a larger role than legal leverage in the debtor's decision to pay one creditor before the next. Debtors may base priority in paying creditors on their need for continued service or financing, for example, rather than on the nature of the creditor's claim.

In judgment lien reform jurisdictions like California, the additional legal leverage held by a creditor with a new judgment lien may change the way the debtor allocates her inadequate resources. That is precisely what makes these reform statutes attractive in the first place: debtors will begin paying judgment creditors sooner than they would otherwise—before they pay others. Yet what impact will such a reordering of priorities have on the debtor's economic survival? Might the reordering squeeze the debtor in a way that accelerates financial demise? As suggested earlier, unsecured creditors in these jurisdictions seem less likely to benefit as much from bankruptcy, including reorganization under Chapter 11, as they might have under more traditional systems. They might therefore be less tolerant of late payments and less flexible in working through difficult periods with the debtor. The resulting increased pressure on the debtor combined with a less viable bankruptcy process could yield an unpredictable economic impact. Legislatures ought to consider carefully whether the new statutes will lessen the chances for economic survival of shaky businesses and, if so, whether that is desirable as a policy matter.

The improved position of the judgment creditor in relation to other creditors in the new systems raises yet another potentially undesirable consequence: claimants' increased use of formal judicial procedures to collect their debts. As developed earlier, in most jurisdictions, creditors with judgments have substantial disincentives to use the judicial process to enforce their judgments. The judicial process is expensive, the results are uncertain, and the debtor can file a bankruptcy petition and render the efforts worthless in any event. The enforcement problems no doubt work their way backwards in many cases to the point where a claimant decides whether or not to bother to get a

judgment. It makes little sense under the present systems to begin a legal action in the first place if it will ultimately yield little or no money. Additionally, in the old systems, claimants and judgment holders are both unsecured creditors without claims to specific personal property. In a case of difficult enforceability, a competing claimant need not worry much about the judgment creditor. Indeed, if the judgment creditor attempted enforcement, the claimant without a judgment might file an involuntary bankruptcy petition[157] and thereby nullify the short-lived competitive advantage.

As suggested earlier, one expects that most creditors with judgments will get liens on personalty under the new systems, because doing so is so inexpensive. Yet the very fact that those with judgments will get liens as a matter of course upsets a kind of equilibrium formerly held by ordinary claimants and judgment creditors without liens. A creditor with a mere claim might have to be more legally competitive in the new jurisdictions because if a second creditor were to get a judgment first, that second creditor could easily get a decisive competitive advantage in the form of an earlier lien. Put another way, because judgments have become more valuable both as against debtors and as against other creditors, one expects the inter-creditor competition for judgments to increase.

If the new statutes will prompt an increased use of the judicial system in debt collection, the value of the provisions comes into serious question at two levels. The first is a question of redistribution of debtor assets, this time among competing claimants. The distributional problem is suggested by the observation that some creditors are able to obtain judgments more easily than others and that the ease in getting judgments seems to have little to do with the nature of the underlying claim. For example, suppose the debtor assaulted claimant one, an individual, and failed to pay a loan installment to claimant two, a finance company. Which of the two seems more likely to get the first judgment and lien? Do we want a system that, in fact, will create a high priority in relation to other creditors for those that, because of the strength of their claims or the size of their legal staffs, can most easily get judgments? As suggested earlier, priority statutes typically award liens to creditors on the basis of a policy judgment that the claimant is somehow deserving.

---

[157] *See* 11 U.S.C. § 303 (1984 and Supp. IV 1986).

The statutes under consideration here have made that kind of distinction between those that get judgments and all others. It seems very naive indeed to assume that all claimants have equal access to judgments and that the most deserving of special priority will be the ones who will obtain them.

The second level of questions goes to the tendency these statutes may have to encourage formal, over informal, debt collection itself. Under the traditional systems, a creditor, with or without a judgment, can get a non-possessory lien on the debtor's personal property by acquiring a consensual Article 9 security interest. To obtain that under the traditional procedures, the debtor and creditor negotiate informally over debt collection. To what extent will the incentives to make use of informal measures decrease under the new provisions? To what extent do the incentives actually encourage resort to the legal system as a first, rather than last, move? Would we want to move away from the present system in which nearly all debts are settled informally and consensually? If use of formal processes increases, would the filing fees associated with increased use of formal processes fully cover the legal system's expanded costs?[158] Might not the total system losses, given enhanced creditor competitiveness, exceed those under more traditional systems?

Indeed, can one even consider the new statutes "efficient" when viewed in the context of the larger system? A seldom used judicial system combined with a heavily used informal collection system (the old system) may well be cheaper and more "efficient" than a frequently used judicial system combined with less reliance on the informal system. Once other values, such as consensual dispute resolution as the preferred approach to debt collection and the policy proposition that the most deserving claimants ought to be the first paid from limited assets, are taken into account, the conclusions that these statutes will yield a bounty in saved costs are open to serious question.

---

[158] None of the costs of informal dispute resolution are directly imposed on citizens as taxes. If the filing and associated fees do not cover the system costs within the formal system, an increased use of that system (and decreased use of the informal system) would redistribute some dispute resolution costs now borne by creditors (and their customers) to taxpayers. While one could develop a policy argument favoring a redistribution of costs from creditors and their customers to taxpayers, the point here is that a legislature should not implicitly decide to redistribute those costs without considering the policy implications.

V. Conclusion

Those who practice or preach state collection law routinely condemn state collection statutes as cumbersome, expensive, and inefficient. As with modern architecture, in judgment collection, lean and trim—efficient—is often thought of as inherently better. The new statutes examined here offer possibly drastic reductions in the inefficiency of the state collection machinery, and one's first instinct is to applaud and embrace the new legislation. The impulse is to see the new provisions as efficient, cost- saving, and innovative legislation that is desirable within the debt collection system.

Yet, further examination of the statutes, particularly within the larger context of the debtor-creditor system, produces ambivalence. It seems likely that the new statutes will have an impact on the bankruptcy process as well as secured lending under Article 9 of the U.C.C., that they will redistribute wealth between debtor and creditor classes in the same way that changing exemption laws redistributes wealth, that they will set priorities among creditors in ways that might not be desirable, that they may have a substantial impact on the informal collection process, that they may shift some costs of dispute resolution from debtors and creditors to taxpayers, and that they may even provoke lawsuits by a debtor's claimants who are fearful that others will use these new provisions first. Indeed, when viewed in this larger context, the statutes have remarkable political implications, and there are serious questions about whether they will produce the increased efficiency we first imagine or whether, in fact, total system losses will be greater.

Even if one believes that "efficient" is "better," that a clean facade is inherently superior to a decorated one, deciding whether a proposal is "efficient" is enormously more complicated than deciding whether a building is in the Second Empire or International Style. In the debtor-creditor field, one cannot simply look to outward appearances: determining whether a proposal is "efficient" requires more than simply seeing whether a procedural innovation will reduce the immediate legal costs.

A review of judgment lien reform statutes illustrates major problems with easy conclusions about legal efficiency based on too narrow a view of impact and too rigid an idea of those things to which we might attach value. It also demonstrates the central need in this area of the law for extensive field work to assess

the actual effects of legal change. These statutes have numerous potential side effects that are not readily apparent and which, if they occur, may not be desirable. The statutes' impact requires close monitoring by both federal and state policy makers. Until we have a better idea of their actual impact, legislatures would be wise to move slowly in embracing these statutes.

ARTICLE

UNION SECURITY AGREEMENTS UNDER THE NATIONAL LABOR RELATIONS ACT: THE STATUTE, THE CONSTITUTION, AND THE COURT'S OPINION IN *BECK*

KENNETH G. DAU-SCHMIDT*

The Supreme Court's recent decision in Communications Workers of America v. Beck *interpreted section 8(a)(3) of the National Labor Relations Act (NLRA) to prohibit the observance of agency shop agreements. By interpreting the statute in this way, the Court avoided the question of whether union security agreements under the NLRA are subject to constitutional scrutiny. The Court's determination that section 8(a)(3) does not allow agency shop agreements was an important decision affecting the enforceability of union security agreements in the vast majority of private sector bargaining agreements.*

In this Article, Professor Dau-Schmidt criticizes the Court's interpretation of section 8(a)(3) in Beck. *The Article examines the Court's NLRA precedents and the legislative history of the NLRA's section 8(a)(3). Several methods of statutory construction and constitutional adjudication are analyzed. Finally, Professor Dau-Schmidt argues that union security agreements under the NLRA are not subject to constitutional scrutiny because there is no state action in their negotiation or observance.*

In its recent opinion, *Communications Workers of America v. Beck*,[1] the Supreme Court announced the governing principle for determining the extent to which union security agreements[2] may be observed under the National Labor Relations Act (NLRA).[3] In *Beck*, the Court faced the questions of whether

* Associate Professor of Law and Economics at the University of Cincinnati College of Law; A.B., University of Wisconsin, 1978; M.A., University of Michigan, 1981; J.D., University of Michigan Law School, 1981; Ph.D. in economics, University of Michigan, 1984. Professor Dau-Schmidt wishes to thank Joseph Biancalana, Jean Braucher, Gordon Christenson, Bob Martineau, Mike Nicholson, Henry H. Perritt, Jr., Ronna Schneider, Ted St. Antoine, Joseph Tomain, and all the members of University of Cincinnati College of Law's Faculty Seminar for helpful comments. Professor Dau-Schmidt extends special thanks to Beverly Moran for selflessly reading and commenting on three different drafts of this Article, John Applegate for particularly useful comments on the structure of the Article, and Michael Morley for very able research assistance in the preparation of this Article.

[1] ——— U.S. ———, 108 S. Ct. 2641 (1988).

[2] A "union security agreement" is an agreement between a union and an employer that the employer will require all employees to undertake a specified level of support for the union as a condition of employment. R. GORMAN, LABOR LAW 639 (1976). *See infra* notes 32–36 and accompanying text.

[3] 29 U.S.C. §§ 141–87 (1982).

section 8(a)(3) of the NLRA[4] allowed a union and an employer to negotiate and observe an agency shop agreement,[5] and if so, whether such agreements violated dissenting employees' first amendment rights.[6] In answering these questions, the Court found that Congress's sole purpose in allowing union security agreements under section 8(a)(3) was to ensure that workers who shared in the benefits of collective bargaining also shared in the costs of that bargaining.[7] Based on this perception of congressional purpose, the Court concluded that section 8(a)(3) allows unions to compel contributions from dissenting employees only to the extent that the union expenditures are "necessarily or reasonably incurred for the purpose of performing the duties of an exclusive [bargaining] representative."[8] The Court's resolution of the statutory issue obviated the need to answer the constitutional question.

The Court's opinion in *Beck* has potentially far-reaching implications. The NLRA is the nation's basic labor statute cov-

---

[4] The pertinent portions of § 8(a)(3) read as follows:

> (a) It shall be an unfair labor practice for an employer—. . . (3) by discrimination in regard to hire or tenure of employment or any term or condition of employment to encourage or discourage membership in any labor organization: *Provided*, That nothing in [the NLRA], or in any other statute of the United States, shall preclude an employer from making an agreement with a labor organization (not established, maintained, or assisted by any action defined in this subsection as an unfair labor practice) to require as a condition of employment membership therein on or after the thirtieth day following the beginning of such employment or the effective date of such agreement, whichever is later, (i) if such labor organization is the representative of the employees as provided in section [9(a) of the NLRA], in the appropriate collective-bargaining unit covered by such agreement when made, and (ii) unless following an election held as provided in section [9(e) of the NLRA] within one year preceding the effective date of such agreement, the Board shall have certified that at least a majority of the employees eligible to vote in such election have voted to rescind the authority of such labor organization to make such an agreement: *Provided further*, That no employer shall justify any discrimination against an employee for nonmembership in a labor organization (A) if he has reasonable grounds for believing that such membership was not available to the employee on the same terms and conditions generally applicable to other members, or (B) if he has reasonable grounds for believing that membership was denied or terminated for reasons other than the failure of the employee to tender the periodic dues and the initiation fees uniformly required as a condition of acquiring or retaining membership[.]

29 U.S.C. § 158(a)(3) (1982).

[5] An "agency shop" is a form of union security agreement in which the employer agrees to require all employees to begin making agency fee payments to the union within a specified period of time after accepting employment, and to continue such payments for the term of their employment. Traditionally the agency fees required under an agency shop agreement are equal to union dues. R. GORMAN, *supra* note 2, at 642.

[6] *Beck*, 108 S. Ct. at 2645.

[7] *Id.* at 2650 (citing Radio Officers v. NLRB, 347 U.S. 17, 41, 74 (1974)).

[8] *Id.* at 2652 (citing Ellis v. Brotherhood of Ry. Clerks, 466 U.S. 435, 447–48 (1984)).

ering virtually the entire private sector. Although section 14(b) of the NLRA allows states to proscribe union security agreements,[9] thirty states, comprising two-thirds of the United States's population, have elected not to do so.[10] In those states over ninety percent of collective bargaining agreements,[11] covering over six million workers, include union security agreements.[12] The Court's decision in *Beck* limits the enforceability of all of these agreements. Although the number of employees who will dissent from full agency shop payments will probably be small,[13] and the reduction in fees that will accompany such dissension will be small for most unions,[14] the bookkeeping and litigation expenses required to resolve the complaints of dissenters may prove a significant drain on the resources of the American labor movement.[15]

The Court's opinion in *Beck* also promises to be controversial. The Court failed to interpret section 8(a)(3) by direct examination of the statute's words, administrative interpretations, or legislative history. Instead, the majority announced an identity

---

[9] Section 14(b) of the NLRA provides that "[n]othing in [the NLRA] shall be construed as authorizing the execution or application of agreements requiring membership in a labor organization as a condition of employment in any State or Territory in which such execution or application is prohibited by State or Territorial law." 29 U.S.C. § 164(b) (1982).

[10] AFL-CIO, ECONOMIC COMPARISONS BETWEEN OPEN SHOP STATES AND FREE COLLECTIVE BARGAINING STATES 1 (1986).

[11] 2 Collective Bargaining Negot. & Cont. (BNA) 87:1 (July 1986).

[12] BUREAU OF LABOR STATISTICS, MAJOR COLLECTIVE BARGAINING SETTLEMENTS IN PRIVATE INDUSTRY, 1988 1 (Jan. 1989). By comparison, the Supreme Court's opinion in International Ass'n of Machinists v. Street, 367 U.S. 740 (1961), holding that agency shop agreements could not be fully observed under the Railway Labor Act, affected only about one million workers, many of whom were not unionized. *See* BUREAU OF LABOR STATISTICS, EMPLOYMENT AND EARNINGS 17 (Jan. 1962). This argument first appeared approximately in this form in the petition for writ of certiorari in *Beck*.

[13] If the experience of private sector unions with voluntary programs to refund the portion of union dues spent on political expenditures is any indication, the number of dissenting employees may be very small indeed. Such a refund program has been in effect for several years at United Auto Workers, and in a typical year dissenting duespayers number only about 100 (approximately 0.01% of their membership). B. TAYLOR & F. WITNEY, LABOR RELATIONS LAW 389 (5th ed. 1987).

[14] Under the *Ellis* formula made applicable to NLRA unions by *Beck*, *see infra* note 69 and accompanying text, the largest expenditures that a union cannot charge dissenting employees are expenditures for organizing and political activities. These expenditures have been estimated at 15% and 5% respectively for the average private sector union. Henkel & Wood, *Limitations on the Uses of Union Shop Funds After* Ellis*: What Activities are "Germane" to Collective Bargaining?* 35 LAB. L.J. 736 (1984); B. TAYLOR & F. WITNEY, *supra* note 13, at 388–89.

[15] *See* Brief of the AFL-CIO at 25, *Beck*, 108 S. Ct. 2641. The *Beck* decision also provides an opportunity for employers to harass the trade union movement through employer organizations ostensibly organized to protect the interests of individual employees. *See infra* note 55.

between section 8(a)(3) of the NLRA and section 2 Eleventh of the Railway Labor Act (RLA), and found that its prior opinion, that section 2 Eleventh allowed the observance of union security agreements only for the recoupment of collective bargaining expenses, was controlling in its interpretation of section 8(a)(3).[16] The Court then battled mightily to make the legislative history of the NLRA fit this limited interpretation of section 8(a)(3).[17] The Court's task was not eased by its earlier failure, when interpreting section 2 Eleventh, to consider the legislative history of section 8(a)(3).[18] This omission is inconsistent with the Court's new-found identity between the two sections, since section 8(a)(3) was passed four years before the passage of section 2 Eleventh. Finally, it should be noted that the Court's earlier interpretation of section 2 Eleventh has itself been criticized as inconsistent with the language of the RLA[19] and affected by the Court's desire to avoid the question of whether agency shop agreements under the RLA violate dissenting employees' first amendment rights.[20]

In this Article, I argue that the Court's interpretation of section 8(a)(3) of the NLRA in *Beck* cannot be supported by direct examination of the statute's words, administrative interpretations, or legislative history. Such a direct examination of the usual intrinsic and extrinsic evidence for statutory interpretation suggests that in enacting section 8(a)(3), Congress intended to allow the full observance of agency shop agreements and rec-

---

[16] *Beck*, 108 S. Ct. at 2648–49. The relevant language of section 2 Eleventh of the RLA reads as follows:

> Notwithstanding any other provisions of [the RLA], or of any other statute or law of the United States, or Territory thereof, or of any State, any carrier or carriers as defined in this Act and a labor organization or labor organizations duly designated and authorized to represent employees in accordance with the requirements of this chapter shall be permitted—a) to make agreements, requiring, as a condition of continued employment, that . . . all employees shall become members of the labor organization representing their craft or class: *Provided*, That no such agreement shall require such condition of employment with respect to employees to whom membership is not available upon the same terms and conditions as are generally applicable to any other member or with respect to employees to whom membership was denied or terminated for any reason other than the failure of the employee to tender the periodic dues, initiation fees, and assessments (not including fines and penalties) uniformly required as a condition of acquiring or retaining membership . . . .

45 U.S.C. § 152 Eleventh (1982).

[17] *Beck*, 108 S. Ct. at 2649–57.

[18] *See* International Ass'n of Machinists v. Street, 367 U.S. 740 (1961).

[19] *See* Ellis v. Brotherhood of Ry. Clerks, 466 U.S. 435, 445–46 (1984).

[20] *See id.*; NLRB v. Catholic Bishop of Chicago, 440 U.S. 490, 500 (1979); Abood v. Detroit Bd. of Educ., 431 U.S. 209, 232 (1977); *Street*, 367 U.S. at 749; Cantor, *Uses and Abuses of the Agency Shop*, 59 NOTRE DAME L. REV. 61, 65–72 (1983).

ognized purposes for union security agreements beyond the mere recoupment of collective bargaining expenses. The Court's interpretation of section 8(a)(3) deviates from Congress's intent because it relies on the Court's prior interpretation of section 2 Eleventh of the RLA and on that interpretation's constitutionally colored view of the purpose and extent of union security agreements allowed under the RLA.[21] The Court's reliance on its prior interpretation of section 2 Eleventh in interpreting section 8(a)(3) amounts to an application of the doctrine of avoiding constitutional questions.[22] Because the *Beck* Court failed to acknowledge or consider the effect of its past constitutional concerns on its interpretation of section 2 Eleventh,[23] it never considered whether imposing this interpretation on section 8(a)(3) was an appropriate application of this doctrine.

I also argue that the Court's interpretation of section 8(a)(3) in *Beck* is an inappropriate application of the doctrine of avoiding constitutional questions. The negotiation and observance of agency shop agreements under the NLRA raises no serious constitutional question to avoid. Under the Court's recent precedents, there is no plausible argument that the actions of a union and a private employer in negotiating and observing such agreements represent state action subject to constitutional scrutiny. Moreover, the Court's interpretation in *Beck* is not supported by the rationales typically given for the doctrine of avoiding constitutional questions. One rationale is that the Court should presume that Congress avoids constitutional controversy in its enactments.[24] The words and legislative history of the NLRA suggest that, in enacting that statute, Congress consciously raised a host of constitutional questions including whether union security agreements infringe dissenting employees' first amendment rights. Another rationale is that the Court should avoid constitutional questions to minimize its encroachment on the powers of the elected legislature.[25] The Court's interpretation of section 8(a)(3) actively encroaches on the domain of the

---

[21] *See Beck*, 108 S. Ct. at 2648 (citing *Street*, 367 U.S. 740).

[22] The doctrine holds that where a "serious question" of a statute's constitutionality has been raised, the Court should, if possible, "fairly" construe the statute to avoid the constitutional question. Crowell v. Benson, 285 U.S. 22, 62 (1935).

[23] *Beck*, 108 S. Ct. at 2657.

[24] W. ESKRIDGE & P. FRICKEY, LEGISLATION, STATUTES, AND THE CREATION OF PUBLIC POLICY 676 (1988).

[25] Ashwander v. TVA, 297 U.S. 288, 346–48 (1935) (Brandeis, J., concurring); Posner, *Statutory Interpretation—In the Classroom and in the Courtroom*, 50 U. CHI. L. REV. 800, 815 (1983).

legislature by specifying both the policy and the form of the statute. Both rationales suggest that, even where application of the doctrine is appropriate, the Court should undertake the minimum deviation from legislative intent necessary to avoid the constitutional question. However, the Court's opinion in *Beck* ignores an alternative interpretation of section 9(a) of the NLRA[26] that would allow the negotiation of agency shop agreements, but would avoid constitutional objections by removing the negotiation of employee fees for non-collective bargaining expenses from the exclusive province of the union.[27]

Finally, I examine the question, avoided in *Beck*, of whether the negotiation and observance of union security agreements under section 8(a)(3) are subject to constitutional constraints. This question is of continuing importance, since the Court will probably soon be asked to decide whether the procedural protections of dissenting employees' rights it has found constitutionally required in the public sector[28] apply to employees governed by the NLRA. I find that under the current precedents of the Court there is insufficient state action in the negotiation and observance of union security agreements under the NLRA to support constitutional objections. There is insufficient grant of authority in the designation of the union as the exclusive bargaining representative to make the union a state actor. Moreover, it is a fundamental tenet of American labor law that the government regulates only the process of collective bargaining, leaving the determination of the actual provisions of collective agreement to the parties.[29] Thus, there is insufficient state en-

---

[26] 29 U.S.C. § 159(a) (1982).

[27] As discussed below, the designation of the union as the exclusive representative of the employees under section 9(a) is the primary argument for state action in the negotiation and observance of union security agreements. Such state action is a prerequisite for dissenting employees' constitutional objections to agency fees for expenses unrelated to collective bargaining. *See infra* text accompanying notes 310–313, 373–380.

[28] Chicago Teachers Union, Local No. 1 v. Hudson, 475 U.S. 292 (1986).

[29] One of [the] fundamental policies [of the NLRA] is freedom of contract. While the parties' freedom . . . is not absolute under the Act, allowing the Board to compel agreement when the parties themselves are unable to agree would violate the fundamental premise on which the Act is based—private bargaining under governmental supervision of the procedure alone, without any official compulsion over the actual terms of the contract.

H.K. Porter Co. v. NLRB, 397 U.S. 99, 108 (1970) (citations omitted). "[T]he Wagner Act became law on the floodtide of the belief that the conflicting interests of management and worker can be adjusted only by private negotiation, backed, if necessary, by economic weapons, without the intervention of law." Cox, *The Right To Engage In Concerted Activities*, 26 IND. L.J. 319, 322 (1951). "The basic philosophy of the Wagner Act was that labor problems would be resolved by a private process of negotiation and contracting, backed up by the threat or use of self-help measures to secure bargaining

couragement of the negotiation of union security agreements to designate observance of such agreements as state action. The fact that the government, in hopes of achieving certain public policies, has fostered a particular system of private negotiation—collective bargaining through elected exclusive representatives—does not change the private nature of these negotiations or their results. The government commonly fosters certain private institutions, for example private corporations, in hopes that these institutions will benefit their participants. To extend constitutional scrutiny to the negotiation and observance of union security agreements made by private unions and private employers under the NLRA would undermine the individual liberty the state action requirement is designed to protect[30] and extend judicial power beyond the bounds circumscribed for it in our Constitution.[31]

I. *Communications Workers of America v. Beck*: A Lack of Union Security in Statutory Rights

A. *The Prelude to* Beck

A union security agreement is an agreement between a union and an employer that the employer will require all employees to undertake some specified level of union support as a condition of employment.[32] Such agreements can take several forms. If the union and the employer negotiate a "closed shop" agreement, the employer agrees to require union membership and the payment of union dues as a condition of both gaining and retaining employment. Under such an agreement, the employer may hire only union members and must require all employees to maintain their union membership and pay union dues.[33] Under

---

advantage." Klare, *The Public-Private Distinction in Labor Law*, 130 U. Pa. L. Rev. 1358, 1390 (1982).

[30] [I]f the Constitution is used to restrict private conduct, its role will be transformed. Rather than retaining its position as the protector of liberty, it will become for many, if not all, a vehicle of regulation and annoyance as the populace is forced continually to look over its collective shoulder in fear that its actions might be in contravention of the judiciary's demarcation of another's constitutional rights.

Marshall, *Diluting Constitutional Rights: Rethinking "Rethinking State Action"*, 80 Nw. U.L. Rev. 558, 569–70 (1985).

[31] L. Tribe, American Constitutional Law § 18-2 (2d ed. 1988).

[32] R. Gorman, *supra* note 2, at 639–41.

[33] *Id.* at 641–42.

a "union shop" agreement, the employer may hire without regard to union membership, but agrees to require all employees to join the union and begin paying union dues within a specified period of time after being hired.[34] An "agency shop" agreement is similar to a union shop agreement, except that the employees are only required to pay dues to the union. Whether the employees actually join the union is left to their discretion.[35] Finally, if the union and employer agree to a "maintenance of membership" agreement, the employer agrees that, if any employees join the union during the term of their employment, the employer will require that they continue their membership and pay dues as a condition of employment.[36] As discussed below, the closed shop is now largely of only historical significance.[37]

For the vast majority of employees in the private sector, the legality of union security agreements is governed by section 8(a)(3) of the NLRA and by state law.[38] Section 8(a)(3) was first enacted in 1935 as section 8(3) of the Wagner Act.[39] At that time, Congress enacted the general prohibition of section 8(a)(3), which forbids employers from "encourag[ing] or discourag[ing] membership in any labor organization" "by discrimination in regard to hire or tenure of employment or any term or condition of employment."[40] By itself, this general prohibition would forbid union security agreements. However, to preserve the negotiation and observance of union security agreements, Congress added the following proviso:

> *Provided,* That nothing in this subchapter, or in any other statute of the United States, shall preclude an employer from making an agreement with a labor organization . . . to require as a condition of employment membership therein . . . if such labor organization is the representative of the employ-

---

[34] *Id.*

[35] *Id.*

[36] *Id.*

[37] In the Taft-Hartley Act, Congress outlawed the closed shop. *See infra* notes 46, 243 and accompanying text. The Supreme Court has also interpreted the second proviso of section 8(a)(3) to prohibit enforcement or observance of the membership requirement of a union shop agreement. NLRB v. General Motors Corp., 373 U.S. 734 (1963). *See infra* notes 154–156 and accompanying text.

[38] Section 8(a)(3) allows certain forms of union security agreements for employees governed by the NLRA. 29 U.S.C. § 158(a)(3) (1982). However, section 14(b) allows state laws prohibiting union security agreements to supersede section 8(a)(3)'s authorization. 29 U.S.C. § 164(b) (1982). For the text of NLRA section 14(b), *see supra* note 9.

[39] National Labor Relations Act § 8(3), ch. 372, Pub. L. No. 198, 49 Stat. 449 (1935) (current version at 29 U.S.C. §§ 151–69 (1982)).

[40] 29 U.S.C. § 158(a)(3) (1982).

> ees as provided in section 159(a) of this title, in the appropriate collective-bargaining unit covered by such agreement when made, ...[41]

Section 159(a) referred to in the proviso is now section 9(a) of the NLRA. It specifies that a union elected by a majority of the employees is the "exclusive representative" of those employees with the right and obligation to represent all of the employees fairly, including non-members, in negotiating and enforcing a collective agreement.[42]

In the Taft-Hartley Act of 1947,[43] Congress relettered section 8(3) as section 8(a)(3)[44] and placed additional limitations on union security agreements. Congress limited the existing proviso by specifying that union security agreements could not require union membership as a condition of employment until "on or after the thirtieth day following the beginning of such employment or the effective date of such agreement, whichever is the later . . . ."[45] By limiting union membership to a post-condition of employment, Congress outlawed the closed shop.[46] Congress imposed further limitations by adding a second proviso to section 8(a)(3):

> *Provided further*, That no employer shall justify any discrimination against an employee for nonmembership in a labor organization (A) if he has reasonable grounds for believing that such membership was not available to the employee on the same terms and conditions generally applicable to other members, or (B) if he has reasonable grounds for believing that membership was denied or terminated for reasons other than the failure of the employee to tender the periodic dues and the initiation fees uniformly required as a condition of acquiring or retaining membership . . . .[47]

---

[41] *Id.*

[42] "Representatives designated or selected for the purposes of collective bargaining by the majority of employees in a unit . . . shall be the exclusive representatives of all the employees in such unit . . . ." 29 U.S.C. § 159(a) (1982); *see also* Steele v. Louisville & N. R.R., 323 U.S. 192 (1944) (union, as exclusive representative, has duty to represent all employees fairly in negotiation and enforcement of the collective agreement); Ford Motor Co. v. Huffman, 345 U.S. 330 (1953) (precedent of *Steele* under RLA applied to section 9(a) of the NLRA).

[43] Labor Management Relations Act, ch. 120, § 101, 61 Stat. 136, 140 (1947) (amending 29 U.S.C. § 158(3) (1946), now codified at 29 U.S.C. § 158(a)(3) (1982)).

[44] Congress relettered section 8(3) because it created a new subsection, 8(b), specifying prohibited or "unfair" labor practices for unions. 29 U.S.C. § 158(b) (1982).

[45] 29 U.S.C. § 158(a)(3) (1982).

[46] *Beck*, 108 S. Ct. at 2649–50.

[47] 29 U.S.C. § 158(a)(3) (1982).

This proviso seeks to limit union discretion under union security agreements to the enforcement of the obligation to pay dues. Finally, Congress enacted sections 8(b)(2) and 14(b) of the NLRA. Section 8(b)(2) prohibits a union from causing or attempting to cause an employer to discriminate against an employee in violation of section 8(a)(3).[48]Section 14(b) states that nothing in the NLRA, including section 8(a)(3), pre-empts state laws prohibiting union security agreements.[49] Section 14(b) was included merely to clarify the pre-emption question since, even under the Wagner Act, the pre-emption of state laws prohibiting union security agreements was not intended.[50]

For employees in the railroad and airline industries, the legality of union security agreements is governed by section 2 Eleventh of the RLA.[51] Under this Act, union security agreements were unlawful until 1951.[52] At that time Congress amended section 2 Eleventh so that it is now very similar to section 8(a)(3) of the NLRA, except that section 2 Eleventh explicitly pre-empts contrary state laws.[53] Indeed, the legislative history of the 1951 amendments suggests that Congress intended to allow the same sort of union security agreements under section 2 Eleventh as it had permitted under section 8(a)(3) of the NLRA.[54]

The statutory constructions of both section 8(a)(3) and section 2 Eleventh sparked disputes between nonunion employees and their employers. The resentment of these dissenting employees over mandatory payments to support unions they opposed, and the ready access to the litigation resources of employer-dominated "right-to-work" organizations,[55] inevitably led to litigation to determine the extent and constitutionality of union security

---

[48] 29 U.S.C. § 158(b)(2) (1982).
[49] 29 U.S.C. § 164(b) (1982).
[50] *See infra* note 217.
[51] 45 U.S.C. § 152 Eleventh (1982).
[52] *Beck*, 108 S. Ct. at 2665 (Blackmun, J., dissenting).
[53] For the text of RLA section 2 Eleventh, *see supra* note 16.
[54] *Beck*, 108 S. Ct. at 2649 (citing various sections of NLRA legislative history).
[55] Almost all of the major union security litigation under the RLA and the NLRA has been supported by the National Right to Work Legal Defense and Education Foundation. Wright, *Clipping the Political Wings of Unions: An Examination of Existing Law and Proposals for Change*, 5 HARV. J.L. & PUB. POL'Y 1, 34 n.203 (1982). The National Right to Work Legal Defense and Education Foundation was founded, and is largely supported, by employer interests. UAW v. National Right to Work Legal Defense & Educ. Found., Inc., 781 F.2d 928, 929, 934 (D.C. Cir. 1986); *see also* UAW v. National Right to Work Legal Defense & Educ. Found., Inc., 590 F.2d 1139 (D.C. Cir. 1978).

agreements allowed under section 8(a)(3) and section 2 Eleventh. The test came first under section 2 Eleventh of the RLA.

In *Railway Employer's Department v. Hanson*[56] the Supreme Court faced the question of whether a union shop agreement under section 2 Eleventh of the RLA violated dissenting employees' first amendment freedom of association and fifth amendment "liberty" to work. The parties did not dispute, nor did the Court question, whether a union shop was allowed by section 2 Eleventh. Although the Court recognized that section 2 Eleventh did not compel the parties to agree to a union shop, it found the prerequisite state action to support the constitutional claims in the section's pre-emption of an applicable state constitutional provision prohibiting union shop agreements.[57] Because of this pre-emption, the Court viewed "the federal statute [as] the source of . . . power and authority by which [the employees'] private rights are lost or sacrificed."[58] However, no violation of the employees' first or fifth amendment rights was found. Rather, the Court found that trade unions and their support strengthen the liberty to work.[59] Moreover, it held that with respect to the periodic dues and fees allowed by the RLA, any infringement of first amendment rights was justified by the state's compelling interest in promoting collective bargaining and industrial peace.[60] The Court specifically reserved the question of whether the imposition of "other conditions . . . or . . . the exaction of dues . . . as a cover for forcing ideological

---

[56] 351 U.S. 225 (1956).

[57] The Court in passing alluded to several theories for state action including Smith v. Allwright, 321 U.S. 649 (1944), and Steele v. Louisville & N. R.R., 323 U.S. 192 (1944). The mention of *Smith* perhaps suggests that because under the RLA the government once prohibited union security, the withdrawal of this prohibition made union security agreements state action. Read, *Minority Rights and the Union Shop: A Basis for Constitutional Attack*, 49 MINN. L. REV. 227, 243 (1964). *Steele* suggests that the union's designation as the exclusive representative under section 9(a) of the NLRA makes it a state actor. The Court, in dicta, also expressed its opinion that if a court were to enforce a union security agreement, a valid state action claim would arise. Railway Employees Dep't v. Hanson, 351 U.S. 225, 232 n.4 (1956) (citing Shelley v. Kraemer, 334 U.S. 1 (1948)). However, the basis of the Court's holding of state action in *Hanson* is the exercise of the Supremacy Clause. *Hanson*, 351 U.S. at 232; Wellington, *The Constitution, the Labor Union, and "Governmental Action"*, 70 YALE L.J. 345, 355–56 (1961).

[58] *Hanson*, 351 U.S. at 232.

[59] *Id.* at 235.

[60] *Id.* at 238. The Court's opinion is ambiguous as to whether it found no infringement of dissenters' first amendment rights or whether it found such infringement justified. The later interpretation of *Hanson* was adopted by the Court in Abood v. Detroit Bd. of Educ., 431 U.S. 209 (1977).

conformity or other action . . ." would be a violation of dissenters' first amendment rights.[61]

Five years later, the Court determined that the very question it had reserved in *Hanson* was raised by the facts in *International Association of Machinists v. Street*.[62] In *Street*, dissenting employees had challenged the constitutionality of an agency shop agreement negotiated under section 2 Eleventh.[63] The trial court found that, in fact, the union had spent a portion of the money it had collected in dues on political activities.[64] Again, neither party disputed the legality of the agency shop agreement under section 2 Eleventh.[65] However, after citing the doctrine of "fairly" construing statutes to avoid serious questions of constitutionality, the Supreme Court held that it did not have to answer the constitutional question. It interpreted section 2 Eleventh to allow the compulsion of dues for collective bargaining purposes, but not for political purposes.[66] To support this conclusion, the Court argued that the legislative history of the 1951 revisions to the RLA showed that Congress's purpose in allowing union security agreements under the RLA was to allow unions to recoup the expenses of their obligations under the RLA in negotiating and enforcing the collective agreement.[67] Thus, dissenters could not be compelled to fund political activities. The Court concluded that this was not only a "fair" but a "reasonable" interpretation of section 2 Eleventh, implying that it could have reached this interpretation without resort to the doctrine of construing statutes to avoid constitutional questions.[68] The holding in *Street* was further refined in *Ellis v. Brotherhood of Railway, Airline and Steamship Clerks*,[69] where the Court interpreted section 2 Eleventh to allow the compulsion of dues only to cover expenses "necessarily or reasonably"

---

[61] *Hanson*, 351 U.S. at 238.
[62] 367 U.S. 740 (1961).
[63] *Id.* at 743–44.
[64] *Id.* at 744–45.
[65] *Id.* at 803 (Frankfurter, J., dissenting).
[66] *Id.* at 749–50.
[67] *Id.* at 764.
[68] *Id.* at 750; *but see id.* at 786 (Black, J., dissenting) ("Yet no one has suggested that the Court's statutory construction of § 2 Eleventh could possibly be supported without the crutch of its fear of unconstitutionality."). "In order to avoid the constitutional issue assumed to be lurking in *Street*, Justice Brennan's opinion tortured the legislative history to find a congressional limitation on the use of union security for political purposes." Cantor, *supra* note 20, at 72.
[69] 466 U.S. 435 (1984).

incurred by the union in the performance of its collective bargaining duties under the Act.[70]

A test of the constitutionality of the negotiation and observance of agency shop agreements next arose in the public sector in *Abood v. Detroit Board of Education*.[71] In *Abood*, the legality of union security agreements for state employees was governed by a state statute modeled after both section 8(a)(3) and section 2 Eleventh.[72] The Court could not avoid the constitutional question because it was bound by a state court interpretation that the statute allowed the negotiation and observance of agency shop agreements.[73] Because the employer in the case was a state government, there was no question that the negotiation and observance of the agency shop agreement constituted state action. The Court held that the negotiation and enforcement of an agency shop agreement infringed dissenting employees' first amendment rights. The majority reasoned that although this infringement was justified with respect to collective bargaining expenses by the state's interest in promoting collective bargaining,[74] it could not be justified with respect to expenses for political activities.[75] Thus, the Court held that where there is state action in the negotiation and observance of a union security agreement, the agency shop violates dissenters' first amendment rights.

B. Communications Workers of America v. Beck

The test of the extent of union security agreements allowed under section 8(a)(3) of the NLRA finally arose in *Communications Workers of America v. Beck*.[76] Plaintiffs, twenty dis-

[70] *Id.* at 448.
[71] 431 U.S. 209 (1976).
[72] *Id.* at 223–24.
[73] *Id.* at 211, 214–15.
[74] *Id.* at 225–26.
[75] *Id.* at 235–36.
[76] 108 S. Ct. 2641 (1988). There is a reason why the test of section 8(a)(3) came after that of section 2 Eleventh. Under the NLRA, primary jurisdiction for the enforcement and interpretation of section 8(a)(3) lies with the National Labor Relations Board. 29 U.S.C. § 160(a) (1982). The Board has consistently given a broad interpretation of the extent of union security agreements allowed under section 8(a)(3). *See infra* notes 169–188 and accompanying text. Thus, the Board has not prosecuted cases to test the limits of section 8(a)(3). There is no corresponding administrative agency under the RLA to restrict individual tests of particular sections of the Act. The plaintiffs in *Beck* escaped the primary jurisdiction of the Board by alleging a violation of the duty of fair representation and of the Constitution, raising the interpretation of section 8(a)(3) as a collateral issue. 108 S. Ct. at 2647. The Supreme Court failed to discuss substantively either the duty of fair representation or the constitutional claim in its opinion in *Beck*.

senting employees, objected to an agency shop agreement that required them to pay agency fees equal to periodic union dues.[77] It was undisputed that a portion of the dues and agency fees collected by the union had been used for political purposes.[78] The plaintiffs argued that the agency shop agreement violated section 8(a)(3) of the NLRA, or in the alternative, that the agreement violated their first amendment rights and that section 8(a)(3) was therefore unconstitutional.[79]

The parties' briefs before the Supreme Court suggest four major areas of contention.[80] First, the parties disputed the relevance of the Supreme Court's interpretations of section 2 Eleventh in *Street* and *Ellis*.[81] The plaintiffs argued that the Court's RLA precedents should govern the interpretation of section 8(a)(3) because of the similarity in the language of the two sections and the legislative history of section 2 Eleventh, which suggested that Congress intended it to have the same scope as section 8(a)(3).[82] The union attempted to distinguish the RLA precedents, arguing that the RLA had a different legislative history[83] and that the Court's interpretations of section 2 Eleventh were colored by its attempt to avoid constitutional questions that do not arise under the NLRA.[84]

---

[77] 108 S. Ct. at 2645. The plaintiffs' cause was supported and financed by the National Right to Work Legal Defense and Education Foundation. *See* Brief for the AFL-CIO at 24, *Beck*, 108 S. Ct. 2641. Many interested parties filed *amicus* briefs before the Supreme Court. The AFL-CIO filed a brief supporting the Communications Workers of America, as did Solicitor General Charles Fried, who was joined by NLRB General Counsel Rosemary Collyer. The Landmark Legal Foundation, the Pacific Legal Foundation (joined by a group of dissident Screen Actors Guild members including Charlton Heston, Claude Akins, and Rory Calhoun), and a group of four Senators (Jessie Helms (R-N.C.), Strom Thurmond (R-S.C.), Dan Quayle (R-Ind.), and Steven Symms (R-Idaho)) filed briefs supporting plaintiffs.

[78] *Beck*, 108 S. Ct. at 2645.

[79] *Id.* at 2645–46. The plaintiffs also argued that the negotiation and enforcement of an agency shop agreement by the union violated the union's duty of fair representation, and that if the NLRA allowed the agency shop, it violated the plaintiffs' fifth as well as first amendment rights. Brief for Respondents at 29–44. However, these arguments were not addressed by the Supreme Court and are beyond the scope of this Article.

[80] Not all of these arguments were raised before the district and circuit courts, but they are presented at this point to facilitate exposition. For a summary of the oral presentation of these arguments before the Supreme Court, see 56 U.S.L.W. 3475 (Jan. 19, 1988).

[81] *See supra* notes 62–70 and accompanying text.

[82] Brief for Respondents at 23–27.

[83] The union argued that the history of the NLRA was one of first allowing all union security agreements and then prohibiting only the worst abuses, while the history of the RLA was one of first prohibiting all union security and then allowing only limited union security. Thus, the union argued that the legislative histories of the two acts with respect to union security are diametrically opposed. Brief for Petitioners at 41–42.

[84] *Id.* at 40–41. The constitutional question which the Court avoided in its RLA cases

Second, the parties differed in their interpretations of the plain language of section 8(a)(3). The plaintiffs argued that the definitions of "labor organization" and "exclusive representative" in the NLRA limited the purpose of unions under the statute to collective bargaining.[85] Because section 8(a)(3) allows only exclusive representatives to negotiate union security agreements, the plaintiffs reasoned that the section limited the "dues" which the representative could compel to those necessary to cover collective bargaining expenses.[86] The union disputed whether the NLRA limited union purposes to collective bargaining and argued that the plain language of section 8(a)(3) allows unions to charge all employees "uniform dues" on pain of loss of their jobs.[87]

Third, the parties had different interpretations of the legislative history of section 8(a)(3). The plaintiffs contended that Congress's purpose in allowing union security agreements under section 8(a)(3) was to prevent "free-riding" by nonunion employees who benefitted from collective bargaining but did not share its cost.[88] Accordingly, they argued that the only fees allowed in union security agreements under section 8(a)(3)

---

was whether an agency shop agreement between a private union as exclusive representative and a private employer violates dissenting employees' first amendment rights. International Ass'n of Machinists v. Street, 367 U.S. 740 (1961). Although a similar constitutional question was raised in *Beck*, the union argued that *Beck* did not pose the same constitutional concern as the RLA cases because it did not share the basis for the finding of state action in the RLA cases. The basis for finding state action in the RLA union security cases was the exercise of the Supremacy Clause by Congress in pre-empting state legislation which prohibited union security agreements. Railway Employees' Dep't v. Hanson, 351 U.S. 225, 232 (1956). There is no such pre-emption under the NLRA. *See* 29 U.S.C. § 164(b) (1982).

[85] Brief for Respondents at 19–20. Section 2(5) of the NLRA defines a "labor organization" as

> . . . any organization of any kind, or any agency or employee representation committee or plan, in which employees participate and which exists for the purpose, in whole or in part, of dealing with employers concerning grievances, labor disputes, wages, rates of pay, hours of employment, or conditions of work.

29 U.S.C. § 152(5) (1982). Section 9(a) of the NLRA gives the following "definition" of "exclusive representative:"

> Representatives designated or selected for the purposes of collective bargaining by the majority of the employees in a unit appropriate for such purposes, shall be the exclusive representatives of all the employees in such unit for the purposes of collective bargaining in respect to rates of pay, wages, hours of employment, or other conditions of employment . . . .

29 U.S.C. § 159(a) (1982).

[86] Brief for Respondents at 18.

[87] Brief for Petitioners at 21–23.

[88] Brief for Respondents at 21.

would be payments for collective bargaining services.[89] The union countered that, prior to the Taft-Hartley amendments, all forms of union security agreements had been legal; and in enacting the amendments, Congress had intended to prohibit only the worst abuses of union security—the closed shop and union discrimination in membership. Thus, Congress did not intend to intrude into the internal affairs of unions or to place restrictions on the use of union funds.[90]

Finally, the parties disputed whether there was sufficient state action to support the plaintiffs' constitutional claim. The plaintiffs argued that the state action prerequisite was met by the union's designation as the exclusive bargaining representative under section 9(a).[91] They quoted dicta from *Steele v. Louisville & Nashville Railroad Co.*[92] suggesting that because the Union, as exclusive representative, is "clothed with power not unlike that of a legislature" it may be a state actor and subject to constitutional scrutiny.[93] The plaintiffs also argued that the designation of the union as the exclusive representative under the NLRA encourages the negotiation of agency shop agreements because it prevents individual employees from negotiating employment contracts without such agreements, and it requires the employer to negotiate such agreements in good faith.[94] The union maintained that its designation as the exclusive bargaining representative did not make it a state actor since the Supreme Court had recently established that governmental grants of mo-

---

[89] *Id.* at 17.

[90] Brief for Petitioners at 38–39.

[91] Brief for Respondents at 5.

[92] 323 U.S. 192 (1944).

[93] Brief for Respondents at 11. The quote from *Steele* is as follows:

> For the representative is clothed with power not unlike that of a legislature which is subject to constitutional limitations on its power to deny, restrict, destroy or discriminate against the rights of those for whom it legislates and which is also under an affirmative constitutional duty equally to protect those rights. If the Railway Labor Act purports to impose on . . . [the nonunion employees] the legal duty to comply with the terms of a contract whereby the representative has discriminatorily restricted their employment for the benefit and advantage of the . . . [union's] own members, we must decide the constitutional questions, which . . . [the nonunion member] raises . . . .

323 U.S. 192, 198–99. *See also id.* at 208 (Murphy, J., concurring). Although *Steele* is a RLA case, the Court has found a similar duty of fair representation under the NLRA, Ford Motor Co. v. Huffman, 345 U.S. 330 (1953), and considers the reasoning of *Steele* equally applicable to NLRA-governed unions.

[94] 29 U.S.C. § 158(a)(5) (1982); Brief for Respondents at 5, *Beck*, 108 S. Ct. 2641. The Supreme Court established that union security agreements are a mandatory subject of bargaining under section 8(a)(5) of the NLRA in NLRB v. General Motors Corp., 373 U.S. 734 (1963).

nopoly status to private institutions in the provision of services did not make those institutions state actors.[95] The union saw no government encouragement of union security agreements since section 8(a)(3) does not require, but merely permits, such agreements.[96] The union closed its argument by noting that the Supreme Court had previously rejected the notion that a union's designation as the exclusive bargaining representative made its actions state action.[97]

The district court, acting *sua sponte*, granted partial summary judgment to the plaintiffs. Without ruling on the legality of the agency shop under the NLRA, the trial court found that "collect[ing] from the plaintiffs amounts beyond that allocable to collective bargaining . . . violates the First Amendment rights of the plaintiffs."[98] On appeal, a divided panel for the Fourth Circuit affirmed the district court's decision regarding the union's liability.[99] Based on the Supreme Court's interpretation of section 2 Eleventh, the majority held that section 8(a)(3) did not allow the compulsion of fees for purposes unrelated to collective bargaining.[100] The majority also stated in dicta that such collection would violate the employees' first amendment rights.[101] The Fourth Circuit granted the union's request for

---

[95] Brief for Petitioners at 17–18, *Beck*, 108 S. Ct. 2641 (citing San Francisco Arts & Athletics v. United States Olympic Comm., 483 U.S. 522 (1987); Jackson v. Metropolitan Edison Co., 419 U.S. 345 (1975)).

[96] Brief for Petitioners at 13–15, *Beck*, 108 S. Ct. 2641.

[97] *Id.* at 18–19 (citing United Steelworkers v. Sadlowski, 457 U.S. 102, 122 n.16 (1982) ("union rule precluding candidates for union office from accepting campaign contributions from nonmembers 'does not involve state action'"); United Steelworkers v. Weber, 443 U.S. 193, 200 (1979) ("collectively bargained affirmative action plan 'does not involve state action'" (dicta)).

[98] Beck v. Communications Workers of Am., 468 F. Supp. 93, 97 (D. Md. 1979), *aff'd in part*, 776 F.2d 1187 (4th Cir. 1985), *aff'd en banc*, 800 F.2d 1280 (4th Cir. 1986), *aff'd*, 108 S. Ct. 2641 (1988). The court ordered the union to return to the plaintiffs all dues for expenses other than "collective bargaining, contract administration and grievance adjustment" since January 1, 1976, and to institute a record-keeping system to segregate expenses for collective bargaining and non-collective bargaining purposes. *Id.* The district court appointed a special master who determined the percent of the union's expenditures allocable to collective bargaining since the beginning of 1976. *Beck*, 776 F.2d at 1191–92.

[99] However, the court, unclear about the standard used by the appointed special master in the district court, held that the "preponderance of evidence" standard, rather than the "clear and convincing" standard, should have been applied to the union's burden of showing that expenses were related to collective bargaining. The case was remanded for further proceedings. *Beck*, 776 F.2d at 1212.

[100] *Id.* at 1201–02. The majority also held that the union's negotiation and enforcement of an agency shop agreement violated its duty of fair representation. *Id.* at 1203.

[101] *Id.* at 1205.

rehearing *en banc*[102] and, by a six to four majority, affirmed the panel majority's disposition of the case.[103] Because the Fourth Circuit's opinion in *Beck* directly conflicted with the Second Circuit's opinion in *Price v. United Auto Workers*,[104] the Supreme Court granted certiorari to resolve the conflict.[105]

The Supreme Court affirmed the judgment of the Fourth Circuit Court of Appeals, holding that section 8(a)(3) of the NLRA did not allow the compulsion of dues for purposes unrelated to collective bargaining.[106] The majority held that the Court's prior interpretation of section 2 Eleventh, announced in *International Association of Machinists v. Street*,[107] was controlling with respect to the interpretation of section 8(a)(3).[108] They described the two provisions as "identical" in "all material respects" and noted that the legislative history of section 2 Eleventh suggests that Congress intended to model it after section 8(a)(3).[109] Based on this identity, the majority imposed on section 8(a)(3) the same narrow purpose for union security agreements the Court had found under section 2 Eleventh in *Street*. They asserted that the "nearly identical language [of the two sections] reflects the fact that, in both, Congress authorized compulsory unionism only to the extent necessary to ensure that those who enjoy union-negotiated benefits contribute to their cost."[110] The majority concluded that "in these circumstances, we think it clear that Congress intended the same language to have the same meaning in both statutes."[111]

---

[102] Beck v. Communications Workers of Am., 800 F.2d 1280 (4th Cir. 1986), *aff'd*, 108 S. Ct. 2641 (1988).

[103] The *en banc* opinion of the court gives little insight into the rationale of the Fourth Circuit in disallowing agency shop payments. The arguments of the parties at the *en banc* hearing focused primarily on whether the NLRB or the federal courts had primary jurisdiction over the case. *Id.* at 1282. This was probably due to the fact that the panel majority erroneously cited RLA cases as authority for jurisdiction even though no such jurisdictional issue arises under the RLA. *Beck*, 108 S. Ct. at 2647. As a result, the *en banc* opinions are expressed in terms of the judges' beliefs regarding the court's jurisdiction, rather than in terms of the merit of the questions at hand. The *per curiam* opinion indicated that five judges found federal jurisdiction over the plaintiffs' claims under both section 8(a)(3) and the duty of fair representation, while one judge found jurisdiction only over the plaintiffs' duty of fair representation claim alone. *Beck*, 800 F.2d at 1282. The opinion indicated that all six of these judges had voted to affirm "the majority panel opinion's disposition of the allocation issue." *Id.*

[104] 795 F.2d 1128 (2d Cir. 1986).

[105] Communications Workers of Am. v. Beck, 482 U.S. 904 (1987) (mem.).

[106] *Beck*, 108 S. Ct. at 2657.

[107] 367 U.S. 740 (1961).

[108] 108 S. Ct. at 2648.

[109] *Id.* at 2648–49.

[110] *Id.* at 2649.

[111] *Id.*

The Court attempted to support its narrow interpretation of the purpose of union security agreements under section 8(a)(3) by referring to the legislative history of the NLRA. The majority argued that during the enactment of the Taft-Hartley amendments, Congress was not only concerned that the closed shop "create[d] too great a barrier to free employment," but was "equally concerned . . . that without such agreements, many employees would reap the benefits that unions negotiated on their behalf without . . . contributing financial support to those efforts."[112] They concluded that these dual purposes, one for limiting union security and one for allowing the continuation of some form of union security, were both represented in section 8(a)(3).[113]

The Court gave a very expansive reading of the Congress's purpose in limiting union security agreements under section 8(a)(3). The union's argument that in enacting section 8(a)(3) Congress sought only to prohibit the worst abuses of union security under the Wagner Act was rejected. Instead, the majority maintained that in 1947 "Congress viewed the Wagner Act's regime of compulsory unionism as seriously flawed,"[114] and did not "set out . . . simply to tinker in some limited fashion with the [Act's] authorization of union-security agreements."[115] According to the majority, Congress retained union security to the extent that it did only because "such agreements promoted stability [in labor relations] by eliminating 'free-riders.'"[116] To strengthen their argument, the majority resorted to the unusual practice of quoting a minority report on the Taft-Hartley Act. The report asserted that the Act allowed "union-shop agreement[s] only under limited and administratively burdensome conditions."[117]

---

[112] *Id.* at 2649–50 (quoting S. REP. No. 105, 80th Cong., 1st Sess. 6 (1947) [hereinafter S. REP. No. 105], *reprinted in* 1 LEGISLATIVE HISTORY OF THE LABOR MANAGEMENT RELATIONS ACT 407, 412 (1974) [hereinafter LMRA LEGISLATIVE HISTORY]).

[113] *Id.* at 2650.

[114] *Id.* at 2653.

[115] *Id.* at 2653–54 (quoting S. REP. No. 105, *supra* note 112, at 7, *reprinted in* 1 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 407, 413).

[116] *Id.*

[117] *Id.* at 2653 (quoting S. REP. No. 105, Pt. 2, 80th Cong., 1st Sess. 8, *reprinted in* 1 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 463, 470 (Minority Report)). This practice seems ill-advised given the natural tendency of opponents to exaggerate the purpose and effect of a bill in seeking its demise. For example, opponents of the Wagner Act alleged that it repealed the thirteenth amendment, creating labor despots and czars. *Hearings Before the Senate Comm. on Education and Labor*, 73d Cong., 2d Sess. 965–66 (1934) [hereinafter *1934 Senate Hearings*], *reprinted in* 1 LEGISLATIVE HISTORY OF

The Court limited the purpose of union security under section 8(a)(3) to merely the recoupment of collective bargaining expenses. The majority did this by reference to the analogy between section 8(a)(3) and section 2 Eleventh. They insisted that the same concern over free-riders that had caused Congress to retain limited union security under section 8(a)(3) had also prompted it to enact section 2 Eleventh.[118] The union's argument, that the different legislative histories of the two sections required that they be interpreted differently, was rejected on the basis of the legislative history of section 2 Eleventh, which suggested that it was to allow the same union security provisions as section 8(a)(3).[119] The union's assertion that the Court's interpretations of section 2 Eleventh had been constitutionally colored was likewise dismissed. The Court maintained, as it had in *Street*, that its interpretation of section 2 Eleventh was "not only 'fairly possible' but entirely reasonable" and thus not distorted by constitutional concerns.[120]

The Court's interpretation of section 8(a)(3) obviated the need to answer the constitutional question. However, in passing the majority made an interesting attribution on the issue of whether the exercise of rights permitted by section 8(a)(3) involves state action. They cited as analogous the holdings in *Steelworkers v. Sadlowski*[121] and *Steelworkers v. Weber*,[122] the cases cited by the union as supporting its argument that there is no state action.[123] Without noting that the relevant language in *Weber* is dicta, the majority attributed to that case the holding that "negotiation of [a] collective bargaining agreement's affirmative action plan does not involve state action."[124] This attribution would seem to defeat any arguments that a union's actions in negoti-

---

THE NATIONAL LABOR RELATIONS ACT 27, 1003–04 (1959) [hereinafter NLRA LEGISLATIVE HISTORY]. Opponents of the Taft-Hartley Act alleged that it would once again allow "yellow dog" contracts between an employer and his employees, restraining them from joining a union during the term of their employment. 93 CONG. REC. 6672 (1947), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1565, 1588–89 (comments of Sen. Pepper (D-Fla.)). Moreover, since it is the majority of the legislature who enact the law, the opinion of the minority as to its purpose and effect is, arguably, irrelevant.

[118] *Beck*, 108 S. Ct. at 2653.

[119] *Id.* at 2654.

[120] *Id.* at 2657 (citing International Ass'n of Machinists v. Street, 367 U.S. 740, 750 (1961)).

[121] 457 U.S. 102 (1982).

[122] 443 U.S. 193 (1979).

[123] *Id.* at 2656–57 (citing United Steelworkers v. Sadlowski, 457 U.S. 102 (1982); United Steelworkers v. Weber, 443 U.S. 193 (1979)).

[124] *Id.* at 2657.

ating and observing a union security agreement are state actions by virtue of the union's designation as the exclusive bargaining representative under section 9(a) of the NLRA.

Justice Blackmun, joined by Justices O'Connor and Scalia, dissented from the Court's interpretation of section 8(a)(3). Blackmun stated that he was unable to join the majority's opinion because he was "unwilling to offend our established doctrines of statutory construction and strain the meaning of the language used by Congress in section 8(a)(3), simply to conform section 8(a)(3)'s construction to the Court's interpretation of similar language in a different later-enacted statute . . . ."[125] He further commented that the interpretation of section 2 Eleventh was itself "not without its difficulties."[126] In criticizing the majority's decision, Blackmun first argued that the plain language of section 8(a)(3) allowed the agency shop, that this had been the consistent interpretation of the statute by the National Labor Relations Board (NLRB), and that the Court had previously approved agency shop agreements.[127] Second, he asserted that the legislative history of the NLRA showed that Congress's purpose in limiting union security under the Taft-Hartley amendments to section 8(a)(3) was to proscribe only the most serious abuses of union security—the closed shop and union discrimination in membership and employment.[128] He urged that the legislative history did not support the majority's conclusion that Congress's "single minded" purpose in continuing to allow union security under section 8(a)(3) was to prevent free-riders.[129] To Blackmun, nothing in the legislative history suggested that Congress intended to outlaw the agency shop or limit the dues that can be collected from dissenting employees.[130] Finally, Blackmun had no problem with interpreting section 8(a)(3) differently from section 2 Eleventh, since he saw these statutes as born of different concerns with different legislative histories.[131] Rather, he reasoned that even if Congress intended in 1951 to enact section 8(a)(3) into the RLA by enacting section 2 Eleventh, the events of 1951 were irrelevant to the determination of what

---

[125] *Id.* at 2658 (Blackmun, J., dissenting) (citing Abood v. Detroit Bd. of Educ., 431 U.S. 209, 232 (1977)).
[126] *Id.*
[127] *Id.* at 2659.
[128] *Id.* at 2661.
[129] *Id.* at 2664.
[130] *Id.* at 2661–62.
[131] *Id.* at 2664–65.

Congress enacted in section 8(a)(3) in 1947.[132] For Blackmun, "[t]he relevant sources for gleaning the 1947 Congress' intent are the plain language of section 8(a)(3), and . . . the legislative history of section 8(a)(3)."[133]

II. An Interpretation of Section 8(a)(3) of the NLRA by Direct Examination of Its Words, Administrative Interpretations, and Legislative History

The Supreme Court's interpretation of section 8(a)(3) in *Beck* does not follow the usual rules of statutory construction. The canons of construction require that a court first examine the "plain meaning" of the words of the statute,[134] and then, if the meaning is ambiguous, the administrative interpretations and legislative history.[135] The *Beck* Court referred to the words of section 8(a)(3) only to note its similarity to section 2 Eleventh of the RLA.[136] The Court ignored the NLRB's *amicus curiae* position that section 8(a)(3) allows agency shop agreements.[137] It alluded to the Board's prior opinions on section 8(a)(3) only to note some errant language which, by the Court's own admission, deviated from the Board's other interpretations of section 8(a)(3).[138] The majority examined the legislative history of section 8(a)(3) only to corroborate its imposition of the RLA precedents rather than to make a direct determination of the section's meaning.[139] Instead of undertaking a direct examination of the words, administrative interpretations, and legislative history of section 8(a)(3), the Court identified section 8(a)(3) with section 2 Eleventh and then imposed on section 8(a)(3) its previous interpretations of section 2 Eleventh.[140] Thus, the Court bypassed the traditional starting points for statutory construc-

---

[132] *Id.* at 2665–66.

[133] *Id.* at 2666.

[134] Caminetti v. United States, 242 U.S. 470 (1916) (a court should first look to the plain meaning of the statutory language); *see also* 2A N. Singer, Sutherland Statutory Construction § 46.01 (4th ed. 1984).

[135] Posner, *supra* note 25, at 807–08; *Ex parte* Collett, 337 U.S. 55, 61 (1949) (extrinsic aids, such as legislative history, are used only when the statute is unclear); *see also* 2A N. Singer, *supra* note 134, § 48.01.

[136] *Beck*, 108 S. Ct. at 2648–49.

[137] Brief for the United States at 12–19, 23, *Beck*, 108 S. Ct. 2641.

[138] *Beck*, 108 S. Ct. at 2652 n.7.

[139] *Id.* at 2649–50.

[140] *Id.* at 2648–49.

tion in favor of the more remote canon of interpreting similar statutory provisions the same.[141]

One need not subscribe to the belief that the canons of construction provide a determinative formula for the discovery of the "true" meaning of a statute to object to the Court's analysis in *Beck*.[142] Almost all critics of the canons of construction have argued that a statute's words and legislative history are the best evidence to examine in discerning a meaning grounded in the intent of the legislature.[143] Some commentators have also argued that, regardless of the canons of construction, administrative interpretations merit deference by the courts where the relevant agency has special expertise in the subject area or where experience in the statute's enactment has given the agency special insight into the legislature's intent.[144] Deference by the courts to the intent of the elected legislature is, of course, a fundamental premise of our democratic government.[145]

It would not be impossible to derive an interpretation of a statute which fairly represented the intent of the legislature by interpreting the statute identically to a previous interpretation of a similar statute without direct examination of the statute's words, administrative interpretations, or legislative history. Presumably the similar language of the two statutes should convey the same meaning to the court and represent the same legislative purposes in their enactment. Moreover, if the similarity of the statutes was recognized when the court first considered their meaning, the court might legitimately consider the words, administrative interpretations, and legislative history of both statutes in its initial interpretation. There may even be some policy

---

[141] For a discussion of this canon, see 2A N. SINGER, *supra* note 134, § 51.02.

[142] This myth was of course "debunked" long ago by Karl Llewellyn. Llewellyn, *Remarks on the Theory of Appellate Decision and the Rules or Canons About How Statutes are to be Construed*, 3 VAND. L. REV. 395 (1950).

[143] R. DWORKIN, LAW'S EMPIRE 313–54 (1986) (statutory interpretation is part of an ongoing process of creating new and evolving meaning from the text of the statute, in which the court should read that text in light of the political history of the statute); H. Hart & A. Sacks, The Legal Process: Basic Problems in the Making and Application of Law 1156–57 (tent. ed. 1957) (in interpreting a statute, a court should examine the statute's words and history to determine legislative purpose); Posner, *supra* note 25, at 817–18 (in interpreting a statute, a court should look at its words, history, and the values of the enacting legislature to reconstruct how the enacting legislators would have wanted the statute applied to the case at bar); *but see* J. CULLER, ON DECONSTRUCTION: THEORY AND CRITICISM AFTER STRUCTURALISM 123–24 (1982) (the meaning of a statute cannot be determined out of context, but because context knows no bounds of description, interpretation is utterly subjective).

[144] *See, e.g.*, Posner, *supra* note 25, at 811.

[145] *See, e.g.*, Hatch, *Modern* Marbury *Myths*, 57 U. CIN. L. REV. 891, 893–94 (1989).

reason expressed or implied by Congress as to why the statutes should be interpreted similarly, regardless of extrinsic evidence to the contrary. However, by imposing its previous interpretation of a similar statute without direct examination of the words, administrative interpretations, and legislative history of the statute in question, a court ignores the best evidence of congressional intent in favor of assumptions and evidence concerning a different statute.

In this Part, I undertake a direct examination of section 8(a)(3)'s words, administrative interpretations, and legislative history to determine Congress's purposes in allowing union security agreements under the NLRA and to determine whether Congress intended to allow the observance of agency shop agreements under section 8(a)(3). The results of this examination can be used to verify whether the Court's circuitous route of statutory interpretation in *Beck* has led to a fair interpretation of Congress's intent.

A. *The Words of the Statute*

Under the canons of construction, the usual starting point for statutory interpretation is the plain meaning of the words of the statute. If the words of the statute are unambiguous, the court ends its inquiry and adopts their plain meaning as its interpretation of the statute.[146] Even divorced from the rhetoric of the canons of construction, the words of the statute are of course the primary intrinsic evidence of congressional intent in the interpretation of a statute.[147]

On its face, the language of section 8(a)(3) of the NLRA permits the negotiation of union security agreements that require all employees to pay dues and initiation fees equal to those of union members.[148] The first proviso of the section allows employers and the unions designated as exclusive bargaining representatives to enter into union security agreements which require all employees to become members of the union as a post-condition of employment.[149] The second proviso prevents the employer from discharging an employee covered by the union

[146] Caminetti v. United States, 242 U.S. 470, 485 (1916).

[147] *See generally* W. ESKRIDGE & P. FRICKEY, *supra* note 24, at 639–46; Posner, *supra* note 25, at 807–08.

[148] 29 U.S.C. § 158(a)(3) (1982).

[149] *See supra* text accompanying note 41.

security agreement if ". . . [union] membership was not available to the employee on the same terms and conditions generally applicable to other members . . ." or if the employee is denied membership for any reason other than "the failure . . . to tender the periodic dues and the initiation fees uniformly required as a condition of acquiring or retaining membership."[150] There is no ambiguity that the dues and fees that can be charged are equal to full membership dues and fees. By adding the word "uniformly" to the proviso, Congress emphasized that all employees are to be charged the same dues and fees.[151] Nothing in the language of section 8(a)(3) suggests that there is any difference between the dues that can be charged dissenting and non-dissenting employees, or that there are any limitations on the union expenses that can be included in these dues and fees.

The only ambiguity in this plain reading of section 8(a)(3) concerns the exact form of union security agreement allowed by the section. The first proviso, in allowing union security agreements requiring membership as a post-condition of employment, defines a union shop agreement.[152] The second proviso, in stating that union security agreements can be enforced only to the extent of requiring the payment of uniform dues, defines an agency shop agreement.[153] This conflict was resolved by the Supreme Court in *NLRB v. General Motors Corp.*,[154] where the Court held that "membership," as used in section 8(a)(3) after the Taft-Hartley amendments, consists only of the obligation to pay dues.[155] Agency shop obligations thus became the "practical equivalent" of membership obligations for the purposes of section 8(a)(3).[156] With the resolution of this question in favor of the agency shop, the language of section 8(a)(3) unambiguously allows employers and unions to negotiate and observe agency shop agreements that require all employees, as

---

[150] *See supra* text accompanying note 47.

[151] *Id.*

[152] *See supra* text accompanying note 34.

[153] *See supra* text accompanying note 35.

[154] 373 U.S. 734 (1963).

[155] *Id.* at 742. In *General Motors*, the employer refused to bargain with the union over an agency shop agreement, arguing that because the agency shop does not require membership, it was not allowed under the first proviso of section 8(a)(3). *Id.* The Court held that the Taft-Hartley amendments to the NLRA had "whittled" the term "membership" down to its "financial core" of merely the obligation to pay dues. *Id.* Accordingly, the Court ordered the employer to bargain with the union over the agency shop agreement. *Id.* at 744–45.

[156] *Id.* at 743.

a condition of employment, to pay dues and fees equal to those paid by union members.

The plaintiffs' argument in *Beck* with respect to the plain language of the statute is unpersuasive. They argued that the language of the NLRA limits the dues that can be charged employees under a union security agreement to collective bargaining expenses because the definitions of "labor organization" and "exclusive representative" are limited to collective bargaining and section 8(a)(3) specifies that only exclusive representatives can negotiate union security agreements.[157] In fact the Act defines a "labor organization" as an organization which exists "*in whole or in part*" for the purposes of collective bargaining.[158] Moreover, in the "findings and declaration of policy" of the NLRA, Congress stated that it is the policy of the United States to "protect . . . the exercise by workers of full freedom of association, self-organization, and designation of representatives . . . for the purpose of negotiating the terms and conditions of their employment *or other mutual aid or protection.*"[159] On its face, the limitation that only an exclusive representative can negotiate a union security agreement appears to strive to ensure the legitimacy of the union and its representation of the majority of employees rather than to limit the type of union security agreement it can negotiate.[160] Finally, both the definitions of "labor organization" and "exclusive representative" and the limitation in section 8(a)(3) were carried forward from the original Wagner Act.[161] Since the Wagner Act allowed all forms of union security, including the closed shop, it does not seem likely that these definitions and limitations were intended to limit the dues and fees a union can charge under a union security agreement.[162]

B. *The Interpretation of the National Labor Relations Board*

Another canon of statutory construction is that courts should give great deference to prior interpretations of the statute by

---

[157] Brief for Respondents at 19–20, *Beck*, 108 S. Ct. 2641.

[158] 29 U.S.C. § 152(5) (1982) (emphasis added).

[159] 29 U.S.C. § 151 (1982) (emphasis added).

[160] *Cf.* THE DEVELOPING LABOR LAW 1362 (C. Morris 2d ed. 1983) [hereinafter DEVELOPING LABOR LAW] (Wagner Act authorized union security agreements only for unions which "legitimately represented" the employees).

[161] *See* National Labor Relations Act, ch. 372, §§ 2(5), 8(3), 9(a), 49 Stat. 449, 450, 452–53 (1935) (current version at 29 U.S.C. §§ 152(5), 158(3), 159(a) (1982)).

[162] *See infra* notes 209–219 and accompanying text.

the administrative agency charged with the enforcement of that statute.[163] This is especially true where the agency has special expertise in the subject area, or where the agency's construction was undertaken by people who helped in the enactment of the statute.[164] Judge Posner has criticized this canon, pointing out that an administrative agency's interpretation of a statute may show more fidelity to the views of the current administration than the intent of the enacting Congress.[165] However, even Judge Posner admits that the interpretation of the administrative agency may be a valuable piece of extrinsic evidence of congressional intent where the agency has special knowledge of the subject matter or legislative history.[166] The NLRB was created precisely to provide a body with special expertise in labor relations to enforce the NLRA.[167] Moreover, the opinion of a General Counsel appointed by President Reagan that favors organized labor by arguing that section 8(a)(3) allows the negotiation of agency shop agreements does not seem subject to the criticism of political bias.

As previously mentioned,[168] the Board and the Department of Justice joined in an *amicus curiae* brief in *Beck*, supporting the union's arguments that section 8(a)(3) allowed the negotiation and observance of agency shop agreements.[169] Because of the Board's active support of the union's position, and because the Board's prior interpretations of section 8(a)(3) are fairly straightforward, those interpretations were not a primary matter of dispute between the parties in *Beck*. They do provide insight, however, into the Board's view of the legislative intent behind section 8(a)(3).

The Board has approved the negotiation and observance of agency shop agreements requiring agency fees equal to full union dues under section 8(a)(3) of the NLRA. In *American Seating Co.*,[170] the Board upheld the legality of a union shop agreement which allowed religious objectors to pay an agency

---

[163] *See, e.g.*, Chevron v. Natural Resources Defense, 467 U.S. 837, 843 (1984); NLRB v. Hendrick County Rural Elec. Membership Corp., 454 U.S. 170, 177, 178–90 (1981); Udall v. Tallman, 380 U.S. 1, 16 (1965).

[164] *Udall*, 380 U.S. at 16.

[165] Posner, *supra* note 25, at 811.

[166] *Id.*

[167] *See generally* J. GETMAN & B. POGREBIN, LABOR RELATIONS (1988); J. GROSS, THE MAKING OF THE NATIONAL LABOR RELATIONS BOARD (1974).

[168] *See supra* text accompanying note 137.

[169] Brief for the United States at 12–19, 23, *Beck*, 108 S. Ct. 2641.

[170] 98 N.L.R.B. 800 (1952).

fee equal to full union dues in lieu of joining the union.[171] The Board reasoned that such an agreement would have been legal under section 8(3) of the Wagner Act,[172] that the enactment of section 8(a)(3) in the Taft-Hartley amendments added only "qualifications not pertinent here," and that "the legislative history of the amended Act indicates that Congress intended not to illegalize the practice of obtaining support payments from nonunion members who would otherwise be 'free-riders.'"[173]

Similarly, in *General Motors Corp.*,[174] the Board approved the negotiation of a union security agreement which required all employees to pay agency fees equal to full union dues.[175] In upholding the lawfulness of the agency shop agreement under consideration, the Board stated: "[W]e are unable to distinguish, *so far as its legality is concerned*, the instant agency-shop proposal from any other union-security proposal which predicates a right of discharge only upon an employee's failure to tender *the equivalent of regular union dues and initiation fees*."[176] The Board stated that it was "impelled" to such a holding by "the clear intention of the Congress as expressed in Section 8(a)(3) of the Act, in the legislative history of the Wagner and Taft-Hartley Acts, and by the Board and court decisions in which that section has been construed."[177]

Moreover, the Board has specifically held that section 8(a)(3) allows unions to charge non-members and dissenters agency fees equal to full union dues without regard to whether such

---

[171] The Board examined the legality of the agreement to determine whether the collective bargaining agreement was valid and could act as a bar to severance of a portion of the bargaining unit by the employer. *Id.* at 801.

[172] *Id.* at 802 (citing Public Service Co. of Colorado, 89 N.L.R.B. 418 (1950)).

[173] *Id.*

[174] 130 N.L.R.B. 481 (1961), *vacated*, 133 N.L.R.B. 451 (1961), *enforcement denied*, 303 F.2d 428 (6th Cir., 1962), *rev'd*, 373 U.S. 734 (1963).

[175] *See supra* notes 154–155 and accompanying text. The Board was asked to decide whether the proviso to section 8(a)(3) allowed only union security agreements requiring "membership," such as the union shop, or whether the employer's duty to bargain extended to the agency shop, a lesser form of union security. *Id.* at 481. The Board initially decided that the proviso did in fact allow only union shop agreements requiring membership. *Id.* at 486, 500. To reach this conclusion the Board had to ignore its precedent under the Wagner Act that section 8(3), the precursor to section 8(a)(3) of the current NLRA, prescribed the maximum union security provisions allowed rather than the minimum. Public Service Co. of Colorado, 89 N.L.R.B. 418 (1950). However, on rehearing after a change in the membership of the Board, the Board followed its precedent under the Wagner Act and upheld the negotiation of the agency shop. General Motors Corp., 133 N.L.R.B. 451, 457 (1961), *enforcement denied*, 303 F.2d 428 (6th Cir. 1962), *rev'd*, 373 U.S. 734 (1963).

[176] General Motors Corp., 133 N.L.R.B. at 459 (original emphasis suppressed, emphasis added).

[177] *Id.* at 457 (citations omitted).

dues are used for collective bargaining or other purposes. In *Detroit Mailers Union No. 40 (Detroit Newspaper Publishers Ass'n)*,[178] the Board rejected arguments that the second proviso of section 8(a)(3) prevented the discharge under a union security agreement of dissenting employees who had tendered a portion of union dues, but who objected to paying the remainder of the dues which had been committed to non-collective bargaining purposes.[179] The Board stated that "[n]either on its face nor in the congressional purpose behind [section 8(a)(3)] can any warrant be found for making any distinction here between dues which may be allocated for collective-bargaining purposes and those earmarked for institutional expenses of the union."[180] Hence, full union dues could be required of dissenting employees under a union security agreement "so long as they are periodic and uniformly required and are not devoted to a purpose which would make their mandatory extraction otherwise inimical to public policy."[181]

The only deviation from this consistent line of rulings upholding the agency shop under section 8(a)(3) came in *Local 959, International Brotherhood of Teamsters*,[182] the Board's opinion cited by the Supreme Court in *Beck*.[183] In *Local 959* the Board stayed the union's attempts to compel payments for a savings and loan program and a building program. The payments were "assessments" which went into separate funds established by the union for a "special purpose," rather than into the union's general funds established to support and maintain the union. According to the Board, these payments did not constitute "periodic dues" within the meaning of section 8(a)(3).[184] The Board supported its finding that the monies were for a "special purpose" by asserting that they would be used "to accomplish ends not encompassed in its duties as a collective-bargaining agent of the employees."[185] The examination of the purpose to which the monies would be spent was justified by arguing that Con-

---

[178] 192 N.L.R.B. 951 (1971). *See also* Great Lakes Dist., Seafarers' Int'l Union (Tomlinson Fleet Corp.), 149 N.L.R.B. 1114, 1120 (1964).

[179] The dues were committed to the fraternal purposes of an old age pension and mortuary fund and to maintain a home for aged and infirm printers. *Detroit Mailers*, 192 N.L.R.B. at 951, 956.

[180] *Id.* at 952.

[181] *Id.*

[182] 167 N.L.R.B. 1042 (1967).

[183] *Beck*, 108 S. Ct. at 2652 n.7.

[184] *Local 959*, 167 N.L.R.B. at 1044.

[185] *Id.*

gress's purpose in allowing union security under the NLRA was to prevent non-members from free-riding with respect to collective bargaining expenses.[186] During the course of this argument the Board concluded: "It would thus appear that the right to charge 'periodic dues' granted unions by the proviso to Section 8(a)(3) is concerned exclusively with the concept that those enjoying the benefits of collective bargaining should bear their fair share of the cost incurred by the collective-bargaining agent in representing them."[187]

However, the Board's opinion in *Detroit Mailers* succeeded its opinion in *Local 959* and clearly repudiates any implication that section 8(a)(3) does not allow an agency shop. In *Local 959*, the Board described a congressional purpose for allowing union security agreements under section 8(a)(3) which is inconsistent with allowing agency shop agreements. The narrow holding of the opinion, however, was that the assessments in question were not "periodic dues" within the meaning of section 8(a)(3). The Board in *Detroit Mailers* distinguished its decision in *Local 959* on precisely these grounds, characterizing the case simply as one in which the union treasury never received the collected monies and in which even the union viewed the charges as an assessment rather than periodic dues.[188] Indeed, this may be the only basis on which *Local 959* can be consistently read, since it would seem counterfactual to disallow charges for the construction of union buildings on the grounds that these buildings would not be used for collective bargaining. When actually confronted in *Detroit Mailers* with the question of whether section 8(a)(3) allowed an agency shop agreement, the Board's response was an unequivocal "yes."

C. *The Legislative History of Section 8(a)(3) of the NLRA*

When the words of a statute are ambiguous, courts traditionally refer to legislative history to determine the legislature's intent.[189] The record of a bill's progress through the legislature,

---

[186] *Id.* at 1044–45 (quoting Radio Officers' Union v. NLRB, 347 U.S. 17, 41 (1954)).

[187] *Id.* at 1045. This was the language quoted in a footnote by the majority in *Beck*, 108 S. Ct. at 2652 n.7.

[188] Detroit Mailers Union No. 40 (Detroit Newspaper Publishers Ass'n), 192 N.L.R.B. 951, 952 (1971).

[189] Griffin v. Oceanic Contractors, Inc., 458 U.S. 564 (1982); 2A N. SINGER, *supra* note 134, § 48.01.

from committee hearings to final passage, is the best extrinsic evidence of the context in which the words of the statute evolved.[190]

1. The History, Effects, and Objectives of Union Security Agreements

In order to understand the legislative history of the NLRA with respect to union security argreements, it is necessary to understand both the law on union security agreements at the time of the enactment of the Wagner Act and unions' objectives in promoting the negotiation of such agreements.

Since their inception in this country, unions have sought to obtain union security agreements.[191] Under the common law, such agreements initially were held unlawful as being in furtherance of an unlawful conspiracy.[192] Once the courts accepted that unions were not unlawful conspiracies,[193] most jurisdictions held that union security agreements, including the closed, union, and agency shop, were lawful.[194] A few states enacted statutes outlawing specific forms of union security, in particular the closed shop, but most state legislatures remained silent on the subject.[195] The federal government took no role in regulating union security agreements outside of the railway industry prior

---

[190] *See* W. ESKRIDGE & P. FRICKEY, *supra* note 24, at 698–99, 709–10, 717–19. Of course, there is a traditional hierarchy of extrinsic evidence in statutory interpretation. Certain sources, such as committee reports or statements by the author, are given greater weight than other sources in the interpretation of the statute. *Id.*, 709, 717, 735.

[191] F. DULLES, LABOR IN AMERICA: A HISTORY 27 (3d ed. 1966).

[192] *See generally* Despres, *The Collective Agreement For The Union Shop*, 7 U. CHI. L. REV. 24, 31–33 (1939); Sayre, *Labor and the Courts*, 39 YALE L.J. 682, 695–97 (1930).

[193] *E.g.*, Commonwealth v. Hunt, 45 Mass. (4 Met.) 111, (1842).

[194] *See, e.g.*, Jacobs v. Cohen, 183 N.Y. 207, 76 N.E. 5 (1905); Parkinson v. Building Trades Council, 154 Cal. 581, 98 P. 1027 (1908); Kemp v. Division 241, Amalgamated Ass'n of Street and Elec. Ry. Employees, 255 Ill. 213, 99 N.E. 389 (1912); Local Branch No. 248, Nat'l Decorators Ass'n v. Solt, 8 Ohio App. 437 (1918); Gasaway v. Borderland Coal Corp., 278 F. 56 (7th Cir. 1921); Ribner v. Rasco Butter & Egg Co., 135 Misc. 616, 238 N.Y.S. 132 (1929); Harper v. Electrical Workers Local 520, 48 S.W.2d 1033 (Tex. Civ. App. 1932); Mississippi Theatres Corp. v. Hattiesburg Local Union No. 615, 174 Miss. 439, 164 So. 887 (1936). *See generally* 1 L. TELLER, THE LAW GOVERNING LABOR DISPUTES AND COLLECTIVE BARGAINING § 170 (1940); Despres, *supra* note 192, at 31–44; RESTATEMENT OF CONTRACTS § 515(c), illustration 18 (1932).

[195] In 1939, only five states (California, Colorado, Louisiana, Maryland, and Nevada) had statutes which prohibited compulsory union membership. Despres, *supra* note 192, at 54–55.

to 1935.[196] In 1933 Congress passed the National Industrial Recovery Act (NIRA).[197] Section 7(a) of the NIRA prohibited employer discrimination on the basis of union affiliation.[198] Some courts, apparently contrary to the intent of Congress,[199] interpreted this provision to prohibit the closed shop.[200] Thus, in the early 1930's the legality of union security agreements was governed by state law, with most states allowing union security agreements and a few errant courts striking down closed shop agreements under section 7(a) of the NIRA.

The unions' objectives in seeking union security agreements are evident from an examination of the effects of such agreements. All union security agreements require that, as a condition of employment, employees represented by the union provide support for the union either through membership and financial

---

[196] The Railway Labor Act of 1934 prohibited all union security agreements. Railway Labor Act of 1934, § 2 Fourth and Fifth, 48 Stat. 1185, 1187–88. However, the Act's prohibition on union security agreements was as much to prevent their use by the railways in preserving company unions, S. REP. No. 2262, 81st Cong., 2d Sess. 3, *cited in* Railway Employees Dep't. v. Hanson 351 U.S. 225, 231 (1956), as to protect dissenting employees' rights. *1934 Senate Hearings*, *supra* note 117 at 157, *reprinted in* 1 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 187; International Ass'n of Machinists v. Street, 367 U.S. 740, 752–53 (1961). In 1951 Congress enacted Section 2 Eleventh of the RLA to allow union security agreements in the railway industry. Act of January 10, 1951, Pub. L. No. 914, 64 Stat. 1238, (now codified at 45 U.S.C. § 152 Eleventh).

[197] National Industrial Recovery Act of 1933, Pub. L. No. 67, 48 Stat. 195. The NIRA was later declared unconstitutional by the Supreme Court in A.L.A. Schecter Poultry Corp. v. United States, 295 U.S. 495 (1935).

[198] Section 7(a) of the National Industrial Recovery Act provided as follows:

> Every code of fair competition, agreement, and license approved, prescribed, or issued under this title shall contain the following conditions: (1) That employees shall have the right to organize and bargain collectively through representatives of their own choosing, and shall be free from the interference, restraint, or coercion of employers of labor, or their agents, in the designation of such representatives or in self-organization or in other considered activities for the purpose of collective bargaining or other mutual aid or protection; (2) that no employee and no one seeking employment shall be required as a condition of employment to join any company union or to refrain from joining, organizing, or assisting a labor organization of his own choosing; and (3) that employers shall comply with the maximum hours of labor, minimum rates of pay, and other conditions of employment, approved or prescribed by the President.

National Industrial Recovery Act, § 7(a), 48 Stat. at 198–99.

[199] *Memorandum Comparing S. 1958, 74th Cong., 1st Sess., with Substitute S. 2926, 73d Cong., 2d Sess.* 29 (Mar. 11, 1935), [hereinafter *Comparison Memorandum*], *reprinted in* 1 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 1319, 1355.

[200] *See, e.g.*, Fryns v. Fair Lawn Fur Dressing Co., 114 N.J. Eq. 462, 168 A. 862 (Ch. 1933) (closed shop violates NIRA § 7(a) right to join rival union); Bayonne Textile Corp. v. American Fed. of Silk Workers, 114 N.J. Eq. 307, 168 A. 799 (Ch. 1933) (strike for closed shop is illegal), *rev'd*, 116 N.J. Eq. 146, 172 A. 551 (N.J. 1934). *Contra, e.g.*, De Agostina v. Parkshire Ridge Amusements, Inc., 155 Misc. 518, 278 N.Y.S. 622 (Sup. 1935) (closed shop lawful); Farulla v. Ralph A. Freundlich, Inc., 153 Misc. 738, 277 N.Y.S. 47 (Sup. 1934) (same).

contribution, or solely through financial contribution. Since the union determines the conditions of membership, the requirement of membership under a closed or union shop can provide the union with some discretion over the constituency of the employer's work force. Furthermore, the requirements imposed on employees by a closed shop, a union shop, or an agency shop increase the interchangeability of workers with respect to union support, as well as providing a source of financial support and perhaps loyalty for the union. Each effect can be examined separately with respect to its union objectives.

The union's discretion over who is a member and works in a closed or union shop allows the union to achieve a variety of objectives. Although the amount of discretion exercised by unions in excluding workers from membership has varied greatly, historically unions have used the closed or union shop to exclude workers who were careless or poorly trained,[201] who were disloyal to the union,[202] or who violated union rules by failing to pay dues, by undertaking a wildcat strike, or by crossing a picket line.[203] Unfortunately, some unions also used their membership rules to exclude workers on the basis of race, gender, religion, or nationality.[204] The closed shop was also used in some cases to control the supply of labor available to the employer in order to raise wages.[205]

The increase in the interchangeability of workers with respect to union support furthers at least three union objectives. First, the increase in interchangeability decreases the employer's incentive to discriminate on the basis of union affiliation. An employer who operates under a union security agreement has less incentive to fire an employee for union activity, since he knows he will just have to replace that worker with another union supporter. This reduction in employer incentive to discriminate protects both the individual employee's right of association and the integrity of the union as an organization from

---

[201] F. DULLES, *supra* note 191, at 27. Union members were particularly concerned that fellow employees undertake their work with the requisite skill and due care to do the job safely during the reign of the "fellow servant" rule prior to the advent of workers' compensation. 79 CONG. REC. 9732 (1935), *reprinted in* 2 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 3231–33 (extension of remarks by Rep. Beiter(D-N.Y.)).

[202] B. TAYLOR & F. WHITNEY, *supra* note 13, at 367–68.

[203] R. GORMAN, *supra* note 2, at.639.

[204] Steele v. Louisville & N. R.R., 323 U.S. 192 (1944); Oliphant v. Brotherhood of Locomotive Firemen, 156 F. Supp. 89 (N.D. Ohio 1957). This practice has since been outlawed by Title VII of the Civil Rights Act of 1964, 42 U.S.C. § 2000e-2 (1982).

[205] F. DULLES, *supra* note 191, at 63.

being undermined by the employer. Second, the increase in interchangeability decreases the resentment of workers who voluntarily join the union toward other workers who otherwise would not share the burdens of solidarity. By reducing the resentment of the voluntarv union members, the union security agreement removes an irritant and a potential obstacle to solidarity with those who do not join voluntarily. Finally, the increase in interchangeability prevents workers who do not voluntarily join the union from free-riding on the benefits of union representation that cannot be consumed exclusively by union members.[206] Workers who do not join the union voluntarily may in fact value the benefits of union representation very highly, but may not wish to pay for them since they know that other workers will support the union and will obtain the benefits of union representation for them anyway. Such free-riding saps the strength of the union and undercuts its ability to provide union benefits desired by the workers.[207]

Finally, financial support and potential loyalty resulting from union security agreements further several other union objectives. The resources provided by a union security agreement allow the union to do a better job in representing the employees: organizing, lobbying elected officials, supporting strikes, negotiating, and enforcing agreements. By doing a better job, the union can protect itself from being undermined by the employer or by rival unions, and can facilitate the organization of other employees. Also, to the extent that the union does a better job at the expense of the insular interests of the employer, the union security agreement shifts bargaining power from the employer to the employees. Thus, union security agreements provide unions with resources to maintain their activities, to grow, and perhaps to alter the balance of power in the workplace. All of these union objectives in pursuing union security agreements were well known among labor leaders, management representatives, and labor relations experts of the 1930's.[208]

---

[206] Such benefits would include benefits which the union must legally provide to all workers in the bargaining unit and public goods such as a safe and healthy workplace or comfortable environment which the union cannot supply to some workers without supplying to all.

[207] In the lexicon of economics, such a free-rider problem would result in the inefficiently low provision of union benefits. M. OLSON, THE LOGIC OF COLLECTIVE ACTION 15–16 (1971); H. VARIAN, INTERMEDIATE MICROECONOMICS 574–76 (1987).

[208] The unions' desire to achieve some control over the constituency of the workplace, whether to ensure safety, to preserve union discipline, to protect the union, or for any of the less venerable reasons, had long been recognized. Despres, *supra* note 192, at

2. The Wagner Act

Against this background, section 8(a)(3) of the NLRA was first enacted in 1935 as section 8(3) of the Wagner Act.[209] As previously stated,[210] this section contained a general prohibition against employer discrimination on the basis of union affiliation and the first proviso, which allowed the negotiation of union security agreements.[211] Neither the limitation on the first proviso—that union membership could not be a pre-condition to employment—nor the limitations of the second proviso were included. Section 8(3) contained only two explicit limitations on the negotiation of union security agreements: that such agreements were to be negotiated by a union designated as the exclusive bargaining representative of the employees under section 9(a) of the Act, and that the union could not be "established, maintained, or assisted" by the employer.[212] Because the first proviso refers only to agreements requiring "membership" as a condition of employment, it was later argued that this proviso did not allow the agency shop.[213] However, the Board and the Supreme Court interpreted section 8(3) of the Wagner Act to allow both the closed shop and all weaker forms of union se-

---

28–30. The protection of individual members' right to engage in union activity against employer impingement was perhaps the first purpose of union security agreements. Brooks, *Stability Versus Employee Free Choice*, 61 CORNELL L. REV. 344, 350 (1976). During the nineteenth and early twentieth centuries, union adherents referred to employees who worked in an organized shop but did not join the union as "rats," in reference to their unpleasant and parasitic nature, and sought union security agreements to avoid association and to protect their union from free-riding. Despres, *supra* note 192, at 27–38. Finally, historically both employees and employers understood that a union security agreement could foster the incumbent union and shift bargaining power from the employer to the employees. *Id.* at 28.

[209] National Labor Relations (Wagner) Act, ch. 372, § 8(3), Pub. L. No. 74-198, 49 Stat. 449, 452 (1935) (codified as amended at 29 U.S.C. § 158(a)(3) (1982)).

[210] *See supra* notes 39–42 and accompanying text.

[211] Section 8(3) of the Wagner Act provided that it would be an unfair labor practice for an employer:

> [B]y discrimination in regard to hire or tenure of employment or any term or condition of employment to encourage or discourage membership in any labor organization: *Provided*, That nothing in this Act . . . or in any other statute of the United States, shall preclude an employer from making an agreement with a labor organization (not established, maintained, or assisted by any action defined in this Act as an unfair labor practice) to require as a condition of employment membership therein, if such labor organization is the representative of the employees as provided in section 9(a), in the appropriate collective bargaining unit covered by such agreement when made.

National Labor Relations (Wagner) Act, *supra* note 209.

[212] *Id.*, §§ 8(2), 8(3).

[213] *See* NLRB v. General Motors Corp., 373 U.S. 734, 738 (1963).

curity not prohibited by state law, including the union shop and the agency shop.[214]

The broad interpretation of the extent of union security allowed under the Wagner Act was warranted by its legislative history. Both supporters and opponents of the Act understood it to allow the strongest form of union security—the closed shop.[215] The language of the statute and statements by supporters of the Act show that it also allowed weaker forms of union security.[216] In contravention of the general prohibition against employer discrimination based on union affiliation, Congress intended the first proviso to preserve the "status quo" with

---

[214]The prevailing administrative and judicial view under the Wagner Act was or came to be that the proviso to § 8(3) covered both the closed and union shop, as well as less onerous union-security arrangements, if they were otherwise legal. The National Labor Relations Board construed the proviso as shielding from an unfair labor practice charge less severe forms of union-security arrangements than the closed or the union shop, including an arrangement in *Public Service Co. of Colorado* [requiring nonunion members to pay an agency fee].

Id. at 739–40 (citing J.E. Pearce Contracting & Stevedoring Co., 20 N.L.R.B. 1061, 1070–73 (1940) (section 8(3) of Wagner Act allows agency shop)); M & J Tracy, Inc., 12 N.L.R.B. 916, 931–34 (1939) (section 8(3) allows arrangement giving hiring preference to union members)) (citations omitted). *See also* Colgate-Palmolive-Peet Co. v. NLRB, 338 U.S. 355, 361 (1949) (closed shop lawful under section 8(3) of the Wagner Act); Algoma Plywood & Veneer Co. v. Wisconsin Employment Relations Bd., 336 U.S. 301, 307 (1949) (Wagner Act did not preclude state regulation of union security agreements, but merely disclaimed former policy "hostile to the closed shop or other forms of union-security"); General Motors Corp., 133 N.L.R.B. 451 (1961) (section 8(a)(3) of NLRA allows agency shop agreements), *enforcement denied*, 303 F.2d 428 (6th Cir. 1962), *rev'd*, 373 U.S. 734 (1963); Public Service Co. of Colorado, 89 N.L.R.B. 418 (1950) (Wagner Act allows agency shop agreements).

[215] *See, e.g., National Labor Relations Board: Hearings on S. 1958 Before the Senate Comm. on Education and Labor*, 74th Cong., 1st Sess. 305 (1935) [hereinafter *1935 Senate Hearings*], *reprinted in* 1–2 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 1373, 1691; *id.* at 602–03, *reprinted in* 2 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 1988–89; 79 CONG. REC. 7673–74 (1935), *reprinted in* 2 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 2345, 2394–95; *Labor Disputes Act: Hearings on H.R. 6288 Before the House Comm. on Labor*, 74th Cong., 1st Sess. 61 (1935) [hereinafter *1935 House Hearings*], *reprinted in* 2 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 2473, 2535.

[216] The language of the proviso does not require that union membership be a *pre*condition to employment. Thus it allows at least the closed shop and the union shop. Furthermore, in commenting on the final bill, Senate counsel indicated in a memorandum that the exemption of the proviso was not limited to the closed or union shop:

> Unless this change is made as provided in S. 1958, most strikes for a closed shop or even for a preferential shop would by this act in effect be declared to be for an illegal purpose . . . . As the legislative history of [National Industrial Recovery Act] § 7(a) demonstrates, nothing in that section was intended to deprive labor of its existing right in many States to contract or strike for a closed or preferential shop . . . . No reason appears for a contrary view here.

Comparison Memorandum, supra note 199, at 29, *reprinted in* 1 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 1319, 1354–55.

respect to union security agreements.[217] Section 8(3) of the Wagner Act was not intended to outlaw any form of union security, but if a form was illegal under state law, section 8(3) would not pre-empt its prohibition.[218] Additionally, Congress intended section 8(3) to clear up the courts' uncertainty over whether section 7(a) of the NIRA outlawed the closed shop and to prevent such confusion under the general prohibition against employer discrimination in section 8(3). The proviso was drafted to preserve

---

[217] Responding to "misconceptions" that the proviso to section 8(3) of the Wagner Act would pre-empt state laws outlawing the closed shop, the Senate report on the bill stated:

> In other words, the bill does nothing to facilitate closed-shop agreements to or [sic] make them legal in any state where they may be illegal; it does not interfere with the *status quo* on this debatable subject but leaves the way open to such agreements as might now be legally consummated . . . .

S. REP. NO. 573, 74th Cong., 1st. Sess. 11–12 (1935), *reprinted in* 2 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 2300, 2311.

Similarly in debate on the proviso Senator Wagner (D-N.Y.) stated:

> The provision will not change the status quo. That is the law today; and wherever it is the law today that a closed-shop agreement can be made, it will continue to be the law. By this bill we do not change that situation.

79 CONG. REC. 7673 (1935), *reprinted in* 2 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 2345, 2395.

[218] The House report on the Wagner Act stated:

> The proviso to the third unfair labor practice, dealing with the making of closed-shop agreements, has been widely misrepresented. The proviso does not impose a closed shop on all industry; it does not give new legal sanctions to the closed shop. All that it does is to eliminate the doubts and misconstructions in regard to the effect of section 7(a) [of the National Industrial Recovery Act prohibiting employer discrimination with respect to union affiliation] upon closed-shop agreements, and the possible repetition of such doubts and misconstructions under this bill . . . . The bill does nothing to legalize the closed-shop agreement in the States where it may be illegal; but the committee is confident that it would not be the desire of Congress to enact a general ban upon closed-shop agreements in the States where they are legal.

H.R. REP. NO. 969, 74th Cong., 1st Sess. 17 (1935), *reprinted in* 2 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 2910, 2927. Similarly, in his major speech to the Senate in support of the bill, Senator Wagner said:

> While outlawing the organization that is interfered with by the employer, this bill does not establish the closed-shop or even encourage it. The much-discussed closed-shop proviso merely states that nothing in any Federal law shall be held to illegalize the confirmation of voluntary closed-shop agreements between employers and workers.

79 CONG. REC. 7570 (1935), *reprinted in* 2 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 2321, 2335. Representative Connery (D-Mass.), House sponsor of the bill, opposed an amendment offered by Representative Taber (R-N.Y.) to strike the proviso allowing union security agreements by alleging that it would allow 51% of the employees in any organization to bring about the discharge of the other 49%:

> Mr. Chairman, I merely rise to say this in opposition: The closed-shop proposition in this bill does not refer to any State which has any law forbidding the closed shop. It does not interfere with that in any way.

79 CONG. REC. 9726 (1935), *reprinted in* 2 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 3112, 3217.

agreements requiring "membership" against "this Act [and] . . . any other statute of the United States" to make it clear that the closed shop was not outlawed by either section 7(a) or section 8(3).[219] The desire to clarify that federal law allowed the closed shop seems to be the only significance of using the word "membership" in section 8(3).

Congress's determination to carve such an exception for union security agreements out of the general prohibition against employer discrimination is not surprising. Unions' objectives in seeking such agreements fulfill many of Congress's purposes in enacting the Wagner Act. Through the Wagner Act, Congress sought to increase employee control over the rules of the workplace, to protect employees' right to organize, to foster unions and collective bargaining, to promote stability in industrial relations, and to shift bargaining power from employers to employees.[220] By encouraging a stable system of collective bargaining with more power in the hands of employees, Congress hoped that employees would gain both a greater voice in the operation of their workplace and a greater share of the fruits of their labor, and that the nation as a whole would enjoy a greater measure of industrial peace and economic prosperity.[221] Union

---

[219] The Senate report on the bill states:

> The reason for the insertion of the proviso is as follows: According to some interpretations, the provision of section 7(a) of the National Industrial Recovery Act [48 Stat. 198], assuring the freedom of employees to organize and bargain collectively through representatives of their own choosing, was deemed to illegalize the closed shop. The Committee feels that this was not the intent of Congress when it wrote 7(a): that it is not the intent of Congress today; and that it is not desirable to interfere in this drastic way with the laws of the several States on this subject.

S. REP. NO. 573, 74th Cong., 1st Sess. 11 (1935), *reprinted in* 2 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 2300, 2311. Similarly the House report states with respect to the proviso in section 8(3) of the Wagner Act:

> All that it does is eliminate the doubts and misconstructions in regard to the effect of section 7(a) [of the NIRA prohibiting employer discrimination on the basis of union affiliation] upon closed-shop agreements, and the possible repetition of such doubts and misconstructions under this bill, by providing that nothing in the bill or in section 7(a) or in any other statute of the United States shall illegalize a closed-shop agreement between an employer and a labor organization . . . .

H.R. REP. NO. 969, 74th Cong., 1st Sess. 17 (1935), *reprinted in* 2 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 2910, 2927.

[220] National Labor Relations Act, § 1, 49 Stat. 449 (codified as amended at 29 U.S.C. § 151 (1982)). *See also* Keyserling, *The Wagner Act: Its Origin and Current Significance*, 29 GEO. WASH. L. REV. 199, 206, 215–18 (1960); Klare, *Judicial Deradicalization of the Wagner Act and the Origins of Modern Legal Consciousness, 1937-1941*, 62 MINN. L. REV. 265, 281–84 (1978); Mikva, *The Changing Role of the Wagner Act in the American Labor Movement*, 38 STAN. L. REV. 1123, 1126–27 (1986).

[221] Keyserling, *supra* note 220, at 218–24.

a longer and more eloquent statement for allowing union security agreements. In extended comments published in the Congressional Record on the day the House first passed the Wagner Act,[228] Representative Beiter argued that the closed shop allowed unions to exercise "collective responsibility" in meeting their obligations under collective bargaining to supply good, efficient, and safe workers.[229] The congressman also argued that the closed shop was necessary to protect individuals and unions from employer discrimination[230] and to promote stability in labor relations.[231] He then concluded that such stability in labor relations was necessary for industrial peace.[232] Finally,

---

[228] 79 CONG. REC. 9732–35 (1935), *reprinted in* 2 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 3228, 3232–34. The final version of the Act was agreed to in conference committee, without relevant changes, and passed by the House on June 27, 1935. 79 CONG. REC. 10300 (1935), *reprinted in* 2 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 3260, 3267.

[229] A well-organized union offers to supply all the labor an employer needs in a certain line. It proposes a contract covering wages, hours, and so forth. It is based on the principle of collective bargaining and, as a necessary corollary, collective responsibility. The union guarantees efficient and good work on the part of the employees. It cannot assume responsibility for outsiders having no control over them.

79 CONG. REC. 9732 (1935), *reprinted in* 2 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 3228, 3231.

[230] The closed shop is the only sure protection for trade agreements and for the defense of the individual. The open shop destroys organization, and in reality is the open door through which the union man goes out and the non-union man takes his place. The open shop means uncertainties, anxiety, and a shifting basis for the principles of industry. Under the open shop, the easy job goes to the non-union man, to the friend of the employer; the hard and dangerous task to the man whose devotion to his fellows incurs the enmity of the boss.

Id. at 9733, *reprinted in* 2 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 3233.

[231] [The principle of the closed shop] asserts that a shop cannot be half union and half non-union and, therefore, it asks the employer who is willing to recognize the union at all, and with it the principle of collective bargaining, to agree to employ none but union labor. The union shop, in other words, is to be closed to non-union workmen, not only in the interest of the contracting employees, but also in the interest of the employer . . . . In practice it was discovered that majority rule was best for employers as well as employees. Workers found it impossible to approach the employer in a friendly spirit if they remained divided among themselves. Employers likewise found it more satisfactory to confer voluntarily with a united and contented group of workers than with a group torn by internal dissention. Singleness of purpose and responsibility on each side gave to business transactions that stability which every employer desires.

Id. at 9732, *reprinted in* 2 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 3231–32.

[232] *Id.* at 9732–33, *reprinted in* 2 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 3232–33.

Representative Beiter also argued that it is the right of union members to contract with their employers to be free from association with people they considered "disloyal to . . . [their] class":

The right of every man to sell his labor as he sees fit is exactly the right on which the closed shop is based. The right to work and to contract for work includes the right to refuse to work except under certain conditions, and the nonemployment of certain classes of labor may very well be one of these

security agreements help to achieve all these purposes as well as to facilitate and to reinforce such a collective bargaining system.[222] Thus, it is reasonable to infer that to the extent union security agreements achieve the general purposes of the Wagner Act, these purposes were behind Congress's preservation of the status quo on union security agreements.

This inference is supported by the express statements of congressmen contained in the legislative history of the Wagner Act.[223] Senator Robert Wagner (D-N.Y.), in testimony before the Senate Education and Labor Committee, defended the proviso's allowance of the closed shop by attributing to "many experts" the belief that the closed shop[224] is "necessary, in some situations, to obtain for labor equality of bargaining power [with management] and substantial justice."[225] In later testimony before the Senate committee, the House Committee on Labor, and the full Senate, the Senator made a similar (perhaps conclusory) attribution that "many" believe that the closed shop "at times may be necessary to advance and preserve the living standards of employees."[226] Comments by Senator Wagner and the Senate Report on the bill suggest that the existing law on union security agreements was working well and that no adequate reason had been advanced to disrupt industrial relations and industrial peace by outlawing extant union security agreements.[227] Representative Alfred Beiter (D-N.Y.), a supporter of the bill, made

---

[222] *See supra* notes 201–208 and accompanying text.

[223] Supreme Court holdings also partially support this inference. *See, e.g.*, Colgate-Palmolive-Peet Co. v. NLRB, 338 U.S. 355, 364 (1949) ("It is quite reasonable to suppose that Congress thought it conducive to stability of labor relations that parties be required to live up to a valid closed-shop contract made voluntarily with the recognized bargaining representative . . . .").

[224] Comments contained in the legislative history most often state the purposes of allowing union security agreements with respect to the closed shop because that was the form of union security put at issue by the errant court interpretations of NIRA section 7(a). Nonetheless, the reasoning behind these statements applies to union security agreements in general.

[225] *1934 Senate Hearings*, *supra* note 117, at 9, *reprinted in* 1 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 39.

[226] *1935 Senate Hearings*, *supra* note 215, at 41, *reprinted in* 1 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 1417. Senator Wagner made similar comments during the Senate debate, 79 CONG. REC. 7570 (1935), *reprinted in* 2 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 2321, 2335, and in the House committee hearings, *1935 House Hearings*, *supra* note 215, at 16, *reprinted in* 2 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 2490.

[227] S. REP. No. 1184, 73d Cong., 2d Sess. 6 (1934), *reprinted in* 1 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 1099, 1105; *1934 Senate Hearings*, *supra* note 117, at 41, *reprinted in* 1 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 1417; *1935 House Hearings*, *supra* note 215, at 16, *reprinted in* 2 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 2490.

the House and Senate committees heard testimony from supporters of the bill arguing that union security agreements were necessary to allow unions to exercise responsibility,[233] to promote stability in labor relations,[234] to prevent employer discrimination,[235] to strengthen unions in order to increase employees' say in the operation of the workplace, and to maintain wages to prevent a worsening of the Great Depression.[236]

3. The Taft-Hartley Amendments

The NLRA was substantially amended in 1947 with the passage of the Taft-Hartley Act.[237] After renumbering section 8(3) as section 8(a)(3),[238] the Act amended the first proviso to limit the allowance of union security agreements to agreements which required membership as a condition of employment "on or after the thirtieth day following the beginning of such employment or the effective date of such agreement, whichever is the later . . . ."[239] A second proviso to section 8(a)(3) was added, prohib-

---

> conditions. The right of the non-union man is not infringed upon when the unionist merely refuses to work beside him or when he asks the employer to choose between them. As to the employer, he has the right to hire anyone he pleases, and he may discriminate at will against union and nonunion labor . . . . The reasons that appeal to a union man for not working with a nonunion man are manifest and obvious. Men instinctively love the society of their kind, whether in work or play, and the man who desires the society of his companions must arrange his life so that his associates are content to live with him.
>
> Trade unionists have for centuries believed that they were upholding the rights of men, protecting the welfare of their class, and promoting the interests of their homes; that without the union shop, their liberty and independence would be gone. This is not a fact of trade unionism alone, but a deep abiding fact in human life. In the last analysis, it is the law of self-defense; and employers have exactly the same feeling toward one of their members who gives his influence to the other side. Both feel that the offending man is disloyal to his class, and just so long as industry is carried on by two classes in hostile camps this feeling must and will continue.

Id. at 9732, *reprinted in* 2 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 3232.

[233] *1935 Senate Hearings*, *supra* note 215, at 121 (statement of William Green, President of the American Federation of Labor), *reprinted in* 1 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 1497.

[234] *Id.*

[235] *Id.* at 180 (statement of Professor H.A. Mills), *reprinted in* 1 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 1560.

[236] *1935 House Hearings*, *supra* note 215, at 86–87 (statement of William Dennison, representative of Society of Designing Engineers), *reprinted in* 2 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 2560–61.

[237] Labor Management Relations (Taft-Hartley) Act of 1947, ch. 120, § 101, 61 Stat. 136.

[238] The renumbering occurred because section 8(a) was used to list employer unfair labor practices while a separate section 8(b) was created to list union unfair labor practices. *Id.*, 61 Stat. at 140–42 (enacting 29 U.S.C. § 158(b)).

[239] *Id.*, 61 Stat. at 140–41 (codified as amended at 29 U.S.C. § 158(a)(3) (1982)).

iting employer discrimination for nonmembership in the union if the employer has "reasonable grounds" for believing that "membership was not available to the employee on the same terms and conditions generally applicable to other members" or "membership was denied or terminated for reasons other than the failure . . . to tender the periodic dues and the initiation fees uniformly required as a condition of . . . membership."[240] Finally, the Act added a requirement that a majority of employees in the bargaining unit had to approve a union security agreement in an election conducted by the Board before the union could negotiate such an agreement.[241] This requirement was repealed in 1951.[242]

The legislative history of the Taft-Hartley Act demonstrates that Congress understood that the new language in section 8(a)(3) prohibited the closed shop and limited the enforceability of the remaining lawful types of union security agreements. The limitation on the first proviso—that union security agreements could require membership only after thirty days of employment—was understood to prohibit the closed shop's pre-condition of union membership for employment.[243] However, the union shop and other lesser forms of union security were to remain lawful.[244] The second proviso was added to prohibit an

---

[240] *Id.*, 61 Stat. at 141.

[241] *Id.* Employees could rescind this approval in a new election at any time at least one year after the approval election. *Id.*, 61 Stat. at 145 (codified as amended at 29 U.S.C. § 159(e) (1982)).

[242] Act of Oct. 22, 1951, ch. 534, § 1(b), (c), 65 Stat. 601, 601–02. The requirement was repealed because unions won approximately 97% of these approval elections with an average vote of 77.5% in their favor. As a result, Congress saw these elections as a futile and expensive exercise. 1951 NLRB ANN. REP. 54 (1952); 1949 NLRB ANN. REP. 6 (1950). *See also* H.R. REP. No. 1082, 82d Cong., 1st Sess. 2–3 (1951). Current law retains such elections only as an option for rescission of approval. 29 U.S.C. § 159(e) (1982).

[243] The House Report on the Hartley bill states:

> The bill bans the closed shop. Under carefully drawn regulations it permits an employer and a union voluntarily to enter into an agreement requiring employees to become and remain members of the union a month or more after the employer hires them or after the agreement is signed.

H.R. REP. No. 245, 80th Cong., 1st. Sess. 9 (1947), *reprinted in* 1 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 292, 300. *See also id.* at 34; S. REP. No. 105, 80th Cong., 1st Sess. 3, 5–7 (1947), *reprinted in* 1 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 407, 409, 411–13; 93 CONG. REC. 5036, 5088 (1947) (remarks of Sen. Taft (R-Ohio)), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1421.

[244] The House Report on the Hartley bill stated that "the bill permits, subject to certain regulations and limitations, 'union security' agreements in the nature of union shops and maintenance of membership, but it bans the closed shop." H.R. REP. No. 245, 80th Cong., 1st Sess. 30 (1947), *reprinted in* 1 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 292, 321. *See also id.* at 34, *reprinted in* 1 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 325. The union security language of the Hartley Act varied somewhat

employer from discharging an employee under a lawful union security agreement if the union discriminatorily excluded the employee from membership or denied that membership for any reason other than failure to pay dues.[245] The legislative history suggests that employees who are discriminatorily excluded from the union are protected from discharge under the first clause of the second proviso even if they do not apply for membership or tender union dues.[246] However, the legislative history is ambiguous as to whether, under the second clause of the second proviso, the employee can be required to apply for union membership as well as pay dues.[247]

---

from the language which was finally passed in the Taft-Hartley Act, but these differences were irrelevant to the question of whether the statute allowed the union shop and lesser forms of union security. The Senate report on the Taft bill stated that the bill "abolishes the closed shop but permits voluntary agreements for requiring such forms of compulsory membership as the union shop or maintenance of membership . . . ." S. REP. NO. 105, 80th Cong., 1st Sess. 3 (1947), *reprinted in* 1 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 407, 409. *See also* H.R. CONF. REP. NO. 510, 80th Cong., 1st. Sess. 41, 44 (1947), *reprinted in* 1 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 505, 545, 548; 93 CONG. REC. 3950, 3952 (1947) (remarks of Sen. Taft), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1005, 1009; 93 CONG. REC. 5036, 5079 (1947) (remarks of Sen. Malone (R-Nev.) to the effect that "the proposal to outlaw the union shop has never been seriously considered by a majority of this body"), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1347, 1405.

[245] In the second place, we have proposed a proviso in the case where a man is refused admittance to a union, . . . In effect, we say, 'If you are going to have a union shop, then you must have an open union . . . .' The bill further provides that if the man is admitted to the union, and subsequently is fired from the union for any reason other than the nonpayment of dues, then the employer shall not be required to fire that man . . . . The employee has to pay the union dues. But on the other hand, if the union discriminates against him and fires him from the union, the employer shall not be required to fire him from the job.

93 CONG. REC. 3950, 3953 (1947) (remarks of Sen. Taft), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1005, 1010; *see also* S. REP. NO. 105, 80th Cong., 1st Sess. 20 (1947) (remarks of Sen. Ellender (D-La.)), *reprinted in* 1 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 407, 426; 93 CONG. REC. 4258 (1947), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1054, 1061–62; 93 CONG. REC. 4317–18 (1947) (remarks of Sen. Taft), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1090, 1096–97; 93 CONG. REC. 5087–88 (1947) (remarks of Sen. Taft), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1347, 1420–21.

[246] Senator Taft's explanation of the first clause of the second proviso does not seem to require that the employee who suffers from discrimination apply for union membership or pay union dues. 93 CONG. REC. 3953 (1947), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1005, 1010. However, the Board has interpreted the first clause of the proviso as not requiring payment of dues by employees who have been discriminatorily denied union membership. Union Starch & Refining Co., 87 N.L.R.B. 779, 784 (1949), *enforced*, 186 F.2d 1008 (7th Cir. 1951), *cert. denied*, 342 U.S. 815 (1951).

[247] The second proviso refers to the employee's membership being "denied or terminated," perhaps implying that the employee must apply for membership. 29 U.S.C. § 158(a)(3) (1982). The examples of the second proviso's application given by Senator Taft seem to contemplate application for membership. 93 CONG. REC. 3952–53 (1947), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1005, 1010. Represen-

Supporters of the Taft-Hartley bill understood that the dues which could be required of employees under the second proviso were the union's full periodic dues and initiation fees. In explaining the amendments to section 8(a)(3), Senator Robert Taft (R-Ohio) stated unequivocally that "[t]he employee has to pay the union dues."[248] Later, in further explanation of the workings of the second clause of the second proviso of section 8(a)(3), the Senator stated:

> The union could refuse a man admission to the union, or expel him from the union; but if he were willing to enter the union and pay *the same dues as other members* of the union, he could not be fired from his job because the union refused to take him.[249]

The same understanding is reflected in comments of other supporters of the bill and in House and Senate reports on the bill.[250]

---

tative Smith (D-Va.) also understood the Taft-Hartley bill to require an application for membership. 93 CONG. REC. A3141 (1947), *reprinted in* 1 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 906. *Cf.* 93 CONG. REC. 3614 (1947), *reprinted in* 1 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 669, 736 (remarks of Rep. Buck (R-N.Y.) that under the Hartley bill, which had somewhat different wording with respect to membership requirements, an employee must join the union and pay dues to keep his job). However, House and Senate reports on the bill and some comments by its supporters and detractors suggest that all that could be required under the proviso was the payment of dues. S. REP. NO. 105, 80th Cong., 1st Sess. 7 (1947), *reprinted in* 1 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 407, 413; H.R. CONF. REP. NO. 510, 80th Cong., 1st Sess. 41 (1947), *reprinted in* 1 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 505, 545; 93 CONG. REC. 3550 (1947) (remarks of Rep. McConnell (R-Pa.) with respect to the Hartley bill), *reprinted in* 1 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 601, 640.

[248] 93 CONG. REC. 3953 (1947), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1005, 1010.

[249] 93 CONG. REC. 4400 (1947), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1129, 1142. *See also* 93 CONG. REC. 4317–18 (1947), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1090, 1096–97, where Senator Taft four times refers to the money which the employee must pay under section 8(a)(3) as "union dues." The Senator also described the rule adopted under section 8(a)(3) as "substantially the rule now in effect in Canada." 93 CONG. REC. 5088 (1947), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1347, 1422. The Canadian "Rand Rule" required the payment of full union dues, but did not require membership application. Justice I.C. Rand of the Supreme Court of Canada set out the "Rand Rule" in an interest arbitration decision. Ford Motor Co., 1 Lab. Arb. (BNA) 439 (1946). At issue in the case was whether the collective bargaining agreement should contain a union shop clause and a dues check-off provision. *Id.* at 444. Basing his judgment on the "principles" held by the "large majority" of Canadians, Justice Rand declined to require a union shop (where employees must apply to or join the union), but did require that the employer check-off or deduct from each employee's wages "union dues," defined as "such sum as may from time to time be assessed by the union on its members according to its constitution, for general union purposes." *Id.* at 444–45. The decision thus required the payment of full periodic union dues by both members and nonmembers.

[250] In debate on the House floor, Representative Kersten (R-Wis.) stated "I also understand that [under] the provisions of the bill, . . . [employees] are merely required to pay reasonable dues which are required for the unions." 93 CONG. REC. 3615 (1947),

The dues and fees that could be required under section 8(a)(3) did not include special assessments for purposes not of general benefit to the employees or not part of the union's periodic dues.[251] However, Congress specifically rejected limitations on the amount a union could charge for periodic dues[252] and limited

---

reprinted in 1 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 669, 739. Representative Smith (D-Va.) understood the Taft-Hartley bill to provide that "the union cannot compel the employer to discharge an employee unless he refuses to join the union or maintain his membership in the union," presumably including full membership dues. 93 CONG. REC. A3141 (1947) (extension of remarks by Rep. Smith), *reprinted in* 1 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 906. Similarly, Senator Thye (R-Minn.) understood the bill to mean that "the employer can hire any man . . . , but after he has been in the plant 30 days he must become a qualified member of the union in order to remain on the payroll," indicating that he believed the employee must meet the qualifications of membership, i.e., paying full union dues. 93 CONG. REC. 5089 (1947), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1347, 1422. *See also* S. REP. NO. 105, 80th Cong., 1st Sess. 6–7 (1947), *reprinted in* 1 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 407, 413; H.R. CONF. REP. No. 510, 80th Cong., 1st Sess. 41 (1947), *reprinted in* 1 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 505, 545.

Moreover, prohibiting unions from collecting full union dues from dissenters would have violated the Eightieth Congress's desire for a simple, unobtrusive statutory scheme. Senator Taft described the amendments to section 8(a)(3) as merely "[making it] an unfair labor practice for a union to try to get an employer to discharge a man who has been improperly fired from the union." 93 CONG. REC. 3954 (1947), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1005, 1012. Senator Taft also stated:

> I believe . . . [the amended section 8(a)(3)] will permit the continuation of existing relationships, and will not violently tear apart a great many long-existing relationships and make trouble in the labor movement; and yet at the same time it will meet the abuses which exist.

93 CONG. REC. 5088 (1947), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1347, 1420. Senator Taft assured the Congress that the Act's limitations on union security did not interfere with the internal operations of the union or require the employer to inquire into internal union affairs in deciding whether to discharge an employee under a union security agreement. 93 CONG. REC. 4318 (1947) ("The pending measure does not propose any limitation with respect to the internal affairs of unions."), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1090, 1097. The Senate report on the Act stated that "[t]he tests provided by the amendment [to prohibit discharge under a union security agreement] are based upon facts readily ascertainable and do not require the employer to inquire into the internal affairs of the union." S. REP. No. 105, 80th Cong., 1st Sess. 20 (1947), *reprinted in* 1 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 407, 426.

[251] *See* Radio Officers' Union v. NLRB, 347 U.S. 17, 41 (1954) ("[The] legislative history clearly indicates that Congress intended to prevent utilization of union security agreements for any reason other than to compel payment of union dues and fees."); Local 959, Int'l Bhd. of Teamsters, 167 N.L.R.B. 1042, 1045 (1976) (special assessment cannot be collected under union security agreement). Among other problems, Congress sought to redress the experience of Cecil B. DeMille, who had been terminated from union membership and his job for refusing to contribute to a fund for a political cause which he opposed. 93 CONG. REC. 4528 (1947) (remarks of Sen. Ellender (D-La.)), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1054, 1061–62. Finally, Senator Taft stated that the rule adopted under section 8(a)(3) was "substantially" the rule in Canada. 93 CONG. REC. 5088 (1947), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1347, 1422. The Canadian "Rand Rule" did not include special assessments. Ford Motor Co., 1 Lab. Arb. (BNA) 439, 445 (1946).

[252] The Hartley bill, as originally passed by the House, prohibited initiation fees in excess of $25 unless the Board approved a greater amount as reasonable. It also

the use of dues only to the extent of prohibiting contributions to federal election campaigns.[253] Apparently, Congress was willing to trust union democracy to provide the necessary restraints on the size and uses of periodic union dues.[254]

---

prohibited "dues or general or special assessments that are not uniform upon the same class of members, or are in excess of such reasonable amounts as the members thereof, . . . by a majority of those voting, . . . shall authorize." H.R. 3020, 80th Cong., 1st Sess. § 8(c)(2) (1947), *reprinted in* 1 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 158, 179–80. However, the conference committee deleted these restrictions with the exception of section 8(b)(5)'s prohibition on a union charging excessive or discriminatory union initiation fees. H.R. CONF. REP. No. 510, 80th Cong., 1st Sess. 46 (1947), *reprinted in* 1 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 505, 550.

[253] Labor Management Relations Act of 1947, ch. 120, § 304, 61 Stat. 136, 159–60. This section of the Act amended § 313 of the Federal Corrupt Practices Act (1925), and was originally codified at 18 U.S.C. § 610. The latter was amended by the Federal Election Campaign Act of 1971, 86 Stat. 3 (1972), and then repealed by the Federal Election Campaign Act of 1976, 90 Stat. 490 (1976). Section 321 of the 1976 Act reenacted limitations on union and corporate political spending, which were codified at 2 U.S.C. § 441(b). This limitation has been held not to apply to union-controlled funds collected by voluntary contributions and kept segregated from union dues funds. Pipefitters Local 562 v. United States, 407 U.S. 385 (1972). *See also* First Nat'l Bank of Boston v. Bellotti, 435 U.S. 765 (1978); Buckley v. Valeo, 424 U.S. 1 (1976) (questioning the constitutionality of such limitations on campaign contributions); 1 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 571–72, 928; 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1526–35, 1603–04, 1609; Cantor, *supra* note 20, at 75.

[254] As discussed above, the House version of the bill originally included a limitation that union dues be "reasonable." *See supra* note 252. The Senate version of the bill contained no such limitation, and the Conference Committee rejected the limitation in the final version of the bill. Senator Taft explained the Senate conferees' reasons for refusing to accept this limitation:

> The Senate conferees refused to agree to the inclusion of this subsection in the conference agreement since they felt that it was unwise to authorize an agency of the Government to undertake such elaborate policing of the internal affairs of unions as this section contemplated without further study of the structure of unions. In the opinion of the Senate conferees the language which protected an employee from losing his job if a union expelled him for some reason other than nonpayment of dues and initiation fees, uniformly required of all members, was considered sufficient protection.

93 CONG. REC. 6601 (1947), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1526, 1540. There is a seeming contradiction in disallowing the charging of special assessments while setting no limitation on the amount of periodic dues and only one limitation on their use. Arguably, in a union the same democratic forces which would regulate the amount and uses of dues would also regulate the amount and uses of special assessments. However, Congress saw the special assessments as fraught with abuse. 93 CONG. REC. 4258 (1947) (remarks of Sen. Ellender (D-La.) concerning the case of Cecil B. DeMille), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1054, 1061–62. Although Congress might also have believed that a portion of periodic dues would be misused, the disallowance of special assessments but not periodic dues was in keeping with Congress's desire for an unobtrusive solution. Congress did not want to impose limitations on internal union affairs, 93 CONG. REC. 4318 (1947) (remarks of Sen. Taft), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1090, 1097, or to make unnecessary trouble for unions, 93 CONG. REC. 5088 (1947) (remarks of Sen. Taft), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1347, 1420. It thus made sense for Congress to prohibit the charging of special assessments to dissenting employees, thereby curing a serious problem with little intrusion into unions' internal affairs. It also made sense to refrain from prohibiting the collection of full union dues from dissenters, a remedy to a less serious problem that would involve a much greater intrusion into internal union affairs.

The Taft-Hartley amendments to section 8(a)(3) reflected the political compromises necessary to enact the legislation. The Eightieth Congress included many strong proponents of labor law reform. They sought to redress what they viewed as the excesses in union power allowed by the Wagner Act in unions' dealings with individual employees.[255] Some supporters of the bill would have liked to outlaw union security agreements altogether.[256] However, their zeal for amending the NLRA was tempered by the knowledge that a successful bill would need the support of a super-majority of Congress to survive a probable veto by President Truman.[257] As a result, the Taft-Hartley amendments to section 8(a)(3) reflected a compromise between those in Congress who opposed union security agreements and those congressmen who wanted no restraints on such agreements.[258]

---

[255] *E.g.*, 93 CONG. REC. 3538 (1947) (remarks of Rep. Hoffman (R-Mich.)), *reprinted in* 1 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 601, 622–23; *id.* at 3544 (remarks of Rep. Barden (D-N.C.)); *id.* at 3547 (remarks of Rep. Schwabe (R-Minn.)); 93 CONG. REC. 3951 (1947) (remarks of Sen. Taft), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1005, 1006–07; *id.* at 3953 (remarks of Sen. Taft). *See generally* Rosenthal, *The National Labor Relations Act and Compulsory Unionism*, 1954 WIS. L. REV. 53, 57–58. Congress also sought to redress what was viewed as excessive union power in bargaining with employers. *See generally* DEVELOPING LABOR LAW, *supra* note 160, at 35–36. However, this concern was addressed in other amendments prohibiting secondary boycotts and hot cargo provisions and not in the amendments to section 8(a)(3). *See* Labor Management Relations Act, ch. 120, § 8(b)(4), 61 Stat. 136, 141–42 (1947) (codified at 29 U.S.C. § 158(b)(4) (1982)); Labor-Management Reporting and Disclosure Act, § 704(b), 73 Stat. 519, 543–44 (1959) (codified at 29 U.S.C. § 158(e) (1982)).

[256] *E.g.*, 93 CONG. REC. 3612 (1947) (amendment by Rep. Hoffman (R-Mich.) to delete the first proviso of section 8(a)(3) allowing union security agreements), *reprinted in* 1 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 669, 732–33; 93 CONG. REC. 5087 (1947) (amendment by Sen. Ball (R-Minn.) to delete the first proviso of section 8(a)(3) allowing union security agreements), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1347, 1418.

[257] F. DULLES, *supra* note 191, at 359.

[258] Rosenthal, *supra* note 255, at 58. While deciding a secondary boycott issue, the Supreme Court in Local 1976, United Bhd. of Carpenters v. NLRB, 357 U.S. 93 (1958) stated:

> It is relevant to recall that the Taft-Hartley Act was, to a marked degree, the result of conflict and compromise between strong contending forces and deeply held views on the role of organized labor in the free economic life of the Nation and the appropriate balance to be struck between the uncontrolled power of management and labor to further their respective interests. This is relevant in that it counsels wariness in finding by construction a broad policy . . . when, from the words of the statute itself, it is clear that those interested in just such a condemnation were unable to secure its embodiment in enacted law. The problem raised by these cases affords a striking illustration of the importance of the truism that it is the business of Congress to declare policy and not this Court's.

Id. at 99–100 (*cited in* Beck v. Communications Workers of Am., 800 F.2d 1280, 1293 (4th Cir. 1986) (Winter, C.J., dissenting), *aff'd*, 108 S. Ct. 2641 (1988)).

The amendments to section 8(a)(3) sought to eliminate only what were perceived as the worst abuses of union security.[259] These abuses were the denial of employment to non-union employees under the closed shop and the loss of employment by workers whose union membership was denied or terminated for arbitrary or discriminatory reasons under a union shop.[260] They were eliminated by outlawing the closed shop and by limiting the enforceability of union shop agreements to cases in which the employee neither had been discriminatorily excluded from union membership nor had failed to fulfill the objective membership requirement of paying dues.[261] Neither the abuses discussed nor the solution proposed related in any way to the payment of union dues or to the enforceability of agency shop agreements.[262]

---

[259] "[T]hese amendments remedy the most serious abuses of compulsory union membership and yet give employers and unions who feel that such agreements promoted stability by eliminating 'free riders' the right to continue such arrangements." S. REP. No. 105, 80th Cong., 1st Sess. 7 (1947), *reprinted in* 1 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 407, 413. "[I]n this bill we are trying to be strictly practical and to meet the actual problems which have arisen, and not to go into the broader fields of the rights of particular persons." 93 CONG. REC. 5088 (1947) (remarks of Sen. Taft), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1347, 1421. "However, both the House and Senate bills correct the worst abuses." 93 CONG. REC. A2378 (1947) (extension of remarks by Sen. Ball (R-Minn.)), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1523, 1524. *See also* NLRB v. General Motors Corp., 373 U.S. 734, 740–41 (1963) (Taft-Hartley intended to correct only the most serious abuses); 93 CONG. REC. 3952–53 (1947) (remarks of Sen. Taft), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1005, 1010–11; 93 CONG. REC. 5087–88 (1947) (remarks of Sen. Taft), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, 1347, 1420–21.

[260] S. REP. No. 105, 80th Cong., 1st Sess. 6–7, 20 (1947), *reprinted in* 1 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 407, 412–13; 93 CONG. REC. 3952–53 (1947) (remarks of Sen. Taft), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1005, 1010–11; 93 CONG. REC. 4258–59, 4262 (1947) (remarks of Sen. Ellender (D-La.)), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1054, 1061–62, 1068.

[261] It was important to Congress that the conditions under which an employee could be discharged under a union security agreement be objective not only from the union's perspective to protect the employee from union discrimination, but also from the employer's perspective to protect the employer from liability under section 8(a)(3). S. REP. No. 105, 80th Cong., 1st Sess. 20 (1947), *reprinted in* 1 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 407, 426.

[262] Senator Taft and Senator Donnell (R-Mo.) made it clear that their primary concern was employees' ability to obtain and retain employment rather than their obligation to pay dues. When questioned by Senator Donnell as to why the union shop should remain lawful when the closed shop was being prohibited, Senator Taft responded, "the great difference is that in the [union shop] . . . a man can get a job without joining the union or asking favors of the union . . . . The fact that the employee will have to pay dues to the union seems to me to be much less important. The important thing is that the man will have the job." 93 CONG. REC. 5088 (1947), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1347, 1422. Senator Donnell responded, "I do not regard the payment of dues as the important point, at all." *Id.*

Congress rejected floor amendments that would have outlawed all forms of union security agreements.[263] A variety of purposes for preserving the legality of the union shop and the enforceability of the obligation to pay dues were offered.[264] Senator Taft defended the legality of the union shop. He argued that it was the "customary" form of union security, of long and widespread usage, that it was not fatally flawed like the closed shop, that his bill had remedied the closed shop's abuses, and that outlawing it would upset established relationships and cause industrial strife.[265] He defended the enforcement of the obligation to pay dues by arguing that it prevented employees from free-riding on union benefits.[266] The Senate Report on the bill stated that the Taft-Hartley amendments to section 8(a)(3) "remedy the most serious abuses of compulsory union membership and yet give employers and unions who feel that such agreements promote[] stability by eliminating 'free riders' the right to continue such arrangements."[267]

In the House, a number of representatives voiced support for union security agreements. Representative Buck (R-N.Y.) defended the union shop by arguing that "millions of men are working on union [shop] terms satisfactory to the men and satisfactory to the employer" and that prohibiting the union shop "would lead to complete chaos in labor relations."[268] Representatives Jennings (R-Tenn.) and Robsion (R-Ky.) raised the free-rider argument, asserting that it was "fair" to require all employees to pay dues to the union since all employees benefit from the union.[269] Representative MacKinnon (R-Minn.) stressed the importance of union security agreements to shifting

---

[263] Amendments were offered on both the House and Senate floors to delete the first proviso of section 8(a)(3), thereby outlawing all forms of union security agreements. *See supra* note 256.

[264] *Id.*

[265] 93 CONG. REC. 5087–88, 5089 (1947) (remarks of Sen. Taft), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, 1347, 1420, 1422; *see also* 93 CONG. REC. 3952–53 (1947) (remarks of Sen. Taft), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1005, 1010–11.

[266] 93 CONG. REC. 5089 (1947) (remarks of Sen. Taft), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1347, 1422; 93 CONG. REC. 3953 (1947) (remarks of Sen. Taft), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1005, 1010.

[267] S. REP. NO. 105, 80th Cong., 1st Sess. 7 (1947), *reprinted in* 1 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 407, 413.

[268] 93 CONG. REC. 3614 (1947) (remarks of Rep. Buck), *reprinted in* 1 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 669, 736.

[269] *Id.* at 3616–17, *reprinted in* 1 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 740–42.

bargaining power to unions so that they can "combat the large concentration of economic power that exists on the other side."[270] Representative Kersten (R-Wis.) opined that union security agreements were necessary in order for unions to be "effective."[271] Finally, Representative Brehm (R-Ohio) argued for the freedom of the union majority contractually to require all employees to support the union.[272]

The free-rider argument raised by the supporters of the bill was not limited in purpose to the recoupment of collective bargaining expenses. Although collective bargaining benefits were used as examples of the benefits on which employees could free-ride,[273] none of the statements limited the argument to such examples.[274] Such a limitation would have been contrary to Congress's understanding that the dues collectible under section 8(a)(3) were full union dues. Indeed, each of the congressmen who used collective bargaining benefits as examples of benefits which are subject to free-riding stated, shortly before or after his argument, that the dues that could be required under section 8(a)(3) were full union dues.[275] Congress was well aware that a portion of periodic union dues was commonly used for political and other non-collective bargaining purposes.[276] No rationale

---

[270] Representative MacKinnon stated: "I . . . believe that . . . we should permit and encourage voluntary [employer-union agreements for] union security, in order that the American worker may fairly combat the large concentration of economic power that exists on the other side." *Id.* at 3613.

[271] *Id.* at 3615.

[272] *Id.* at 3614.

[273] S. REP. No. 105, 80th Cong., 1st Sess. 7 (1947), *reprinted in* 1 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 407, 413; 93 CONG. REC. 3616 (1947) (remarks of Rep. Jennings), *reprinted in* 1 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 669, 740; *id.* at 3617 (remarks of Rep. Robsion); 93 CONG. REC. 5089 (1947) (remarks of Sen. Taft), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1347, 1422.

[274] *See supra* notes 262–272 and accompanying text. *See also* 93 CONG. REC. 3614 (1947) (remarks of Rep. Buck), *reprinted in* 1 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 669, 736; 93 CONG. REC. 3953 (1947) (remarks of Sen. Taft), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1005, 1010.

[275] Representative Jennings understood that dissenters would have to "contribute dues like the others." 93 CONG. REC. 3616 (1947), *reprinted in* 1 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 669, 740. Representative Robsion referred to the amount which the dissenters must pay as "union dues" and argued that it was fair that the dissenters contribute their "equal share" in securing union benefits. *Id.* at 3617, *reprinted in* 1 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 741. Senator Taft argued that under the provisions of the bill "a man can get a job . . . if, in effect, he joins the union and pays the union dues." 93 CONG. REC. 5088 (1947), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1347, 1421.

[276] *See, e.g.*, 93 CONG. REC. 6593–98 (1947) (Senate debate concerning union political contributions), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1526, 1526–35; Cantor, *supra* note 20, at 74. In fact, several legislators noted in the debates over the Taft-Hartley amendments that organized labor was spending considerable sums

for why a union could collect only collective bargaining expenses from dissenters was discussed.[277] Such a limitation constrains the free-rider argument well short of its full logical force. Employees can free-ride on any public good provided by the union, including political representation and organizing as well as collective bargaining services.[278]

4. The Purposes of Union Security Agreements Under the National Labor Relations Act: A Synthesis

To discuss the synthesis of Congress's purposes in preserving union security agreements under the Wagner and Taft-Hartley Acts as represented in the present NLRA, it is illuminating to place these purposes in the context of the previous discussion of the effects of union security agreements and the purposes of promoting such agreements.[279]

In the Wagner Act, Congress preserved all of the effects of union security agreements: union control over the constituency of the workplace, increased interchangeability of workers, and increased support (financial or otherwise) for unions. It made sense for Congress to preserve union security agreements because many of the unions' objectives in seeking such agreements supported Congress's general purposes in enacting the Wagner Act. The Act's legislative history suggests that Congress preserved union security agreements to achieve all, or almost all, of the laudable union objectives in seeking such agreements: increasing union control over the quality of the work force, decreasing employer incentive to discriminate against union supporters, avoiding resentment of union supporters for non-supporters, preventing free-riders, providing sufficient support for

---

to oppose the passage of the Act. 93 CONG. REC. 6605–06 (1947), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1526, 1549–51.

[277] Cantor, *supra* note 20, at 81–82; Gaebler, *Union Political Activity or Collective Bargaining?: First Amendment Limitations on the Uses of Union Shop Funds*, 14 U.C. DAVIS L. REV. 591, 603 (1981).

[278] It seems doubtful that in 1947 Congress envisioned the distinction the Supreme Court would draw in Ellis v. Brotherhood of Ry. Clerks, 466 U.S. 435, 447–48 (1984), that unions are entitled to compensation for collective bargaining expenses they incur in the course of performing their duties as exclusive representatives but not their other expenses. The doctrine of the duty of fair representation and the legal obligation of the exclusive representative to provide collective bargaining services to dissenters were not extended to the NLRA until 1955. Syres v. Oil Workers Int'l Union, Local 23, 350 U.S. 892 (1955) (per curiam) (reversing 223 F.2d 739 (5th Cir. 1955)).

[279] *See supra* notes 201–208 and accompanying text.

the union, shifting bargaining power from the employer to the employees, and promoting stability in labor relations.

In the Taft-Hartley Act, Congress sought to place limitations on union security agreements to eliminate union control over the constituency of the workplace. The abuses that the Taft-Hartley amendments were designed to eliminate were by-products of such control. By prohibiting the closed shop and by limiting the enforcement of the union shop to cases in which the union did not discriminate against the employee and the employee failed the objective requirement of paying union dues, Congress sought to remove union discretion over who did and did not work. Congress did not seek to constrain or eliminate the effectiveness of union security agreements in increasing the interchangeability of workers or providing support for the union. Neither the abuses at which the Act was aimed nor the solution devised by Congress were related to the agency shop or the obligation to pay dues.

Through the Taft-Hartley Act, Congress sought to amend the Wagner Act, not to repeal it. To the extent that Congress's purposes in enacting the Taft-Hartley Act conflict with its purposes in enacting the Wagner Act, the purposes of the Taft-Hartley Act must be given pre-eminence. By seeking to abolish union control over the constituency of the workplace, the Taft-Hartley Act countermanded Congress's prior purpose under the Wagner Act of allowing union control over the quality of the work force. However, in enacting the Taft-Hartley Act, Congress did not intend to abandon its prior purposes under the Wagner Act, which related to increased interchangeability of workers and support for unions. Thus, Congress's purposes under the Wagner Act of decreasing employer incentive to discriminate, preventing free-riders, providing sufficient support for unions, shifting bargaining power from the employer to the employees, and promoting stability in labor relations survive as purposes behind today's NLRA section 8(a)(3). Indeed, in response to proposed amendments to prohibit union security agreements, the advocates of the Taft-Hartley Act reiterated many of the same purposes for preserving union security agreements expressed by advocates of the Wagner Act.[280]

---

[280] The only argument in favor of union security agreements related to their effectiveness at increasing worker interchangeability or providing support to the union which was raised in the passage of the Wagner Act but not raised by the Congress which passed the Taft-Hartley Act was the argument that union security agreements decrease

In conclusion, there are many purposes behind Congress's limited preservation of union security under section 8(a)(3) of the NLRA. Most of Congress's purposes under the Wagner Act continue as purposes behind the present section 8(a)(3) of the NLRA. Far from limiting its purpose merely to preventing free-riding on collective bargaining benefits, Congress implicitly and explicitly stated purposes ranging from the lessening of incentive for employer discrimination to the promotion of stability in labor relations. Moreover, Congress's purpose of the prevention of free-riders was never limited to collective bargaining benefits, nor could it logically be so limited.

III. *BECK* AND THE DOCTRINE OF CONSTRUING STATUTES TO AVOID CONSTITUTIONAL QUESTIONS

The Supreme Court's interpretation of section 8(a)(3) varies greatly from the interpretation suggested by the words, administrative interpretations, and legislative history of the NLRA. In *Beck*, the Court drew an identity between section 8(a)(3) and section 2 Eleventh of the Railway Labor Act[281] and imposed on section 8(a)(3) the same limited interpretation of the purpose and enforceability of union security agreements as it had previously found under section 2 Eleventh. The Court held that Congress's sole purpose in preserving union security agreements under section 8(a)(3) was to allow the prevention of free-riding on collective bargaining benefits. This limited vision of statutory purpose was used to determine that only those expenses "necessarily or reasonably" related to collective bargaining can be charged to dissenting employees. The Court therefore held that section 8(a)(3) does not allow the compulsion of dues

employer incentive to discriminate on the basis of union affiliation. However, there is evidence that the Taft-Hartley Congress was aware of this argument and rejected the notion that it was no longer relevant to the purposes of section 8(a)(3). Arguing in favor of his amendment to prohibit all union security agreements, Senator Ball (R-Minn.) stated that the only purpose for allowing such agreements was to prevent employer discrimination and that section 8(a)(3)'s general prohibition on employer discrimination adequately solved this problem. 93 CONG. REC. 5087 (1947), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1347, 1419. Senator Ball's amendment was soundly defeated, 57 to 21. *Id.* at 5092. There is no evidence in the legislative history to suggest that Congress sought to limit the purpose of allowing union security agreements to prevent employer discrimination any more than necessary to eliminate union control over the constituency of the workplace.

[281] *See supra* notes 106–111 and accompanying text.

for non-collective bargaining purposes under an agency shop agreement.[282]

Direct examination of the words, administrative interpretations, and legislative history of section 8(a)(3) suggests a much broader interpretation of the purpose and enforceability of union security agreements. These sources establish that Congress intended section 8(a)(3) to allow full enforcement of an agency shop agreement. The legislative histories of the Wagner and Taft-Hartley Acts demonstrate that Congress preserved union security agreements under section 8(a)(3) because such agreements lessen employer incentive to discriminate on the basis of union affiliation, lessen animosity of union supporters towards non-supporters, prevent free-riding on public goods provided by unions, provide unions with resources necessary to be effective, shift bargaining power from the employer to the employees, and promote stability in labor relations.[283]

The Court's interpretation of section 8(a)(3) is inconsistent with the best evidence of congressional intent in the enactment of the NLRA. Why did the Court's reliance on its prior interpretations of section 2 Eleventh produce such a poor interpretation of section 8(a)(3)? There are several possible explanations.

The Court may have been mistaken that Congress intended an identity between section 8(a)(3) and section 2 Eleventh. However, this explanation appears doubtful. The similar language and legislative histories of the two statutes suggest that Congress did intend to allow the same forms of union security agreements under the two statutes.[284] Moreover, although when it enacted section 2 Eleventh Congress may not have understood all the purposes for union security agreements represented in section 8(a)(3), it has been argued persuasively that Congress understood section 8(a)(3) to allow agency shop agreements.[285]

Perhaps the Court's interpretation of section 8(a)(3) has gone awry because of its reliance on an indirect method of interpretation: analogizing to a similar statute rather than directly examining section 8(a)(3)'s words, administrative interpretations, and legislative history. However, the Court's interpretation

---

[282] *Beck*, 108 S. Ct. at 2652.

[283] *See supra* text accompanying notes 279–280.

[284] *Beck*, 108 S. Ct. at 2649 (quoting 96 CONG. REC. 17055 (1951) (remarks of Rep. Brown (R-Ohio))).

[285] Cantor, *supra* note 20, at 72–73.

seems to suffer less from reliance on second-best evidence than from a design to avoid a straightforward interpretation of section 2 Eleventh and consequently of section 8(a)(3). In interpreting the second of two similar statutes, one would expect that the Court would rely on a substantive analysis of its language and on consideration of the intrinsic and extrinsic evidence relevant to both statutes.[286] The Court's interpretation of section 2 Eleventh does not rest on such considerations. By the Court's own admission, its opinions interpreting section 2 Eleventh do not give merit to the plain language of the RLA.[287] Moreover, those opinions do not recognize an identity between section 2 Eleventh and section 8(a)(3), nor do they consider the interpretations or legislative history of section 8(a)(3) although section 8(a)(3) was passed four years before section 2 Eleventh.[288] Thus, the sole basis for the interpretation of section 2 Eleventh imposed on section 8(a)(3) in *Beck* is the legislative history of section 2 Eleventh.[289] This history chronicles events which occured four years after the passage of section 8(a)(3) and are arguably irrelevant to its interpretation.

To answer why the *Beck* opinion makes such a pronounced deviation from the congressional intent of section 8(a)(3), we must consider why the Court so narrowly constrained its consideration of intrinsic and extrinsic evidence in its interpretation of section 2 Eleventh. The Court may have just been mistaken in its interpretation of section 2 Eleventh and consequently of section 8(a)(3). It may have decided to construe these statutes according to its own predilections without proper deference to legislative intent. However, the most likely explanation is that the Court's interpretation of section 2 Eleventh was influenced by its desire to avoid the constitutional question of whether the negotiation and observance of agency shop agreements under the RLA violates dissenting employees' constitutional rights. Because the Court's interpretation of section 8(a)(3) relies so heavily on its prior interpretation of section 2 Eleventh, it also reflects these constitutional concerns.[290]

---

[286] *See supra* notes 134–135, 143–145 and accompanying text.

[287] See Ellis v. Brotherhood of Ry. Clerks, 466 U.S. 435, 445–46 (1984).

[288] *Ellis*, 466 U.S. 435; International Ass'n of Machinists v. Street, 367 U.S. 740 (1961); Railway Employees' Dep't v. Hanson, 351 U.S. 225 (1956).

[289] The Court does not have the benefit of administrative rulings on section 2 Eleventh since there is no equivalent of the NLRB under the RLA.

[290] Whether the *Beck* decision is a mistake, an abuse of power, or an application of the doctrine of avoiding constitutional questions, it should be noted that the case

A. *The* Beck *Decision as a Statutory Application of the Doctrine of Avoiding Constitutional Questions*

The Court has developed a doctrine of constitutional adjudication that where a "serious question" of a statute's constitutionality has been raised, the Court should, if possible, "fairly" construe the statute to avoid the constitutional question.[291] To

represents a continuation of the deradicalization of the NLRA. *See* Klare, *supra* note 220. For a further discussion on deradicalization, *see* Finkin, *Revisionism in Labor Law*, 43 MD. L. REV. 23 (1984); Klare, *Traditional Labor Law Scholarship and the Crisis of Collective Bargaining: A Reply to Professor Finkin*, 44 MD. L. REV. 731 (1985); Finkin, *Does Karl Klare Protest Too Much?*, 44 MD. L. REV. 1100 (1985); Klare, *Lost Opportunity: Some Concluding Thoughts of the Finkin Critique*, 44 MD. L. REV. 1111 (1985). As described by Klare, this deradicalization has occurred because, in its interpretation of the Wagner Act, the Court has repeatedly favored the purposes of the Act which are consistent with liberal capitalism and ignored the Act's more radical purposes. Klare, *supra* note 220, at 292–93. This process of deradicalization is present in the Court's emphasis on contractualism in collective bargaining, the development of the "public right" doctrine in the enforcement of the NLRA, and the limitation on the protection of employee concerted activity. *Id.* at 293. The limitation on the protection of employee concerted activity was achieved by separating the concept of union activity from the concept of employee activity and by limiting the extent of "legitimate" union activity which the Act would protect. *Id.* at 320–21. In *Beck*, the Court once again ignored Congress's more radical purposes, focusing only on a limited purpose of promoting collective bargaining. This limited congressional purpose was then used to resolve a conflict in which the activity of the union was conceived of as separate from that of the employees, with a resulting limitation on the extent of union activity allowed under the NLRA. The conflict between the preservation of the rights of the collective and the "privileging" of the rights of the individual is a central theme of Critical Legal Studies literature. *See generally* Kennedy, *Form and Substance in Private Law Adjudication*, 89 HARV. L. REV. 1685 (1976). The only distinctions between *Beck* and the examples cited by Klare are that in *Beck* the Court has directly limited the purposes of the NLRA it will consider rather than obscuring this limitation in a discussion of doctrine, and that in *Beck* the Court has undertaken the process of deradicalizing the Taft-Hartley Act as well as the Wagner Act.

[291] When the validity of an act of the Congress is drawn in question, and even if a serious doubt of constitutionality is raised, it is a cardinal principle that this Court will first ascertain whether a construction of the statute is fairly possible by which the question may be avoided.

Crowell v. Benson, 285 U.S. 22, 62 (1935); *see also* St. Martin Evangelical Lutheran Church v. South Dakota, 451 U.S. 772, 780 (1981); United States v. Rumley, 345 U.S. 41, 45 (1952) (Frankfurter, J.); Ashwander v. TVA, 297 U.S. 288, 348 (1936) (Brandeis, J., concurring); Lucas v. Alexander, 279 U.S. 573, 577 (1929); Richmond Screw Anchor Co. v. United States, 275 U.S. 331, 346 (1928); Blodgett v. Holden, 275 U.S. 142, 148 (1927) (Holmes, J., concurring). Some judges and commentators view this doctrine as a canon of statutory construction. *Ashwander*, 297 U.S. at 346–48 (Brandeis, J., concurring); W. ESKRIDGE & P. FRICKEY, *supra* note 24, at 676. However, the doctrine seems more properly viewed as a doctrine of constitutional adjudication, since to apply the doctrine the court must construe the Constitution at least to the extent of determining whether there is a serious constitutional question and how to avoid it. *See infra* note 330. Some respected justices and judges agree. *See* Textile Workers Union v. Lincoln Mills, 353 U.S. 448, 477 (1957) (Frankfurter, J., dissenting); *Rumley*, 345 U.S. 41, 45 (1953) (Frankfurter, J.); United States v. Lovett, 328 U.S. 303, 319–20 (1946) (Frankfurter, J., concurring); H. FRIENDLY, BENCHMARKS 210–11 (1967). Also, treating the doctrine as one of statutory interpretation rather than constitutional adjudication can result in the needless frustration of congressional intent and inequity in the treatment of people affected by different statutes. *See infra* note 309 and accompanying text.

limit this doctrine, the Court has admonished itself not to "press statutory constructions 'to the point of disingenuous evasion' even to avoid a constitutional question."[292]

Two rationales have been put forth to support this doctrine. First, where alternatives exist, Congress will not choose to enact a statute that raises a serious constitutional question for fear the statute will inadvertently trample constitutional rights and turn out to be merely a wasteful legislative exercise when it is struck down by the courts.[293] Second, the unelected judiciary should minimize its intrusion on the power of the elected Congress by declining to invoke its constitutional power of legislative review.[294] The first rationale sets up a presumption as to what Congress might have intended the statute to say. The second constitutes a directive, arguably contained in the Constitution, to choose an interpretation which avoids the constitutional question despite a certain amount of evidence suggesting that the statute has raised the question. If accepted, these rationales can justify either a modest application of the doctrine—that among equally likely interpretations of a statute, one of which raises the constitutional question and one of which does not, a court should select the one which avoids the question;[295] or a more radical application—that the doctrine effectively countermands some positive evidence which suggests an interpretation that raises the constitutional question.[296]

---

[292] United States v. Lock, 471 U.S. 84, 96 (1985) (quoting George Moore Ice Cream Co. v. Rose, 289 U.S. 373, 379 (1933)).

[293] W. ESKRIDGE & P. FRICKEY, *supra* note 24, at 676.

[294] Ashwander v. TVA, 297 U.S. 288, 346–48 (1936) (Brandeis, J., concurring); Posner, *supra* note 25, at 815.

[295] The cardinal principle of statutory construction is to save and not to destroy. We have repeatedly held that as between two possible interpretations of a statute, by one of which it would be unconstitutional and by the other valid, our plain duty is to adopt that which will save the act. *Even to avoid a serious doubt the rule is the same.*

NLRB v. Jones & Laughlin Steel Corp., 301 U.S. 1, 30 (1937) (emphasis added).

[296] "This rule of Constitutional adjudication is normally invoked to narrow what would otherwise be the natural but constitutionally dubious scope of the language." Textile Workers Union v. Lincoln Mills, 353 U.S. 448, 477 (1957) (Frankfurter, J., dissenting); *see also* Association of Salaried Employees v. Westinghouse Elec. Corp., 348 U.S. 437, 453 (1955), *overruled on other grounds*; Smith v. Evening News Ass'n, 371 U.S. 195 (1962); H. FRIENDLY, *supra* note 291, at 210. The doctrine of construing statutes to avoid constitutional questions should be distinguished from the doctrine of construing statutes, where possible, to avoid unconstitutionality. Under the doctrine of avoiding constitutional questions, the court determines whether a likely interpretation of the statute raises a serious constitutional question, and then fairly construes the statute to avoid the question. Under the doctrine of avoiding unconstitutionality, the court determines whether a likely interpretation of the statute violates the Constitution, and then

Despite its protests to the contrary,[297] there is strong evidence that the doctrine of avoiding constitutional questions lurks just below the surface of the court's interpretation of section 2 Eleventh in *Street*. The Court cited and discussed the doctrine at length as a preface to its interpretation of section 2 Eleventh.[298] There seems no doubt that the Court's interpretation of section 2 Eleventh varies greatly from the ordinary interpretation of that statute.[299] The Court itself has admitted that its interpretation of section 2 Eleventh is "not without difficulties,"[300] while individual justices and commentators have described the Court's interpretation as "strained" and even "tortured."[301] Moreover, the Court's interpretation, allowing variation in the amount of dues that can be charged employees based on individual dissent,

---

construes the statute so that it is constitutional. Posner, *supra* note 25, at 814. The benefits of actually deciding the constitutional question are discussed below. *See infra* note 325. The rationale of avoiding unconstitutionality is similar to the first rationale of avoiding constitutional questions. *Lincoln Mills*, 353 U.S. at 477 (Frankfurter, J., dissenting) (Congress would not intentionally enact an unconstitutional statute); Posner, *supra* note 25, at 814–15 (Congress prefers that courts not nullify its efforts). As with the doctrine of avoiding constitutional questions, this rationale will support either a modest "tip the scales" application or a more radical "countermanding" application. *See Jones & Laughlin Steel Corp.*, 301 U.S. at 30 (1937) ("tip the scales" application); United States v. Johnson, 323 U.S. 273, 276 (1944); H. FRIENDLY, *supra* note 291, at 210 (application where doctrine countermands intrinsic and extrinsic evidence).

[297] In *Street* the Court maintained that its interpretation of section 2 Eleventh was "reasonable," intimating that it was more than "fairly possible" and thus did not rely on the doctrine of avoiding constitutional questions. *Street*, 367 U.S. at 750. The Court repeated this claim in *Beck*. 108 S. Ct. at 2657.

[298] International Ass'n of Machinists v. Street, 367 U.S. 740, 749–50 (1961).

[299] The Supreme Court has recognized that, contrary to its interpretation, an ordinary interpretation of the language and legislative history of section 2 Eleventh suggests that it allows the full observance of agency shop agreements. Ellis v. Brotherhood of Ry. Clerks, 466 U.S. 435, 445 (1984).

[300] Abood v. Detroit Bd. of Educ., 431 U.S. 209, 232 (1977). The *Abood* Court stated that the *Street* interpretation of the Railway Labor Act was embraced "precisely to avoid facing the constitutional issues presented by the use of union shop dues for political and ideological purposes unrelated to collective bargaining." *Id.* at 232 (citations omitted).

[301] *Abood*, 431 U.S. at 248 (Powell, J., concurring). Cantor, *supra* note 20, at 67–68, 72; Cantor, *Forced Payments to Service Institutions and Constitutional Interests in Ideological Non-Association*, 36 RUTGERS L. REV. 3, 9 (1983); Read, *supra* note 57, at 256; Recent Development, *Union Shop Provision of the Railway Labor Act Held Not to Authorize Use of Union Dues for Political Purposes*, 61 COLUM. L. REV. 1513, 1517 (1961) [hereinafter Note, *Union Shop Provision*]. Justice Black, in his dissent in *Street*, said that "no one has suggested that the Court's statutory construction of [section] 2, Eleventh could possibly be supported without the crutch of its fear of unconstitutionality." *Street*, 367 U.S. at 786. Similarly, Justices Frankfurter and Harlan pointed out in their dissent in *Street* that the Court's interpretation of section 2 Eleventh deviated so far from an ordinary reading of the statute's language and legislative history that none of the parties to the case had urged that reading before the Court. *Id.* at 803. *See also Ellis*, 466 U.S. at 445–46; NLRB v. Catholic Bishop of Chicago, 440 U.S. 490, 500 (1979); *Street*, 367 U.S. at 786 (Black, J. dissenting).

smacks of constitutional remedy rather than statutory interpretation. Indeed, the Court's interpretation of section 2 Eleventh is very similar to the remedy that was found constitutionally required for the public sector in *Abood*.[302] It strains credulity to argue that this peculiar interpretation avoids the constitutional question merely by fortunate coincidence. Dissenting justices, commentators, and even the Court itself have recognized that the reason for the Court's strained interpretation of section 2 Eleventh was the Court's desire to avoid the constitutional question.[303]

The Court's imposition of the *Street* interpretation on section 8(a)(3) in *Beck* might still be justified if *Street* were an appropriate exercise of the doctrine of avoiding constitutional questions, and if there were important statutory reasons for interpreting sections 8(a)(3) and 2 Eleventh the same way. Regardless of the appropriateness of the Court's decision in *Street*,[304] there is no reason to interpret these sections identically. Although the legislative history of section 2 Eleventh suggests that Congress intended to allow the same forms of union security agreements under the RLA as were allowed under section 8(a)(3), Congress's only expressed purpose for this uniformity was extension of the benefits of a system that was working well under the NLRA.[305] Neither the Court nor the plaintiffs in *Beck* expressed any policy reason for interpreting section 8(a)(3) and section 2 Eleventh the same. Indeed, even after *Beck*, the extent of union security agreements allowed under the NLRA and the RLA differs in states that have prohibited union security agreements under section 14(b) of the NLRA,[306] since such prohibition is pre-empted for agreements governed by the RLA.[307] For the purpose of consistency it is desirable that similar statutes be

---

[302] *Abood*, 431 U.S. at 235–36.

[303] "*Street* embraced an interpretation of the Railway Labor Act not without its difficulties precisely to avoid facing the constitutional issues presented by the use of union shop dues for political and ideological purposes unrelated to collective bargaining." *Abood*, 431 U.S. at 232 (citations omitted); *see also* Ellis v. Brotherhood of Ry. Clerks, 466 U.S. at 445–46 (1984); NLRB v. Catholic Bishop of Chicago, 440 U.S. 490, 500 (1979); *Street*, 367 U.S. at 786 (Black, J., dissenting); Cantor, *supra* note 20, at 67–68; Note, *Union Shop Provision*, *supra* note 301, at 1517.

[304] The Court's opinion in *Street* is subject to many of the objections to the application of the doctrine of avoiding constitutional questions raised by *Beck*. *See* Cantor, *supra* note 20, 70–75; *see also infra* text accompanying notes 310–336.

[305] *Beck*, 108 S. Ct. at 2649 (quoting 96 CONG. REC. 17055 (1951) (remarks of Rep. Brown)).

[306] 29 U.S.C. § 164(b) (1982).

[307] 45 U.S.C. § 152 Eleventh (1982).

interpreted the same by the courts. However, where one statute's language and purpose has been attenuated due to constitutional concerns, such consistent interpretation serves only to frustrate legislative intent when these concerns are absent from the second statute.

The Court's imposition of the *Street* interpretation on section 8(a)(3) in *Beck* might also be justified if that interpretation were an appropriate exercise of the doctrine of avoiding constitutional questions with respect to section 8(a)(3). In fact, the Court's opinion in *Beck* is perhaps best understood as an application of the doctrine of avoiding constitutional questions as a doctrine of statutory interpretation rather than constitutional adjucation.[308] The Court's *constitutional* concerns about union security agreements under the RLA molded its interpretation of section 2 Eleventh in *Street*. That interpretation was then treated as a reasonable *statutory* interpretation and controlling extrinsic evidence in the interpretation of section 8(a)(3) in *Beck*. As evidenced in the *Beck* decision, the mistake in treating an application of the doctrine as one of statutory interpretation rather than constitutional adjudication is that there is no examination of whether the case at hand raises the same constitutional concerns.[309] Without such an examination, the Court may defeat the legislative intent of a statute without constitutional justification.

B. *Can* Beck *Be Justified as an Application of the Doctrine of Avoiding Constitutional Questions?*

To determine whether the *Beck* decision can be justified as an appropriate application of the doctrine of avoiding constitu-

[308] *See supra* note 291.

[309] Treating the doctrine as one of statutory interpretation also can cause inequitable treatment of people affected by different laws due to the dynamic nature of constitutional interpretation. W. ESKRIDGE & P. FRICKEY, *supra* note 24, at 687–88. For example, the Court may avoid a constitutional question by restrictively interpreting a statutory right or power, and then find no constitutional infirmity with the right or power when forced to answer the constitutional question under a different statute. If the interpretation of the first statute is viewed as statutory and thus does not evolve with the constitutionally-based decision, people affected by the first statute will be governed by a different constitutional standard than people affected by the second statute. Justice Black noted this problem in discussing the inequity in the different treatment afforded statutory union security rights and integrated state bars. *Street*, 387 U.S. at 785 (Black, J., dissenting). Apparently this inequity continues to this day. *See* Levine v. Heffernan, 864 F.2d 457 (7th Cir. 1989) (compelled bar membership does not violate the first amendment). One might argue that the aggrieved parties need only return to Congress and seek amendment of the relevant statute under the new constitutional standard. However, in most cases the problems and costs of returning to Congress to re-enact a negated statute will be significant or perhaps even prohibitive.

tional questions, we must first examine whether agency shop agreements under the NLRA raise a serious constitutional question.[310] Although this question seems similar to the question avoided in *Street*, the underlying constitutional concerns are quite different. In *Street* the existence of state action in the negotiation of the agency shop agreement had been established by the Court's decision in *Hanson*.[311] The underlying constitutional concern was whether, given state action, charging dissenting employees for political expenses violated their first amendment rights. The legitimacy of this concern was verified by the Court's decision in *Abood v. Detroit Board of Education*.[312] Under the NLRA, the constitutional concern is whether there is state action in the negotiation and observance of an agency shop agreement.[313]

There seems no serious question as to the constitutionality of agency shop agreements under the NLRA. Although the Court once may have entertained notions that a union's activities as exclusive representative constituted state action,[314] the Court's conception of state action has since narrowed so that this is no longer a real possibility. The Court now requires substantial grants of government authority or direct government coercion in promoting the activity in question to elevate a private party's actions to state action. No such grant of authority or coercion can be found in the NLRA.[315] *Hanson* cannot act as precedent on the question of state action under the NLRA because its finding of state action is based on section 2 Eleventh's pre-emption of contrary state laws.[316] The NLRA does not pre-empt state laws prohibiting union security agreements.[317] *Abood* cannot act as precedent for state action under the NLRA because it was a public sector case.[318] Even Justices Brennan and Marshall, who retain a broad vision of the state action doctrine,[319] seem to have abandoned the idea that the union's actions as

---

[310] *See infra* notes 337–339 and accompanying text.

[311] Railway Employees Dep't v. Hanson, 351 U.S. 225, 232 n.4 (1956); Wellington, *supra* note 57, at 354–59.

[312] 431 U.S. 209, 232–37 (1977).

[313] *See infra* notes 337–345 and accompanying text.

[314] Steele v. Louisville & N.R.R. Co., 323 U.S. 192, 198–99 (1944).

[315] *See infra* notes 388–399 and accompanying text.

[316] Railway Employees' Dep't. v. Hanson, 351 U.S. 225, 232 (1956).

[317] 29 U.S.C. § 164(b) (1982).

[318] *See infra* notes 340–343 and accompanying text.

[319] San Francisco Arts & Athletics v. United States Olympic Comm., 483 U.S. 522, 548–60 (1987) (Brennan, J., dissenting); Jackson v. Metropolitan Edison Co., 419 U.S. 345, 365–74 (1974) (Marshall, J., dissenting).

exclusive representative constitute state action.[320] With no real prospect for establishing the prerequisite state action, there is no serious question concerning the constitutionality of agency shop agreements under the NLRA.

The second question to consider is whether the Court's interpretation of section 8(a)(3) in *Beck* constitutes a "fair" interpretation or a "disingenuous evasion."[321] This inquiry is a subjective matter. However, it may be fruitful to examine the majority's interpretation in light of the rationales of the doctrine of avoiding constitutional questions.[322] If the Court's interpretation of section 8(a)(3) cannot be supported by these rationales, then it cannot be an appropriate application of the doctrine of avoiding constitutional questions.

It would indeed seem disingenuous to justify the Court's opinion on the basis of a general presumption that Congress avoids constitutional questions. The language and legislative history of section 8(a)(3) plainly allow the negotiation of an agency shop, overcoming any possible presumption. Moreover, the legislative history of the NLRA establishes that Congress enacted the statute with full knowledge that it raised many constitutional questions, including whether union security agreements violate dissenting employees' first amendment rights.[323] The passage of

---

[320] In Justice Brennan's opinion for the majority in United Steelworkers of Am. v. Weber, 443 U.S. 193 (1979), he stated in dicta that a union's collective agreement "does not involve state action." *Id.* at 200. Similarly, Justice Marshall wrote the Court's opinion in United Steelworkers of Am. v. Sadlowski, 457 U.S. 102 (1982), which held that a union's internal rules governing the procedures of its elections were not state action. *Id.* at 121 n.16.

[321] *See supra* note 292 and accompanying text.

[322] *See supra* notes 293–296 and accompanying text.

[323] The opponents of the Wagner Act raised a host of constitutional objections to its enactment. Their primary objection was that the regulation of labor relations exceeded Congress's power under the Commerce Clause. *1934 Senate Hearings*, *supra* note 117, at 390–94, *reprinted in* 1 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 424–48 (statement of James H. Emery, General Counsel of the National Association of Manufacturers). *See also 1935 Senate Hearings*, *supra* note 215, at 243–53, *reprinted in* 2 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 1629–39 (statement of Mr. Emery); *1934 Senate Hearings*, at 692–94, *reprinted in* 1 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 730–32 (statement of Wagner Fisher, an employer); NATIONAL LAWYERS' COMMITTEE OF THE AMERICAN LIBERTY LEAGUE, REPORT ON THE CONSTITUTIONALITY OF THE NATIONAL LABOR RELATIONS ACT (1935), *reprinted in Amendments to the National Labor Relations Act: Hearings Before the House Comm. on Labor*, 76th Cong., 1st Sess. 2242–45 (1939) [hereinafter REPORT ON THE CONSTITUTIONALITY OF THE NLRA]. This argument was based soundly on the precedents of the Court. *See, e.g.*, United Mine Workers v. Coronado Coal Co., 259 U.S. 344 (1922); Hammer v. Dagenhart, 247 U.S. 251 (1918). The opponents also argued that the Act violated (1) the employer's fifth amendment due process rights because the prohibited unfair labor practices were vague and interfered with the employer's freedom of contract in prohibiting anti-union discrimination; (2) the fourth, fifth, and seventh amendments, and article

the Wagner Act represented the pinnacle of an historic conflict between the legislative aspirations of the New Deal Congress and the constitutional interpretations of the pre-New Deal Court.[324] It would therefore seem inappropriate to interpret the NLRA on the presumption that Congress shied away from constitutional controversy in its enactment.[325]

III, by delegating responsibilities of the courts to the NLRB; and (3) the employee's first and fifth amendment rights by designating the union as the exclusive representative and allowing union security agreements. *1934 Senate Hearings*, *supra* note 117, at 397–400, *reprinted in* 1 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 431–34 (statement of Mr. Emery); *see also 1935 Senate Hearings*, *supra* note 117, at 244, *reprinted in* 2 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 1630 (statement of Mr. Emery); *1934 Senate Hearings*, *supra* note 117, at 690–91, *reprinted in* 1 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 728–29 (statement of Mr. Fisher); *1934 Senate Hearings*, *supra* note 117, at 762–64, *reprinted in* 1 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 800–02 (statement of Earnest T. Weir, Chairman of National Steel Corp.); 79 CONG. REC. at 7677–80 (1935) *reprinted in* 2 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 2403–11 (remarks of Sen. Hastings (R-Del.)); REPORT ON THE CONSTITUTIONALITY OF THE NLRA, *supra* at 2242–45. Senator Hastings commented that one did not need to be a "constitutional lawyer" but merely a "law student to reach the conclusion that the proposed act is unconstitutional." 79 CONG. REC. 7676 (1935), *reprinted in* 2 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 2403. These arguments also had some support in the decisions of the Court. *See, e.g.*, Adair v. United States, 208 U.S. 161 (1908) (statute outlawing employer discrimination against union members violates fifth amendment).

[324] Approximately one month prior to the final passage of the Wagner Act, the Supreme Court struck down the National Industrial Recovery Act of 1933, 48 Stat. 19, in A.L.A. Schechter Poultry Corp. v. United States, 295 U.S. 495 (1935). The Court believed that NIRA's codes of fair competition, which included the right to select exclusive representatives and to bargain collectively, exceeded Congress's power under the Commerce Clause. *Id.* at 548–50. Congress was then inundated with a "barrage" of letters and opinions from employers and their counsel that the Wagner Act was unconstitutional and should be abandoned. 79 CONG. REC. 8540 (1935), *reprinted in* 2 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 3011–12 (remarks of Rep. Connery (D-Mass.)). After the enactment of the Wagner Act, employer resistance and disregard for the Act was widespread because of its presumed unconstitutionality. Maden, *The Origin and Early History of the National Labor Relations Board*, 29 GEO. WASH. L. REV. 234, 242–46 (1960); DEVELOPING LABOR LAW, *supra* note 160, at 30–31. The Supreme Court finally ended the controversy in NLRB v. Jones & Laughlin Steel Corp., 301 U.S. 1 (1937), finding that the Wagner Act was a lawful exercise of Commerce Clause power without violating article III or employer fifth or seventh amendment rights. The *Jones & Laughlin* decision is recognized as the turning point in the Court's interpretation of congressional Commerce Clause power from its previously restrictive view. Stern, *The Commerce Clause and the National Economy*, 59 HARV. L. REV. 645, 674–85 (1946). The decision, along with its companion decision of similar issues under the RLA, Virginian Ry. Co. v. Sys. Fed'n No. 40, 300 U.S. 515 (1937), is significant also with respect to the delineation of employer fifth and seventh amendment rights.

[325] Judge Friendly has taken issue with the doctrine's rationale that Congress generally shies away from enacting statutes that raise constitutional questions. H. FRIENDLY, *supra* note 291, at 210. The Judge has argued that Congress has little reason to avoid questions of constitutionality since they are usually resolved in the Congress's favor and, even if they are not, the Court will salvage any constitutional portion of the statute by applying the doctrine of avoiding unconstitutionality. *Id.*; *see supra* note 296 and accompanying text. The Judge puzzles as to why Congress would avoid such a "heads-I-win, tails-you-lose" situation. H. FRIENDLY, *supra* note 291, at 210. One response is that Congress abhors forcing people to suffer the temporary deprivation of rights and

The Court's interpretation of section 8(a)(3) in *Beck* is also inconsistent with the doctrine's rationale of minimizing the Court's intrusion into the powers of the elected legislature. Traditionally, the Court describes the broad outlines of permissible legislation while Congress specifies a statute's precise policy and form.[326] When the Court interprets a statute to avoid a constitutional question, at least the initial determination of the law's policy and form is shifted away from Congress and to the Court. Where a possible constitutional problem admits to only one solution with little need of specification or where Congress can easily amend the statute, this initial determination poses a small intrusion into the domain of the legislature. In *Beck*, however, the Court chose one among several possible legislative solutions.[327] The solution chosen requires a great deal of specification as to what expenses a union may charge dissenters, the necessary union bookkeeping procedures, and the allowable fee reduction plans.[328] Moreover, given employers' desire and recent ability to resist reform of the NLRA,[329] the Court's choice of policy and form in this instance is likely to prevail for some time. Thus, the Court's opinion in *Beck* not only impinges on the role of the legislature, it usurps that role in its entirety.[330]

---

expense involved in redressing their constitutional rights through the courts. This argument, however, is insufficient to support the general presumption of avoidance. It seems just as likely to Judge Friendly that Congress would enact what it viewed as reasonable legislation and leave it to the courts to determine the finer points of constitutional law. *Id.* Judge Friendly argues that considering the constitutional question is "very likely just what Congress thinks the Justices are paid to do." *Id.* Indeed, it may be part of a legislative compromise to defer questions of constitutionality to the courts. Easterbrook, *Statutes' Domain*, 50 U. CHI. L. REV. 533, 544–46 (1983).

[326] *See* Hatch, *supra* note 145, at 894.

[327] In response to a finding of constitutional infirmity, Congress could have (1) limited the collection of dissenters' dues to collective bargaining expenses, as the Court did; (2) limited the collection of dissenters' dues to non-political and non-ideological expenses, the constitutional line for the compulsion of dues; (3) omitted the compulsion of political or ideological contributions from the exclusive domain of the union, thereby avoiding state action in the negotiation of union security agreements; or (4) required agency fee payments like those in solutions (1) or (2) for any bargaining units in which a majority of employees voted for a union security agreement. In fact there was some sentiment to require union security agreements upon a vote of the employees during the enactment of the Taft-Hartley Act. Amendment to S. 1126 by Sen. Malone, 93 CONG. REC. 5077 (1947), *reprinted in* 2 LMRA LEGISLATIVE HISTORY, *supra* note 112, at 1400 (requiring employers to adopt a union shop on three-fourths vote of employees).

[328] *See, e.g.*, Ellis v. Brotherhood of Ry. Clerks, 466 U.S. 435 (1984); Brotherhood of Ry. Clerks v. Allen, 373 U.S. 113 (1963).

[329] This ability is evidenced in the demise of the Labor Law Reform Act of 1977, H.R. 8410, 95th Cong., 1st Sess. (1977). *See Filibuster v. U.S. Labor Law Reform Bill*, 34 CONG. Q. ALMANAC 284 (1978); *see also* DEVELOPING LABOR LAW, *supra* note 160, at 66–67.

[330] A sound argument can be made that the doctrine's rationale of avoiding constitu-

The rationales of the doctrine also suggest that in interpreting the statute to avoid the constitutional question, the Court should do the least possible harm to the legislature's intent. The Court should not rely on a presumption of congressional avoidance of constitutional questions in the face of contrary expressions of legislative intent any further than necessary to avoid the constitutional question. Moreover, if the rationale of the doctrine is to minimize the incursion of the judiciary into the power of the legislative, that rationale itself requires a minimalist application. Properly applied, the doctrine requires more than a search for a superficially plausible interpretation which avoids constitutional questions; it requires an endeavor to give the fullest expression to legislative intent subject to avoidance of the constitutional question.

The Court's decision in *Beck* is not a proper application of the doctrine of avoiding constitutional questions because the Court could have interpreted section 9(a) of the NLRA to avoid the constitutional question without frustrating Congress's intent to allow the observance of agency shop agreements under section 8(a)(3). The Court could have concluded that although agency shop agreements are allowed by section 8(a)(3), the

---

tional questions to restrain judicial power is fallacious. The application of the doctrine does not avoid constitutional interpretation or exercise of judicial power. To apply the doctrine the Court must interpret the Constitution to determine whether a serious constitutional question exists. Justice Frankfurter and Judge Friendly, among others, perceived this, recognizing the doctrine as one of constitutional adjudication. H. FRIENDLY, *supra* note 291, at 211. Moreover, this determination is used in the same way as a determination of superficial unconstitutionality is used under the doctrine of avoiding unconstitutionality, *see supra* note 296, to choose among equally likely interpretations or to amend the interpretation away from unconstitutionality. The only saving afforded by the doctrine is an exact determination of the constitutional rights or limitations in question.

Judge Posner has pointed out that by retaining uncertainty as to what the constitution says, the application of the doctrine of avoiding constitutional questions creates a "judge-made 'penumbra'" around the Constitution, exaggerating its prohibitory effect and enlarging rather than diminishing the Court's exercise of judicial power. Posner, *supra* note 25, at 816. He recommends the abandonment of the doctrine in favor of application of the doctrine of avoiding unconstitutionality because the latter doctrine achieves the benefits of avoiding superficial unconstitutionality without expanding the prohibitory reach of the Constitution. *Id.*

However, it is the treatment of the doctrine as one of statutory interpretation rather than the penumbra of uncertainty which caused the needless erosion of unions' statutory rights in *Beck*. Had the Court in *Beck* recognized the application of the doctrine of avoiding constitutional questions in *Street*, treated the doctrine as one of constitutional adjudication, and examined whether it was appropriate to apply the doctrine in *Beck*, the Court would have decided against application of the doctrine and interpreted section 8(a)(3) to allow agency shop agreements. Thus, if the Court had properly acknowledged and applied the doctrine, Professor Posner's penumbra of uncertainty would never have come into play.

payment of fees for non-collective bargaining purposes as a condition of employment is not within the definition of "rates of pay, wages, hours of employment, or other conditions of employment" under section 9(a) of the NLRA.[331] As a result, while the negotiation of union security agreements would remain a mandatory subject of bargaining over which the union was the exclusive representative, the extension of such agreements to payments for non-collective bargaining purposes would be a permissive subject on which individual bargaining was possible.[332] Because the union would not be the exclusive representative with respect to the negotiation of fees for non-collective bargaining expenses, there would be no colorable argument of state action or constitutional violation. This interpretation of section 9(a) is at least as plausible as the Court's interpretation of section 8(a)(3) in *Beck*. The plain language of section 9(a) does not include such payments, and it takes no greater perversion of legislative history to exclude them from the coverage of section 9(a) than it does from the coverage of section 8(a)(3).[333] Indeed, unless the Court would require employers to bargain over agency shop agreements which cannot be observed or enforced under section 8(a)(3), its interpretation of section 8(a)(3) in *Beck* requires a similar interpretation of section 9(a).

---

[331] 29 U.S.C. § 159(a) (1982).

[332] The employer's duty not to bargain with individuals extends only to subjects delineated in section 9(a) of the NLRA. J.I. Case Co. v. NLRB, 321 U.S. 332, 339 (1944); Allied Chem. Co. & Alkali Workers of Am. v. Pittsburgh Plate Glass Co., 404 U.S. 157, 164 (1971); R. Gorman, *supra* note 2, at 380.

[333] The legislative history of section 9(a) reveals that the principle of exclusive representation was meant to encompass all subjects of collective bargaining, including union security. The Senate Report on the Act noted:

> Majority rule carries the clear implication that the employers shall not interfere with the practical application of the right of employees to bargain collectively through chosen representatives by bargaining with individuals or minority groups in their own behalf, after representatives have been picked by the majority to represent all. But majority rule, it must be noted, does not imply that any employee can be required to join a union, *except through the traditional method of a closed-shop agreement, made with the assent of the employer.*

S. REP. No. 573, 74th Cong., 1st Sess. 13 (1935), *reprinted in* 2 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 2300, 2313 (emphasis added); *see also* H.R. REP. No. 1147, 74th Cong., 1st Sess. 20–21 (1935), *reprinted in* 2 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 3046, 3070–71 (majoritarian principle of § 9(a) permits, but does not automatically establish, closed shop); *1935 House Hearings*, *supra* note 215, at 16–17, *reprinted in* 2 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 2490–91 (statement of Sen. Wagner (D-N.Y.) that closed shop may only be established by agreement between employer and § 9(a) representative). However, the legislative history is silent on the exact extent of the exclusive representative's authority on union security agreements, leaving room for the manuevers of a creative court.

The practical effect of this interpretation of section 9(a) on the legislative scheme for union security agreements under the NLRA would be miniscule: to avoid making the fine distinctions between collective and non-collective bargaining expenses and to avoid engaging in individual bargaining, private employers probably would choose to negotiate the traditional agency shop agreements envisioned by the Act.

Finally, the doctrine of avoiding constitutional questions should not be applied to avoid a question which will soon have to be answered in another case. The Court will soon be asked to decide whether union security agreements under the NLRA are subject to the same procedural protections of dissenters' rights that the Court found constitutionally mandated for the public sector in *Chicago Teachers Union, Local No. 1 v. Hudson.*[334] Since the language of section 8(a)(3) and its legislative history admit to no such procedural protections,[335] the Court should finally decide whether agency shop agreements under the NLRA are subject to constitutional scrutiny. Any delay in answering this constitutional question gained by the Court's decision in *Beck* seems pointless.

Thus, the Court's interpretation of section 8(a)(3) in *Beck* would not be an appropriate application of the doctrine of avoiding constitutional questions. There is no serious constitutional question to avoid and the Court's interpretation would seem to "[carry] the doctrine . . . to a wholly unjustifiable extreme"[336] because the rationales of the doctrine do not support the interpretation. Moreover, any avoidance of the constitutional question in *Beck* seems futile because the Court will soon have to address the same constitutional question in deciding whether to extend *Hudson* to union security agreements covered by the NLRA. As a result, the doctrine of avoiding constitutional questions cannot justify the Court's deviation in its interpretation of

---

[334] 475 U.S. 292 (1986).

[335] *See supra* text accompanying notes 148–278. To interpret such procedural restrictions into the language of section 8(a)(3) would seem particularly at odds with Congress's desire for unobtrusive regulation of union security agreements which does not impinge on union internal affairs. *See supra* notes 251, 254.

[336] International Ass'n of Machinists v. Street, 367 U.S. 740, 784 (1961) (Black, J. dissenting) (quoting Clay v. Sun Ins. Office Ltd., 363 U.S. 207, 213 (1960) (Black, J., dissenting)); Note, *Union Shop Provision, supra* note 301, at 1517; Wellington, *Machinists v. Street: Statutory Interpretation and the Avoidance of Constitutional Issues*, 1961 SUP. CT. REV. 49, 73. Although these sources refer to the Court's interpretation of section 2 Eleventh in *Street*, they are equally applicable to the extension of that interpretation to section 8(a)(3) in *Beck*.

section 8(a)(3) from Congress's intent, as represented in the statute's words, administrative interpretations, and legislative history.

IV. The Question Left Unanswered: Does the Negotiation and Voluntary Observance of Union Security Agreements Under the NLRA Constitute State Action?

As discussed at the outset of this Article, the *Beck* case involved two questions: whether section 8(a)(3) allows the negotiation and observance of agency shop agreements and, if so, whether such negotiation and observance violates dissenting employees' first amendment rights. The Court's determination that section 8(a)(3) did not allow agency shop agreements precluded the Court's consideration of the constitutional question. However, since I conclude in my analysis that the Court should have interpreted section 8(a)(3) to allow agency shop agreements, it is appropriate to proceed to an examination of the constitutional question.

In order to determine whether agency shop agreements under the NLRA violate dissenting employees' first amendment rights, one has to examine three questions. The first is whether the union's or employer's activity in negotiating and observing the agreement constitutes "state action."[337] If state action is found, the inquiry proceeds to the question of whether compulsory financial support for a union infringes upon dissenters' first amendment rights.[338] Finally, one must examine whether any state infringement which does occur can be justified by a compelling state interest and whether the state's activity is as narrow as possible to avoid the infringement of the dissenters' first amendment rights.[339] If there is insufficient state interest or

---

[337] *See* Hudgens v. NLRB, 424 U.S. 507, 513 (1976).

[338] This question is not as simple as it might first appear. Professor Cantor has analogized a union's use of dissenters' dues, even for political purposes, to the government's use of tax revenue for programs or the expression of ideas ideologically offensive to the taxpayer. Based on this analogy, Professor Cantor argues that even if there is state action in the negotiation and observance of an agency shop agreement, such an agreement does not infringe upon dissenting employees' first amendment rights. Cantor, *supra* note 20, at 70–71. *But see* Gaebler, *First Amendment Protection Against Government Compelled Expression and Association*, 23 B.C.L. Rev. 995, 1003–06 (1982).

[339] *See* L. Tribe, *supra* note 31, § 12-23.

unnecessarily broad state action, a first amendment violation has been established.

The Court has decided the last two of these questions in *Abood v. Detroit Board of Education*.[340] There the Court held that compulsory financial support of union activities does infringe upon dissenters' first amendment rights.[341] Although the Court found sufficient state interests to justify such infringement in the case of collective bargaining expenses, it found no such justification in the case of compulsory financial support of union political activities.[342] However, *Abood* was a public sector case, not governed by the NLRA, which raised no serious question as to state action.[343] Similarly, the existing precedent holding that the negotiation and observance of a union security agreement under the RLA constitutes state action is easily distinguishable from the problem posed under the NLRA.[344] Thus, the question that remains unanswered is whether the negotiation and observance of an agency shop agreement under the NLRA constitutes state action.[345]

---

[340] 431 U.S. 209 (1977).

[341] *Id.* at 234–35.

[342] *Id.* at 222, 234–36. The Court read Railway Employees' Dep't v. Hanson, 351 U.S. 225 (1956), as establishing that compulsory financial support of a union did infringe on dissenters' first amendment rights, yet was justified in the case of collective bargaining expenses by the state's interest in a system of labor relations based on collective bargaining. *Abood*, 431 U.S. at 222. This reading of *Hanson*, however, is far from obvious. An equally plausible reading of *Hanson* is that the Court found that the compulsory financial support of a union does not infringe on dissenters' first amendment rights. *See Hanson*, 351 U.S. at 238.

[343] In the public sector the government employer's agreement to the union security provision clearly satisfies the state action requirement. The issue was not even discussed in *Abood*.

[344] The basis of the Supreme Court's finding of state action under the RLA in *Hanson* was the exercise of the Supremacy Clause in section 2 Eleventh. *Hanson*, 351 U.S. at 232. There is no such exercise of the Supremacy Clause in section 8(a)(3) of the NLRA. 29 U.S.C. § 164(b) (1982).

[345] Another question which remains open is whether court enforcement of a union security agreement provides the state action necessary to give rise to a constitutional violation. *See* Shelley v. Kraemer, 334 U.S. 1 (1948) (judicial enforcement of racially restrictive covenant is state action). This question was not raised by the facts in *Beck*, since the employer had voluntarily complied with the union security agreement. The case was brought by dissenting employees who objected to this compliance. The fact that the vast majority of collective bargaining agreements rely in the first instance on private arbitration rather than on the courts for enforcement of the agreement's terms substantially decreases the importance of this question to determining the constitutionality of union security agreements. Moreover, the Court's opinion in *Shelley* has been subject to some criticism. *See, e.g.*, Wechsler, *Toward Neutral Principles of Constitutional Law*, 73 HARV. L. REV. 1, 29–31 (1959). One might question the continuing validity of the opinion in light of the Court's recent narrowing of the scope of the state action doctrine. *See infra* note 355 and accompanying text. In this Article I will concern myself only with the question raised in *Beck*—whether the negotiation and private observance of a union security agreement constitute state action.

A. *The State Action Doctrine*

1. State Action—An Historical Perspective

The first amendment, like almost all of the Constitution's guarantees of individual rights, protects the individual only against *government* infringement.[346] A finding that the complained of activity constitutes "state action" is a prerequisite to any first amendment claim.[347] The restriction of the Constitution's protection to government action is both confining and liberating for the individual. Although other private parties need not respect a person's constitutional rights, the private person himself is not bound by constitutional standards in his dealings with other people. Thus, "[c]areful adherence to the 'state action' requirement preserves an area of individual freedom by limiting the reach of federal law and federal judicial power."[348]

The concept of state action extends beyond the official acts of government officers and employees. The acts of an ostensibly private party may be state action if the party engages in a joint venture with the state or performs a public function with the authority of the state, or if the state coerces or significantly encourages the party's action.[349] The Supreme Court has not developed a unified test or doctrine to determine when ostensibly private activities constitute state action. Indeed, such a

---

[346] L. TRIBE, *supra* note 31, § 18-1.

[347] *Id.*

[348] Lugar v. Edmondson Oil Co., 457 U.S. 922, 936 (1982). Another benefit or purpose of the limitation of the Constitution's protections to government action is that this limitation reinforces federalism and the separation of powers. The requirement does this by limiting the range of wrongs the federal judiciary can redress in the absence of valid congressional legislation, thus creating a zone of action which is reserved to the states unencumbered by federal supremacy. L. TRIBE, *supra* note 31, § 18-2.

[349] Robinson v. Florida, 378 U.S. 153 (1964) (state encouragement found where private restaurant owners segregated their restaurants pursuant to a state statute requiring separate toilet facilities for blacks); Lombard v. Louisiana, 373 U.S. 267 (1963) (decision of private store owners to have sit-in demonstrators arrested for trespass pursuant to state encouragement to use trespass laws in a discriminatory manner constitutes state action); Peterson v. City of Greenville, 373 U.S. 244 (1963) (state coercion found where private restaurant owners segregated their restaurants pursuant to state law); Burton v. Wilmington Parking Auth., 365 U.S. 715 (1961) (state action found in refusal by private coffee shop to serve blacks, because coffee shop was leasing space in a state facility, making the state and coffee shop joint venturers); Marsh v. Alabama, 326 U.S. 501 (1946) (company town that carried out all the public functions of a municipality subject to constitutional restraints); Smith v. Allwright, 321 U.S. 649 (1944) (state political primary held to be a government function delegated by state to private parties and thus subject to constitutional restraints).

test is probably impossible.[350] "Only by sifting facts and weighing circumstances can the nonobvious involvement of the State in private conduct be attributed its true significance."[351] However, if the focus of the Court's inquiry is capable of summary, it seems to be "whether there is a sufficiently close nexus between the State and the challenged action of the [private party] so that the action of the latter may be fairly treated as that of the State itself."[352]

The Court's willingness to define a private party's activities as state action has undoubtedly changed over time. Prior to World War II, the concept of state action remained largely limited to actions taken by formal governmental actors—the legislature, the executive, and the judiciary.[353] After the war, the simultaneous increase in the power of private entities, in the welfare state, and in the concern for civil rights led the Vinson and Warren Courts to expand the concept of state action to include an ever larger area of previously private activity.[354]

---

[350] Professor Tribe asserts that a unified theory of state action is impossible under our current constitutional doctrine since the doctrine does not have an affirmative theory of individual liberty. Without an affirmative theory of individual liberty, he argues, it is impossible to distinguish when state inaction allowing private infringement of constitutional rights should be subject to constitutional standards. L. TRIBE, *supra* note 31, § 18-2. Professor Klare has argued that formulating any determinative test as to what constitutes state action is impossible since the distinction between public and private action is without determinative content. Klare, *supra* note 29, at 1415–21. The Court itself has acknowledged that "formulating an infallible test" of state action is "an 'impossible task.'" Reitman v. Mulkey, 387 U.S. 369, 378 (1967).

[351] Burton v. Wilmington Parking Auth., 365 U.S. at 722.

[352] Jackson v. Metropolitan Edison Co., 419 U.S. 345, 351 (1974).

[353] Phillips, *The Inevitable Incoherence of Modern State Action Doctrine*, 28 ST. LOUIS U.L.J. 683, 685–89 (1984). *See, e.g.*, The Civil Rights Cases, 109 U.S. 3 (1883) (fourteenth amendment only prohibits action taken directly by the state).

[354] Schneider, *State Action—Making Sense Out of Chaos—An Historical Approach*, 37 U. FLA. L. REV. 737, 739–43 (1985). *See, e.g.*, Smith v. Allwright, 321 U.S. 649 (1944) (resolution of state Democratic convention excluding blacks from the Democratic primary constitutes state action, *overruling* Grovey v. Townsend, 295 U.S. 45 (1935)); Marsh v. Alabama, 326 U.S. 501 (1946) (decision of company town to have religious proselytizer arrested under trespass law constitutes state action); Terry v. Adams, 345 U.S. 461 (1953) (exclusion of blacks from voting in primary elections of political association constitutes state action); Burton v. Wilmington Parking Auth., 365 U.S. 715 (1961) (refusal by private coffee shop leasing space in state parking facility to serve blacks constitutes state action); Peterson v. City of Greenville, 373 U.S. 244 (1963) (segregation of restaurants by private restaurant owners pursuant to state law constitutes state action); Lombard v. Louisiana, 373 U.S. 267 (1963) (decision of private store owners to have sit-in demonstrators arrested for trespass pursuant to state encouragement to use trespass laws in a discriminatory manner constitutes state action); Robinson v. Florida, 378 U.S. 153 (1964) (segregation of restaurants by private restaurant owners pursuant to state statute requiring separate toilet facilities for blacks constitutes state action); Amalgamated Food Employees' Union Local 590 v. Logan Valley Plaza, Inc., 391 U.S. 308 (1968) (decision of private shopping mall to have union picketers arrested under trespass laws constitutes state action).

This expansion of the state action doctrine into the realm of ostensibly private activity has suffered a sharp reversal under the Burger and Rehnquist Courts.[355] This reversal has been accomplished largely without any change in the expressed state action doctrine, since none of the Vinson and Warren Courts' expansive holdings has been overturned. Instead, the contraction has been achieved through a significant narrowing of the relevant activity the Court will consider in determining whether there is state action.[356] The Court will no longer accept arguments for state action where the state's influence or involvement in the discriminatory activity is indirect, but instead requires actual coercion or direct involvement of the government in the challenged decision or act.[357] Thus, under the current Court, a nexus of actual state coercion or direct state interjection into the specific discriminatory act must exist in order to scrutinize the actions of a private party as those of the state.

2. State Action and Unions—An Historical Perspective

The argument that a union's negotiation and observance of a collective agreement constitutes state action has been raised in several contexts, but each time it has been avoided by the Court.

---

[355] Schneider, *supra* note 354, at 739–43. The only major case since 1969 in which the Court has held that the actions of a private party were state action was Lugar v. Edmondson Oil Co., 457 U.S. 922 (1982) (invocation of state prejudgment attachment procedure and aid of state officers by a private creditor constitute state action). That case was explicitly limited to its facts. *Id.* at 939 n.21. With the exception of *Lugar*, the Court has consistently found that the actions of a private party are not state action. *See, e.g.*, NCAA v. Tarkanian, 109 S. Ct. 454 (1988) (decision by an unincorporated association whose members consisted of both public and private universities and colleges which resulted in suspension of basketball coach of a state university does not constitute state action); San Francisco Arts & Athletics, Inc. v. United States Olympic Comm., 483 U.S. 522 (1987) (decision not to allow use of word "Olympic" by federal corporation with federal grant to exclusive use of the word does not constitute state action); Blum v. Yaretsky, 457 U.S. 991 (1982) (decision to discharge or transfer Medicaid recipients by committees of private doctors required, reviewed, and supported by federal government does not constitute state action); Rendel-Baker v. Kohn, 457 U.S. 830 (1982) (decision to discharge teacher by private school subject to state regulation and supported by state does not constitute state action); Jackson v. Metropolitan Edison Co., 419 U.S. 345 (1974) (decision of state regulated public utility to discontinue service does not constitute state action); Flagg Bros., Inc. v. Brooks, 436 U.S. 149 (1978) (sale of goods by private warehousemen to recover debt pursuant to self-help provision of state law does not constitute state action); Moose Lodge No. 107 v. Irvis, 407 U.S. 163 (1972) (refusal to serve blacks by private dinner club with state liquor license does not constitute state action).

[356] Schneider, *The 1982 State Action Trilogy: Doctrinal Contraction, Confusion, And A Proposal For Change*, 60 NOTRE DAME L. REV. 1150, 1156–57 (1985).

[357] *Id.*

First, in *Steele v. Louisville & Nashville Railroad*[358] the plaintiffs contended that the negotiation of a racially discriminatory collective agreement by the exclusive representative under the RLA violated black employees' fifth amendment rights. Perhaps foreshadowing the union security cases, the Supreme Court avoided this constitutional question by interpreting the RLA as imposing a statutory duty on the union to represent all employees in the bargaining unit fairly.[359] This duty of fair representation was violated by such discriminatory activity. Next, in *Railway Employees' Department v. Hanson*[360] the plaintiffs argued that there was state action in the negotiation of union security agreements under the RLA. Although state action was found in *Hanson*, it was based on Congress's exercise of the Supremacy Clause in section 2 Eleventh of the RLA to pre-empt inconsistent state laws, not on the union's role in negotiating the collective agreement.[361] Lastly, the plaintiffs in *Oliphant v. Brotherhood of Locomotive Firemen and Enginemen*[362] argued that a union's discriminatory exclusion of blacks from membership violated their fifth amendment rights. Although this issue never reached the Supreme Court, two lower courts held that there was no state action and thus no constitutional infringement in a union's discriminatory denial of membership.[363]

Despite the Court's avoidance of the issue, it has suggested rather strongly in dicta that it would not find that the union's role in negotiating and observing the collective agreement constitutes state action. Although the issue was not raised by the parties to the dispute, in *United Steelworkers of America v. Weber*,[364] the Court volunteered the observation that a collective

---

[358] 323 U.S. 192 (1944).

[359] *Id.* at 198–99; *see also* Sayres v. Oil Workers Int'l Union, Local No. 23, 223 F.2d 739 (5th Cir. 1955), *rev'd*, 350 U.S. 892 (1955) (mem.) (parallel case to *Steele* dealing with a discriminatory collective agreement governed by the NLRA). This avoidance, to be sure, can be criticized on the basis that it expands, perhaps unevenly, the prohibitory effect of the Constitution. Yet the avoidance of the constitutional question in the case of the duty of fair representation seems more appropriate than in the case of union security agreements, since there seems to be no realistic legislative alternative and the question of whether a union's representation has been fair is one that is amenable to judicial determination.

[360] 351 U.S. 225 (1956).

[361] *Id.* at 232.

[362] 156 F. Supp. 89 (N.D. Ohio 1957), *aff'd*, 262 F.2d 359 (6th Cir. 1959), *cert. denied*, 359 U.S. 935 (1959).

[363] *Id.*; *but see* Betts v. Easley, 161 Kan. 459, 169 P.2d 831 (1946) (exclusive bargaining representative under RLA may not exclude members of the bargaining unit on the basis of race).

[364] 443 U.S. 193 (1979).

agreement "does not involve state action."[365] In *United Steelworkers of America v. Sadlowski*,[366] the Court, citing *Weber*, held that the union's internal rule prohibiting candidates for union office from accepting campaign contributions from nonmembers did not constitute state action.[367] Finally, in *Beck* both *Sadlowski* and *Weber* were cited in passing, although the resolution of the statutory question obviated the need to resolve the constitutional question. In a parenthetical phrase, *Weber* was cited for the proposition that "negotiation of a collective bargaining agreement's affirmative action plan does not involve state action."[368] If the Court really believes that *Weber* can be cited for the proposition that an affirmative action plan negotiated by an exclusive representative is not state action, then it would seem very difficult to argue successfully that a union security agreement negotiated by the exclusive representative is state action.[369]

The argument over whether a union's negotiation and observance of a collective agreement constitute state action is not merely of historical significance. In *Chicago Teachers Union, Local No. 1 v. Hudson*,[370] the Court held that in the public sector the protection of dissenters' first amendment rights requires that the union's procedure for accommodating dissenters under a union security agreement minimize the risk that dissenters' contributions might be used for impermissible purposes, provide adequate justification to dissenters for the remaining fee after any advance reduction in dues, and offer a reasonably prompt decision by an impartial decision-maker as to disputes over the allowable fee that can be charged dissenters.[371] It seems inevitable that some dissenting employees cov-

---

[365] *Id.*

[366] 457 U.S. 102 (1982), *reh'g denied*, 459 U.S. 899 (1982).

[367] *Id.* at 121 n.16. The issue of whether a union's internal rules covering its voluntary members constitute state action is arguably distinguishable from the issue of whether its activities under section 9(a) affecting dissenting employees constitute state action.

[368] *Beck*, 108 S. Ct. at 2657.

[369] *See also* Black v. Cutter Laboratories, 351 U.S. 292, 298–99 (1956) (just cause provision in a collective bargaining agreement allowing discharge for Communist Party affiliation does not raise federal issue); American Communications Ass'n v. Douds, 339 U.S. 382, 402 (1950) (dicta, "We do not suggest that labor unions which utilize the facilities of the National Labor Relations Board become government agencies or may be regulated as such."); *but see* Steele v. Louisville & N.R.R. Co., 323 U.S. 192, 208 (1944) (Murphy, J., concurring) (negotiation of discriminatory collective agreement by exclusive representative is state action).

[370] 475 U.S. 292 (1986).

[371] *Id.* at 304–09.

ered by a union security agreement under the NLRA will argue that without procedural safeguards equal to those outlined in *Hudson*, their interest in preventing the impermissible use of their contributions has not received adequate protection under the Constitution. The Court would then have several options: it could refuse to take the question, allowing possibly conflicting lower court opinions to stand; it could interpret section 8(a)(3) to include the procedural safeguards of *Hudson*, a feat from which I hope the Court would shrink;[372] or it could finally answer the question of whether there is sufficient state action in the negotiation and observance of a union security agreement under the NLRA to give rise to constitutional scrutiny.

There are two plausible arguments to support the contention that the negotiation and observance of a union security agreement by a union and private employer constitute state action. The first is that the union's designation as the exclusive bargaining representative under section 9(a) of the NLRA makes the union an arm of the state, subject to constitutional scrutiny in all of its activities as the exclusive representative. The second is that by promoting collective bargaining and shifting bargaining power to unions through various provisions of the NLRA, the federal government coerces or substantially encourages unions and employers to reach collective agreements containing terms favorable to unions—including union security agreements. I shall now consider each of these arguments.

B. *The Exclusive Representative as State Actor*

It is often asserted that the strongest argument for state action in the negotiation and observance of agency shop agreements lies in the designation of the union as the exclusive representative under section 9(a) of the NLRA.[373] When a union is elected the exclusive representative, it has the right to negotiate and enforce a collective bargaining agreement with the employer for all employees in the bargaining unit, including non-members

---

[372] Of course, the words of section 8(a)(3) say nothing of any such procedures. The legislative history shows that Congress never considered any such procedure. In fact, Congress desired an unobtrusive solution to the problem of dissenting employees that would not interfere with the internal affairs of unions. *See supra* notes 251, 254.

[373] *See, e.g.*, Fried, *Individual and Collective Rights in Work Relations: Reflections on the Current State of Labor Law and Its Prospects*, 51 U. CHI. L. REV. 1012, 1026 (1984).

and dissenters.[374] The employer, in turn, must bargain in good faith with the union to reach such an agreement.[375] The NLRA does not protect employee collective action which is independent of, or opposed to, the exclusive representative.[376] All individual employee contracts which are inconsistent with the collective agreement are superseded by it.[377]

Two arguments may be advanced as to how the designation of the union as the exclusive representative provides the necessary nexus between the state and the union's acts to make the latter state action. First, from the perspective of dissenting employees, the union is granted, over their opposition, a monopoly on the negotiation of their conditions of employment by operation of federal law.[378] It is argued that the exercise of this state grant of exclusive authority to bargain is state action. However, such a simple formulation of the problem ignores the fact that the union is elected by the employees and that even dissenters have an equal say in whether they are represented by a union.

A more sophisticated argument analogizes the elected union to an elected legislature, the employer to the executive, the negotiations between the union and the employer to the legislative process, and the collective agreement to legislation.[379] Under this scenario, unions and employers are performing a delegated government function in negotiating a collective agreement. Therefore, collective agreements—the "industrial legislation" of unions and employers—should be subject to constitutional restraints, just like the legislation of federal and state governments.[380]

---

[374] Steele v. Louisville & N.R.R., 323 U.S. 192 (1944).

[375] 29 U.S.C. § 158(a)(5), (d) (1982).

[376] Emporium Capwell Co. v. Western Addition Community Org., 420 U.S. 50 (1975) (black employees who picketed employer on their own, believing that their union did not adequately represent their interests, not protected by NLRA).

[377] J.I. Case Co. v. NLRB, 321 U.S. 332 (1944).

[378] Amicus Brief of Senators Jesse Helms (R-N.C.), Strom Thurmond (R-S.C.), Dan Quayle (R-Ind.), and Steven Symms (R-Idaho) at 16–17, *Beck*, 108 S. Ct. 2641.

[379] The Court itself has analogized collective bargaining to "industrial self-government." United Steelworkers v. Warrior & Gulf Navigation Co., 363 U.S. 574, 580 (1960). *See also* Steele v. Louisville & N.R.R. Co., 323 U.S. 192, 198 (1944) (making a similar analogy to find a statutory duty of fair representation under the RLA); NLRB v. Allis-Chalmers Mfg., 388 U.S. 175, 180 (1967) (quoting *Steele* and drawing a similar analogy for the purposes of statutory interpretation under the NLRA).

[380] Symposium, *Individual Rights in Industrial Self-Government—A "State Action" Analysis*, 63 NW. U.L. REV. 4, 8–19 (1968). *Cf.*, Blumrosen, *Group Interests in Labor Law*, 13 RUTGERS L. REV. 432, 482–83 (1959).

This analogy, however, is tenuous. Employers make peculiar executives, since they are unelected, unimpeachable and, if unincorporated, thoroughly private. Should employers really be held to constitutional standards? Amending the argument to hold only unions to constitutional standards not only fractures the analogy, but forces us to deal with the fact that the union is not entirely responsible for the results of collective bargaining. Unlike "real" government, the union does not always have the power to impose contract terms on the other party to its agreements. Indeed, it has been persuasively argued that under the current formulation of the NLRA, the employer's role in constructing the collective agreement exceeds that of the union.[381] Should the union be held responsible for unconstitutional contract terms sponsored by the employer? If not, how are we to separate which party is responsible for an offending term in the give-and-take of collective bargaining?[382] Furthermore, the process of collective bargaining is not analogous to the legislative process. A strike seems a poor analogy to a vote to override a veto, and no legislative analogue exists for an employer's threat to subcontract work or close the plant altogether. Thus, it would seem a mistake to elevate the metaphor of industrial legislation to the level of constitutional doctrine without deeper examination.

If we apply current Supreme Court doctrine to the problem, we find little hope that the union's status as the exclusive representative under section 9(a) would supply the necessary nexus between the state and the union's actions to make them state action. In *Jackson v. Metropolitan Edison Co.*,[383] the Supreme Court held that a state grant of monopoly status to a private utility company did not make the utility's decision to terminate service to customers state action.[384] Similarly, in *San Francisco*

[381] Stone, *The Post-War Paradigm in American Labor Law*, 90 YALE L.J. 1509, 1546–47 (1981); *see also* First Nat'l Maintenance Corp. v. NLRB, 452 U.S. 666, 676 (1981) (under the NLRA, Congress did not intend to make unions an "equal partner" with the employer in running the business); Schatzki, *Majority Rule, Exclusive Representation, and the Interests of Individual Workers: Should Exclusivity be Abolished?*, 123 U. PA. L. REV. 897, 901–02 (1975).

[382] For example, if the union asks for a grievance procedure, and the employer agrees to such a procedure on condition of a broad no-strike clause prohibiting job actions and demonstrations of any kind, could the no-strike clause be considered a violation of the first amendment perpetrated by the union?

[383] 419 U.S. 345 (1974).

[384] *Id.* at 351–52. Although the Court found some nebulous reasons to doubt the utility's monopoly status, it held that even if the utility did have a grant of pure monopoly, this was not determinative in considering whether the utility's decision to terminate service was state action. *Id.*

Arts & Athletics v. United States Olympic Committee[385] the Court held that the grant of a corporate charter or trademark did not make the actions of the recipients of such grants state action.[386]

One might try to distinguish *Jackson* by arguing that the grant of authority examined in that case is narrower than the grant of authority to a union under section 9(a). Arguably, the provision of gas and electricity has a less intimate effect on people's lives than the determination of their conditions of employment. However, there are some compelling counterarguments. The grant of authority to provide utility service is a powerful one, since a home, under modern conditions, is likely to become uninhabitable if service is denied.[387] The self-help remedy for individuals dissenting from the selection of a public utility is moving to a new town. In general, moving is more burdensome than finding a new job, the self-help remedy for individuals dissenting from the selection of an exclusive representative. Moreover, because of the union's elected status, the nexus between the state and the challenged activity under section 9(a) is weaker than in *Jackson*. Unlike employees under the NLRA, the utility customers in *Jackson* were not allowed to vote directly on the selection or retention of Metropolitan Edison as their public utility. Given the precedent of *Jackson* and the current disposition of the Court to define state action narrowly, it seems doubtful that state action could be established on the basis of the state's grant of exclusive authority under section 9(a).

Nor does current state action doctrine hold much promise for the argument that industrial self-government represents a delegated government function. The government-function argument under section 9(a) is based on an analogy between the process and product of collective bargaining and the process and product of state legislation. However, to determine what is a government function for the purposes of defining state action, the Court looks not at process and product, but instead at the subject matter of the function performed. The activities of a private party performing a function assigned by the state are considered state action only if the party performs a function which is "tra-

---

[385] 483 U.S. 522 (1987).

[386] *Id.* at 543–44. *See also* Moose Lodge No. 107 v. Irvis, 407 U.S. 163, 176–77 (1972) (a private club's discrimination against blacks not state action despite the fact that the state had granted the club one of a limited number of liquor licenses).

[387] *Jackson*, 353 U.S. at 361 (Douglas, J., dissenting).

ditionally the exclusive prerogative of the State."[388] The Supreme Court has also pointed out that "[w]hile many functions have been traditionally performed by governments, very few have been 'exclusively reserved to the State.'"[389] Although the state has sometimes passed laws regulating minimum wages, maximum hours, and working conditions,[390] the determination of the terms of employment among private employers and employees has never been its exclusive prerogative. Traditionally, the determination of private terms of employment has been left to the private parties.[391] Thus, under section 9(a) of the NLRA, unions exercise a private function rather than a delegated government function which would constitute state action.[392]

The failure, under existing precedent, of the state action argument based on the union's designation as the exclusive representative is consistent with the state action doctrine's purpose of preserving the Constitution's balance of individual freedom. Certainly an unbounded application of *Jackson*, exempting all authorizations of private power in areas which are not the traditional prerogative of the state, would be wrong. When the state assigns a monopoly to a private party, it substitutes regu-

---

[388] Blum v. Yaretsky, 457 U.S. 991, 1005 (1982); Rendell-Baker v. Kohn, 457 U.S. 830, 842 (1982); Flagg Bros., Inc. v. Brooks, 436 U.S. 149, 157–61 (1978); *Jackson*, 419 U.S. at 353.

[389] *Flagg Bros.*, 436 U.S. at 158.

[390] *E.g.*, Fair Labor Standards Act of 1938, 52 Stat. 1060 (current version codified at 29 U.S.C. §§ 201–19 (1982)) (regulating minimum wage); Occupational Safety and Health Act of 1970, 84 Stat. 1590 (current version codified at 29 U.S.C. §§ 651-78) (regulating workplace safety).

[391] Even if we assume that the union's grant of exclusive bargaining authority under section 9(a) makes the union a state actor, a centuries-old tradition of private determination of the terms of employment in this country preceded the passage of the NLRA. This tradition has continued after the passage of the NLRA in the unorganized bulk of the private sector.

[392] This analysis brings to the forefront what is really the most interesting constitutional question posed by section 9(a), not whether the state's assignment to the union of the private function of bargaining makes the union a state actor, but whether the state *can* reassign the right to bargain from the private employees to the union without infringing upon the employees' first amendment, fifth amendment, or freedom of contract rights. Although the Court has never addressed this issue from the perspective of dissenting employees' constitutional rights, the Court has found no violation of the employer's fifth amendment and freedom of contract rights in the designation of the union as the exclusive representative under section 9(a). NRLB v. Jones & Laughlin Steel Corp., 301 U.S. 1, 44–45 (1937); Virginian Ry. Co. v. System Fed'n No. 40, 300 U.S. 515, 557–59 (1937). There have been some lower court opinions upholding the concept of exclusive representation over the constitutional objections of dissenting employees. *See, e.g.*, Local 858 American Fed'n of Teachers v. School Dist. No. 1, 314 F. Supp. 1069 (D. Colo. 1970). *See generally* Zwerdling, *The Liberation of Public Employees: Union Security in the Public Sector*, 17 B.C. INDUS. COM. L. REV. 993, 1001–02 nn.57–58 (1976).

lation for whatever discipline on individual behavior the market might impose. The regulation may increase or decrease the individual freedom of the monopolist or its customers over what the market may provide. Although the results of current or past markets are not constitutionally protected,[393] if the state unilaterally assigned a complete and unfettered monopoly over an aspect of some importance to people's lives, one must wonder whether the Constitution would require that this monopoly be subject to constitutional constraints.[394] On the other hand, almost every law involves some grant of authority to a private party which affects the individual freedom of the party or the people with whom it deals.[395] Unless every law is to create a state actor, some rule or line must be constructed to distinguish when the exercise of a grant of authority to a private party constitutes state action and when it does not. The construction of such a general rule is beyond the scope of this Article. However, I would argue that the union's designation as exclusive representative is not such an extensive grant of authority that the union's acts should be subject to constitutional scrutiny.

The grant of authority to the exclusive representative is far from that of a unilaterally imposed, unfettered monopoly. In fact, one could argue that the union's designation under section 9(a) is not a grant of authority at all. Long before the passage of the NLRA, unions bargained for and received enforceable agreements of employer recognition that they were the employees' exclusive bargaining representatives.[396] At times such recognition was achieved without majority support of the represented employees or after a strike. By enacting the election and exclusive representation provisions of section 9 and by prohibiting recognition strikes in section 8(b)(7), Congress did not create a new status; it created only a fairer and more peaceful

---

[393] "[A] constitution is not intended to embody a particular economic theory, whether of paternalism and the organic relation of the citizen to the State, or of *laissez faire*." Lochner v. New York, 198 U.S. 45, 75 (Holmes, J., dissenting).

[394] For example, if the NLRA required that all employees be represented by worker committees consisting of private employees designated by the government in each shop and that these committees have complete and unilateral power over the setting of wages, hours, and working conditions in the shop, I would think that there would be a good argument that the acts of these worker committees would be state action. Such is not the case under the current NLRA.

[395] For example, in labor law the simple rule that collective agreements are enforceable grants unions authority in their relationships with employers.

[396] S. REP. NO. 573, 74th Cong., 1st Sess. 11–12 (1935), *reprinted in* 2 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 2300, 2311–12 (1949).

system for allowing unions to achieve the status of exclusive representative.[397] The election process, in which each employee has an equal say, decreases any grant of authority or state nexus by interjecting the employees' private decision as to whether to select or retain the union as their exclusive representative. Finally, the union's authority as exclusive representative has significant limitations. The union must bargain with the employer to determine the terms of the collective agreement and must work with the employer in its enforcement. The union is also bound in its treatment of the employees and their concerns by its duty of fair representation[398] and by specific limitations in the NLRA and the Labor Management Reporting and Disclosure Act (LMRDA).[399] Far from being unfettered, the union's grant of authority under section 9(a) is carefully confined and constrained to protect dissenting employees' rights.

Similarly, the Court's emphasis on subject matter over process and product in the determination of what is a government function for the purposes of finding state action makes sense in light of the purpose of the state action doctrine of preserving individual liberty. Democratic procedures and bargaining are commonly used by private parties in our society. Partnerships, cooperatives, corporations, associations, and fraternal organizations all use, and sometimes are required by law to use, democratic procedures for making decisions. Bargaining is the predominant mode of exchange in our economy. Moreover, many of the documents produced by these institutions, including charters, by-laws, rules, and contracts, can be analogized to legislation. It would not only unduly impinge the liberty of such parties but also extend the Constitution beyond any reasonable reading to hold that they are subject to constitutional constraints. On the other hand, if the government were allowed to assign its traditional functions without subjecting the performance of those functions to constitutional scrutiny, the Consti-

---

[397] *See* DEVELOPING LABOR LAW, supra note 160, at 29, 43, 56–58.

[398] *See, e.g.*, Steele v. Louisville & N.R.R., 323 U.S. 192 (1944) (duty of fair representation); 29 U.S.C. § 158(b) (1982) (union unfair labor practices); 29 U.S.C. § 411 (1982) (union members' "bill of rights"); 29 U.S.C. § 501 (1982) (union officers subject to fiduciary duties). Some may think it ironic that the duty of fair representation, which the Court found implied in section 9(a) to avoid the question of whether the exclusive representative's acts constitute state action, should now be used to argue that the exclusive representative's acts do not constitute state action. Nevertheless, as currently formulated, the duty of fair representation is a significant statutory limitation on the grant of authority to exclusive representatives under section 9(a).

[399] 73 Stat. 519 (codified as amended at 29 U.S.C. §§ 401–531 (1982 & Supp. 1987)).

tution's protection of individual liberty would soon be undermined. The government could merely delegate its functions of conducting elections, tax collection, benefit distribution, and law enforcement to private parties with all of the authority of the government, but with none of its constraints or responsibilities.

C. *State Encouragement of Union Security Agreements*

The second argument in favor of finding state action in the negotiation and observance of agency shop agreements under the NLRA is that section 8(a)(3) and other provisions of the Act encourage collective bargaining and agency shop agreements, and therefore the negotiation of such an agreement is an act attributable to the state. As previously discussed, congressional purposes in enacting the NLRA included promoting collective bargaining and shifting bargaining power to employees.[400] These purposes are reflected in several provisions of the Act. Not only does the ban on unfair labor practices listed in section 8(a) prohibit a variety of employer techniques for discouraging employee organization,[401] but the election procedure of section 9 arguably facilitates such employee organization.[402] Once the employees select a representative, that union is designated under section 9(a) as the exclusive representative for all employees in the unit.[403] The employer is required to bargain in good faith with this exclusive representative over wages, hours, and working conditions, including union security agreements.[404] Union security agreements, including agency shop agreements, are authorized by section 8(a)(3), exempting them from the general prohibition against employer discrimination on the basis of union affiliation.[405] One could argue that if the Act encourages collective bargaining and shifts bargaining power to employees, the predictable result is that unions will achieve more of their

---

[400] 29 U.S.C. § 151 (1982); DEVELOPING LABOR LAW, *supra* note 160, at 27–28; Keyserling, *supra* note 220, at 206–08.

[401] 29 U.S.C. § 158(a) (1982).

[402] *See* 29 U.S.C. § 159 (1982). *But see* Weiler, *Promises to Keep: Securing Workers' Rights to Self-Organization Under the NLRA*, 96 HARV. L. REV. 1769 (1983) (election procedure under NLRA lends itself to employer delay and obstruction).

[403] 29 U.S.C. § 159(a) (1982).

[404] 29 U.S.C. §§ 158(a)(5), (d) (1982).

[405] 29 U.S.C. § 158(a)(3) (1982).

collective bargaining objectives—including agency shop agreements.[406]

This argument is subject to empirical criticism. Although the NLRA's purposes of promoting collective bargaining and shifting bargaining power to employees are well established in the lore of labor law, it is questionable whether these objectives have in fact been achieved by the statute. Professor Weiler has argued persuasively that the lack of adequate penalties for employer unfair labor practices and the protracted election procedures under the NLRA allow employers effectively to discourage employee organization.[407] Comparisons of employee organizing under the NLRA with employee organizing in the public sector and in other countries support this assertion and suggest that employer recalcitrance has contributed significantly to the recent precipitous decline in the percentage of employees organized in the private sector.[408] There is no substantial penalty to enforce the employer's obligation to bargain under the NLRA.[409] Moreover, as Professor Stone has noted, the Supreme Court has weakened the employer's obligation to bargain by excluding such "management decisions" as subcontracting and the introduction of new technology from the range of subjects which the employer must negotiate.[410] Stone argues that this circumscription has lessened unions' bargaining power.[411] Finally, Professor Klare has argued that in interpreting the NLRA the Court has "deradicalized" the statute by emphasizing the congressional purposes consistent with liberal capitalism while

---

[406] Brief of Respondents at 8–9, *Beck*, 108 S. Ct. 2641. It is also sometimes argued that Section 8(a)(3) itself encourages the negotiation of union security agreements. *Id.* This argument, however, seems baseless. The language of Section 8(a)(3) merely authorizes union security agreements, it does not encourage or coerce them. 29 U.S.C. § 158(a)(3) (1982). Indeed, before the passage of Section 8(a)(3) all manner of union security agreements both existed and were enforceable, perhaps even by a minority union. Reply Brief of Petitioners at 1–2, *Beck*, 108 S. Ct. 2641. In its current form, Section 8(a)(3) prohibits the closed shop, limits the enforceability of the union shop and conditions the negotiation of a union security agreement on the union's status as the duly elected or recognized exclusive representative. 29 U.S.C. § 158(a)(3) (1982). Thus, by itself, Section 8(a)(3) can be viewed as a limitation on the negotiation of union security agreements.

[407] *See* Weiler, *Striking a New Balance: Freedom of Contract and the Prospects for Union Representation*, 98 HARV. L. REV. 351, 358–62 (1984); Weiler, *supra* note 402, at 787–97.

[408] R. FREEMAN & J. MEDOFF, WHAT DO UNIONS DO? 221 (1984).

[409] *See, e.g.*, Ex-Cell-O Corp., 185 N.L.R.B. 107 (1970) (denying NLRB power to order employer who refused to bargain to compensate employees by the amount they would have received if there had been good faith bargaining).

[410] Stone, *supra* note 381, at 1547–52.

[411] *Id.* at 1557–58.

ignoring Congress's more radical purposes,[412] including the equalization of bargaining power between employees and employers.[413]

The argument that the NLRA encourages agency shop agreements is also subject to criticism on theoretical grounds. It seems clear that encouragement of collective bargaining would encourage, or at least create more opportunities for, the negotiation of agency shop agreements. However, an increase in employees' bargaining power due to the NLRA would not necessarily encourage the negotiation of agency shop agreements. If agency shop agreements are worth more to employees than they cost employers, presumably even a very powerful employer will offer to include such a provision in the collective agreement in exchange for a wage decrease of greater value to him, but of lesser value to the employees than the agency shop agreement. Similarly, if agency shop agreements are worth less to employees than they cost employers, even a very powerful union would agree to omit an agency shop agreement in exchange for a wage increase of greater value to the employees, but less cost to the employer than the agency shop agreement. In other words, whether an agency shop agreement is included in the collective agreement will depend on whether it is an "efficient" contract term, in that its benefits to the employees outweigh its costs to the employer, and not on the respective bargaining power of the parties.[414] At least in a simple economic model, bargaining power is used to determine the division of any economic profits or rents, not to determine the inclusion or exclusion of any specific contract term.[415]

---

[412] Klare, *supra* note 220, at 265–70.

[413] *Id.* at 292–93. *See also* First National Maintenance Corp. v. NLRB, 452 U.S. 666, 676 (1981) (congressional purpose was to create system for resolution of conflicts, not to make labor an "equal partner" in the operation of the business).

[414] Schwab, *Collective Bargaining and the Coase Theorem*, 72 CORNELL L. REV. 245, 245–49 (1987); *see also* Coase, *The Problem of Social Cost*, 3 J.L. & ECON 1 (1960).

[415] Schwab, *supra* note 414, at 245–49. There are some subtle but potentially serious shortcomings in this argument. First, it ignores potential wealth effects. That is to say it assumes that as employees' bargaining power and wealth grow, their valuation of an agency shop agreement will remain the same. *See* Regan, *The Problem of Social Cost Revisited*, 15 J.L. & ECON. 427, 432–33 (1972). Whether or not such wealth effects would be important in the case of union security agreements is an empirical question. *See* Schwab, *A Coasean Experiment on Contract Presumptions*, 17 J. LEG. STUD. 237, 265 (1988). The argument also assumes that employers and unions always bargain to an efficient solution analogous to what would prevail in a competitive market. *See* Coleman, *Efficiency, Exchange, and Auction: Philosophic Aspects of the Economic Approach to Law*, 68 CAL. L. REV. 221, 244 (1980); Cooter, *The Cost of Coase*, 11 J. LEG. STUD. 1, 14 (1982). Such an assumption seems a gross simplification of the bargaining relationship between employers and unions.

Once again, if we apply current Supreme Court doctrine to the argument that the NLRA encourages agency shop agreements, we find little hope that the Court would find the necessary nexus to establish state action. State encouragement of a private party's action will cause it to be considered state action only if the state "has exercised coercive power or has provided such significant encouragement, either overt or covert," that the action "must in law be deemed to be that of the State."[416] "Mere [state] approval of or acquiescence in the initiatives of a private party is not sufficient" for a finding of state action.[417] Moreover, during the tenures of the Burger and Rehnquist Courts, the tendency has been to examine the activity in question very narrowly, requiring state coercion or significant encouragement of the specific action or decision.[418] A general argument that the NLRA encourages collective bargaining and shifts bargaining power to employees, resulting in more agency shop agreements, would not prevail under the current doctrine. Instead, one would have to show that the government coerced or significantly encouraged the specific decision by the union and the employer to enter into an agency shop agreement.

It is "the fundamental premise" of the NLRA that the government regulates only the process of collective bargaining,

---

[416] Blum v. Yaretsky, 457 U.S. 991, 1004 (1982) (citing Flagg Bros., Inc. v. Brooks, 436 U.S. 149, 166 (1978); Jackson v. Metropolitan Edison Co., 419 U.S. 345, 357 (1974); Moose Lodge No. 107 v. Irvis, 407 U.S. 163, 173 (1972); Adickes v. S.H. Kress & Co., 398 U.S. 144, 170 (1970)). *See also* San Francisco Arts & Athletics v. United States Olympic Comm., 483 U.S. 522, 543–44 (1987); Rendell-Baker v. Kohn, 457 U.S. 830, 840 (1982).

[417] *Blum*, 457 U.S. at 1004–05, (citing *Flagg Bros.*, 436 U.S. at 164–65; *Jackson*, 419 U.S. at 357 (1974)). *See also San Francisco Arts & Athletics*, 483 U.S. at 543–44.

[418] For example, in Blum v. Yaretsky, 475 U.S. 991 (1982), the Court found that the decisions of committees of private doctors to discharge and transfer Medicare patients were not state action because the actual decisions to discharge and transfer were made by the doctors on the basis of professional standards not established by the state. *Id.* at 1008. No state action was found, despite the fact that the federal government paid for the patients' medical care, regulated their treatment, required nursing homes and hospitals to set up the committees, prescribed the form the committees would use in deciding on discharges and transfers, penalized nursing homes and hospitals if they did not make appropriate decisions in discharges and transfers, reviewed the decisions to discharge or transfer, and reduced payments for services in accordance with the committee decisions. *Id.* at 1006–10. Similarly, in Rendell-Baker v. Kohn, 457 U.S. 830 (1982), the decision of a private school to discharge a teacher was not considered state action because the state did not coerce or significantly encourage the specific decision to fire the teacher. *Id.* at 841. In *Rendell-Baker* the school performed a public function (although not one exclusively performed by the state) and the state sent the students to the school, provided 90% of its funding, and had the right of approval in some hiring. *Id.* at 832–35. *See also San Francisco Arts & Athletics*, 483 U.S. at 543–44; *Flagg Bros.*, 436 U.S. at 166; *Jackson*, 419 U.S. at 357.

leaving the determination of the specific terms of the collective agreement to the private parties.[419] Since the passage of the Wagner Act, both Congress and the courts have understood that the NLRA's obligation to bargain in good faith does not require the parties to make any specific concessions; the Act does not allow the government to supervise the terms of collective agreements.[420] With the passage of the Taft-Hartley Act in 1947, Congress has expressly excepted any obligation to make a concession from the obligation to bargain in good faith.[421] As described by Senator Walsh (D-Mass.), chairman of the Committee on Education and Labor during the enactment of the Wagner Act, the Act "indicates the method and manner in which employees may organize . . . and leads them to the office door of their employer, . . . [but] does not go beyond the office door . . . ."[422] Whatever encouragement for collective bargaining and the promotion of employee interests can be found in the NLRA, it is not used to coerce or significantly encourage the inclusion of any specific term in the collective agreement. Consistent with this approach, section 8(a)(3) merely authorizes the parties to enter into union security agreements, neither requiring nor encouraging them.[423]

---

[419] While the parties' freedom of contract is not absolute under the Act, allowing the Board to compel agreement when the parties themselves are unable to do so would violate the fundamental premise on which the Act is based—private bargaining under governmental supervision of the procedure alone, without any official compulsion over the actual terms of the contract.

H.K. Porter Co. v. NLRB, 397 U.S. 99, 108 (1970) (footnote omitted). *See also* Howard Johnson Co. v. Hotel Employees, 417 U.S. 249, 254 (1974).

[420] The committee wishes to dispel any possible false impression that this bill is designed to compel the making of agreements or to permit governmental supervision of their terms. It must be stressed that the duty to bargain collectively does not carry with it the duty to reach an agreement, because the essence of collective bargaining is that either party shall be free to decide whether proposals made to it are satisfactory.

S. REP. NO. 573, 74th Cong., 1st Sess. 12 (1935), *reprinted in* 2 NLRA LEGISLATIVE HISTORY, *supra* note 117, at 2300, 2312. Similarly, in the first major test of the constitutionality of the NLRA, the Court said:

> The Act does not compel agreements between employers and employees. It does not compel any agreement whatever . . . . The theory of the Act is that free opportunity for negotiation with accredited representatives of employees is likely to promote industrial peace and may bring about the adjustments and agreements which the Act in itself does not attempt to compel.

NLRB v. Jones & Laughlin Steel Corp., 301 U.S. 1, 45 (1937).

[421] 29 U.S.C. § 158(d) (1982).

[422] 79 CONG. REC. 7659 (1935), *reprinted in* 2 NLRA LEGISLATIVE HISTORY, supra note 117, at 2373.

[423] *See supra* notes 406–415 and accompanying text.

The analysis above ignores the fact that in *Railway Employees' Department v. Hanson*,[424] the Supreme Court found that the negotiation of a union security agreement under section 2 Eleventh of the RLA (the RLA provision analogous to section 8(a)(3)) was sufficiently encouraged to constitute state action.[425] However, the Court's finding of state action in that case turned on the exercise of the Supremacy Clause in section 2 Eleventh.[426] The RLA authorizes union security agreements notwithstanding contrary state laws. By pre-empting contrary state laws, the RLA legalizes otherwise unlawful private action and thus is said to give union security agreements the imprimatur of federal law.[427] There is no similar exercise of the Supremacy Clause under the NLRA. Section 14(b) of the Act allows contrary state law to supersede the authorization of union security agreements in section 8(a)(3).[428]

The denial of a finding of state action based on the NLRA's general promotion of collective bargaining and employee interests is consistent with the purpose of the state action doctrine to preserve individual liberty. To allow the state to resort to any form of encouragement or coercion of specific private acts without subjecting them to constitutional scrutiny would undermine the Constitution's check on government power. The government could escape its constitutional limitations merely by coercing powerful private parties to carry out its forbidden objectives.[429] However, it is commonplace for states to pass laws designed to foster a type of private institution or activity with the hope of obtaining certain desired results. General encouragement of this sort is much less subject to abuse by the government. Although I will not construct a general dividing line between these two cases, the encouragement of collective bargaining under the NLRA clearly falls much closer to the latter case than to the former. Moreover, the level of government encouragement under the NLRA is well below the level of encouragement that has been tolerated in other instances. For example, the government has undertaken extensive efforts to foster private enter-

---

[424] 351 U.S. 225 (1956).
[425] *Id.* at 232.
[426] *Id.*
[427] *See id.*; *see also* Abood v. Detroit Bd. of Educ., 431 U.S. 209, 218 n.12 (1987).
[428] 29 U.S.C. § 164(b) (1982).
[429] For example, the state might silence a dissenting author by offering great rewards or punishments to private publishers according to their refusal or agreement to publish the author's work.

prise in hopes of promoting economic well-being. Incorporation laws, partnership laws, tax laws, government loans and grants, and even our conceptions of private property and theft are all arguably designed to protect and support private enterprise and economic growth.[430] The government's efforts to foster employee organization and collective bargaining in hopes of promoting employee interests and industrial peace pale by comparison. If the Court were to find that collective bargaining and its results are attributable to the state, it would also have to consider whether a vast portion of private enterprise and its results are attributable to the state.[431]

What seems at odds with the purpose of the state action doctrine is the Supreme Court's holding in *Hanson*. In section 2 Eleventh of the RLA, Congress exempted union security agreements from a general prohibition against employer discrimination on the basis of union affiliation. Thus, Congress decided not to prohibit this type of conduct. Moreover, because section 2 Eleventh pre-empted contrary state laws, Congress decided that states would not prohibit such conduct. It seems curious that a decision *not* to regulate an area of private activity gives rise to a finding of state action. By finding state action this way, the Court interjects constitutional standards that restrict the private parties' individual freedom which are "uninvited" by affirmative legislation.[432] The fact that Congress exercised the Supremacy Clause in its decision not to prohibit union security agreements seems irrelevant to the determination of whether there is state action. The Supremacy Clause was fashioned to cope with problems entirely unrelated to the question of when

---

[430] *See, e.g.*, R. CUNNINGHAM, W. STOEBUCK & D. WHITMAN, THE LAW OF PROPERTY § 1.1 (1984) (property is a legally protected expectation of deriving enjoyment from a thing); R. POSNER, ECONOMIC ANALYSIS OF LAW § 3.1 (3d ed. 1986) (creation of exclusive property rights is a necessary condition for the efficient use of resources); Cohen, *Dialogue on Private Property*, 9 RUTGERS L. REV. 357, 363–65 (1954) (discussion of property as an allocation of wealth).

[431] There have been commentators who believe corporations should be subject to constitutional constraints. *See, e.g.*, A. MILLER, THE MODERN CORPORATE STATE 182–87 (1976); Berle, *Constitutional Limitations on Corporate Activity—Protection of Personal Rights from Invasion Through Economic Power*, 100 U. PA. L. REV. 933, 933–34, 942 (1952); Latham, *The Commonwealth of the Corporation*, 55 NW. U.L. REV. 25, 25–29 (1960). There have also been commentators who disagree. Manning, *Corporate Power and Individual Freedom: Some General Analysis and Particular Reservations*, 55 NW. U.L. REV. 38, 38–54 (1960); Marshall, *supra* note 30, at 569–70.

[432] The fact that this constitutional interjection is "uninvited" by statutory regulation suggests that *Hanson* violates the second purpose of the state action doctrine—preserving federalism and the separation of state and federal powers. *See supra* note 348 and accompanying text.

constitutional restraints should apply to private actions. Indeed, I know of no other case in which the Court has made a similar finding. Confusion of the Supremacy Clause with the state action doctrine can yield ridiculous results. Professor Wellington has suggested that the logic of *Hanson* requires that if Congress legislates to pre-empt state law in a given area, it must affirmatively outlaw all private behavior contrary to the Constitution.[433]

Holding that a union's negotiation and observance of a collective agreement is state action would result in a radical change in our conception of unions under the NLRA. Unions are currently viewed by both the Court and Congress as *private* centers of employee power under the NLRA.[434] Although existing limitations on unions already subject them to constraints similar to those contained in the Constitution,[435] a change in the status of unions from private to state actors would fundamentally affect our labor law and the current conduct of industrial relations. All of a union's actions in the negotiation and enforcement of the collective agreement would be subject to constitutional scrutiny. Would the negotiation of a grievance arbitration procedure in place of the right to litigate contract violations violate employees' seventh amendment rights?[436] The substitution of such

---

[433] Wellington, *supra* note 57, at 356–57.

[434] "[T]he fundamental premise on which the [NLRA] . . . is based [is] . . . *private* bargaining under government supervision of the procedure alone, without any official compulsion over the actual terms of the contract." H.K. Porter Co. v. NLRB, 397 U.S. 99, 108 (1970) (emphasis added).

[435] For example, the duty of fair representation and the employee's "bill of rights" in the LMRDA, 73 Stat. 519 (codified at 29 U.S.C. §§ 401–531) (1982).

[436] Under current labor law, arbitration of contract disputes is favored over litigation because the parties, who are considered private actors, have voluntarily chosen the arbitrator to be an official interpreter of their contract. United Steelworkers v. Warrior & Gulf Navigation Co., 363 U.S. 574, 580–82 (1960). If the union were viewed as a state actor, the typical contract grievance procedure most likely would be considered to violate the grievant's seventh amendment right to a trial by jury. While the state may establish a compulsory non-jury forum for litigation, it must provide a right to trial *de novo* before a jury to avoid seventh amendment invalidity. *See, e.g.*, Capital Traction Co. v. Hof, 174 U.S. 1 (1898) (upholding statute requiring trial before justice of the peace, with appeal to jury trial before court of record); Woods v. Holy Cross Hosp., 591 F.2d 1164 (5th Cir. 1979) (upholding Florida statute requiring referral of malpractice claims to nonbinding "mediation panel"); Kimbrough v. Holiday Inn of Lionville, Inc., 478 F. Supp. 566 (E.D. Pa. 1979) (upholding local rule establishing compulsory nonbinding arbitration). Because the typical grievance arbitration is subject only to limited review, there would be serious seventh amendment problems if the union were considered a state actor. *But see* Keith Fulton & Sons, Inc., v. New England Teamsters and Trucking Indus. Pension Fund, 762 F.2d 1124 (1st Cir. 1984) (compulsory arbitration under Multiemployer Pension Plan Amendments Act not invalid under seventh amendment because Congress may commit enforcement of new public rights to alternative tribunals such as an administrative agency).

grievance arbitration procedures for court procedures in the settlement of contract disputes has been recognized as one of the primary achievements of collective bargaining.[437] It is a cornerstone in the current conduct of labor relations. Finally, a change in the status of unions would result in a fundamental change in the way we make national labor policy. If unions are viewed as state actors, power over our federal labor policy would shift from Congress to the courts.[438] The courts are ill-suited for this task, due to their lack of expertise in the area of labor relations and lack of political mandate to make policy decisions.[439]

V. CONCLUSION

The Court's interpretation of section 8(a)(3) of the NLRA in *Beck* cannot be justified by recourse to the usual intrinsic and extrinsic resources for statutory interpretation. The words, administrative interpretations, and legislative history of the statute all indicate that Congress intended to allow agency shop agreements under section 8(a)(3) and intended purposes for union security agreements beyond the mere recoupment of collective bargaining expenses.

Nor can the Court's interpretation be justified by recourse to the Court's prior interpretations of section 2 Eleventh of the RLA or to the doctrine of avoiding constitutional questions. The Court's interpretation of section 2 Eleventh was affected by its desire to avoid the question of whether the negotiation and observance of agency shop agreements under the RLA violated dissenting employees' constitutional rights. Regardless of whether the Court's constitutionally colored interpretation of section 2 Eleventh was appropriate, *Beck* is an inappropriate case for application of the doctrine of avoiding constitutional questions. There is no serious question of the constitutionality of agency shop agreements to avoid in *Beck*, and the Court's

---

[437] John Wiley & Sons, Inc. v. Livingston, 376 U.S. 543, 549–50 (1964); L. HILL & C. HOOK, MANAGEMENT OF THE BARGAINING TABLE 199 (1945). *See generally* N. CHAMBERLAIN & J. KUHN, COLLECTIVE BARGAINING 138–65 (3d ed. 1986); S. SLECHTER, J. HEALY & E. LIVERNASH, THE IMPACT OF COLLECTIVE BARGAINING ON MANAGEMENT 692–738 (1960).

[438] Wellington, *supra* note 57, at 361.

[439] *Id.* at. 361–64.

interpretation of section 8(a)(3) is not supported by the rationales of the doctrine.

An examination of the constitutional question avoided in *Beck* discloses that there is no state action to support the claim that negotiation and observance of agency shop agreements under the NLRA violate dissenting employees' first amendment rights. Under the Court's current formulation of the state action doctrine, there is insufficient grant of authority to the union as exclusive representative and insufficient encouragement of the negotiation of union security agreements for the negotiation and observance of such agreements to constitute state action. Indeed, it is well established in American labor law that the parties to the collective agreement, not the government, determine the provisions of that agreement. The determination that there is no state action in the negotiation and observance of union security agreements is consistent with the purpose of preserving individual freedom represented in the state action doctrine. If constitutional scrutiny were extended to the negotiation and observance of union security agreements, it would logically extend to a host of other activities now considered private and not subject to constitutional constraints.

The Court's opinion in *Beck* amounts to judicial legislation with no basis in the statute or the Constitution. If the Court had properly performed its traditional role as statutory and constitutional umpire, it would have interpreted section 8(a)(3) to allow the negotiation and observance of agency shop agreements and would have determined that such agreements do not violate dissenting employees' first amendment rights. Instead, because of its failure to acknowledge the effect of past constitutional concerns on its interpretation of section 2 Eleventh and to consider whether constitutional concerns required a similar result in interpreting section 8(a)(3), the Court abandoned its traditional role and usurped the legislature's role by specifying both the policy and form of section 8(a)(3). The *Beck* decision limits the enforceability of agency shop agreements in collective bargaining agreements covering over six million workers and opens a potentially significant drain on the resources of the national labor movement through bookkeeping and litigation expenses required to resolve complaints of dissenters. One can only hope that the mistakes of the *Beck* decision will not be repeated.

STATUTE

A DRAFT BILL TO ALLOW CHOICE BETWEEN NO-FAULT AND FAULT-BASED AUTO INSURANCE*

JEFFREY O'CONNELL**

Recently, furor over auto insurance has rekindled. Tort awards for accidents and insurance rates have escalated in recent years, due in large part to the greater ability of accident victims to bring tort suits in states that have formally adopted no-fault insurance systems. No-fault states typically make the purchase of no-fault insurance mandatory; however, they allow plaintiffs to bring tort suits if their injuries exceed a designated threshold level. As inflation erodes these thresholds and attorney involvement artificially inflates the amount of damages plaintiffs seek, the use of the court system has increased dramatically, and the advantages offered by no-fault insurance have been circumvented.

Professor O'Connell, a leading advocate of auto no-fault insurance, proposes to alleviate many of the problems associated with this crisis. O'Connell suggests that states that allow for suits in tort are closing off an effective method of controlling insurance premium costs and erratic recoveries by accident victims. O'Connell proposes that states instead should maximize the benefits that a "pure" no-fault system can offer by allowing their citizens to choose between fault-based and no-fault auto insurance and then drastically limiting, with very few exceptions, the ability of no-fault insureds to sue in tort. He provides a model statute that delineates how the two insurance schemes can co-exist in one state with the advantages of providing consistent recovery at the least cost and maximizing consumer choice. Professor O'Connell's proposal is a provocative one sure to attract the attention of policymakers and legislators who have struggled with this problem in recent years.

I. INTRODUCTION[1]

Auto insurance is back in the headlines after being eclipsed for years by controversies raging around other areas of tort law. The rise in auto insurance rates and consequent bitter consumer reaction in some states have brought the issue of auto insurance back to the center of public discussion. In particular, as interest in and opinions about auto insurance strengthen, the debate

---

* This Article develops a draft bill originally suggested in O'Connell & Joost, *Giving Motorists a Choice Between Fault and No-Fault Insurance*, 72 VA. L. REV. 61 (1986). I am very grateful to the Master and Fellows of Downing College, Cambridge, where, as the Thomas Jefferson Visiting Fellow, I worked extensively on this draft bill during the Easter Term, 1989.

** John Allan Love Professor of Law, University of Virginia. B.A., Dartmouth College, 1951; J.D., Harvard University, 1954.

[1] I am very indebted to Peter Spiro, J.D., University of Virginia Law School, 1988, for his help on this introduction.

over the relative merits of no-fault insurance schemes is receiving a great deal of attention in insurance trade journals, the popular press, and legislative halls.[2] With no-fault insurance, persons injured in automobile accidents are compensated without regard to fault. At least to some extent, this insurance coverage replaces tort actions against the insured.

No-fault insurance appears to be the most viable cure to the ills of the auto insurance industry. This is confirmed by a thorough statistical analysis compiled by the U.S. Department of Transportation, which indicates that no-fault is more effective than the tort system in dealing with auto accident claims.[3]

Specifically, as compared with fault-based systems, no-fault compensates more people (roughly twice as many)[4] with greater benefits (seventy-nine percent more)[5] in quicker payments (almost all no-fault payments, as opposed to only half of tort awards, are made during the first year following injury).[6] Drivers also get more coverage for their no-fault premium dollar, the logical result of significant administrative cost reductions. Furthermore, a substantial percentage of accident-related small claims have been kept out of court (auto-accident suits have been reduced by as much as two-thirds in some states), representing millions in saved taxpayer dollars.[7]

Despite these favorable statistics, accident victims have not yet received the greatest possible benefits that no-fault insurance could provide. This is because none of the roughly two dozen states[8] that have adopted no-fault has, strictly speaking, a pure no-fault regime. No single state has moved so dramatically as to bar lawsuits in the most serious accident cases,[9] and most do not go nearly that far. Under many state no-fault laws, motorists

---

[2] *See, e.g.*, N.Y. Times, Mar. 19, 1989, § 12, at 1, col. 1.

[3] U.S. DEP'T OF TRANSP., COMPENSATING AUTO ACCIDENT VICTIMS: A FOLLOW-UP REPORT ON NO-FAULT AUTO INSURANCE EXPERIENCES (1985) [hereinafter DOT REPORT]. For a summary of key portions of the DOT REPORT, see O'Connell & Joost, *Giving Motorists a Choice Between Fault and No-Fault Insurance*, 72 VA. L. REV. 61, 63–75 (1986) [hereinafter O'Connell & Joost].

[4] DOT REPORT, *supra* note 3, at 3.

[5] *Id.* at 6.

[6] *Id.* at 4.

[7] *Id.* at 5, 113–17.

[8] ALL-INDUSTRY RESEARCH ADVISORY COUNCIL (AIRAC), COMPENSATION FOR AUTOMOBILE INJURIES IN THE UNITED STATES 133 (1989) [hereinafter AIRAC].

[9] *Cf.* Automobile Insurance Act, QUE. REV. STAT. ch. A-25 (1977), *amended by* ch. 38, 1980 QUE. STAT. 383, ch. 25, 1981 QUE. STAT. 425, ch. 52, 1982 QUE. STAT. 1033, ch. 59, 1982 QUE. STAT. 1173; O'Connell & Tenser, *North America's Most Ambitious No-Fault Law: Quebec's Auto Insurance Act*, 24 SAN DIEGO L. REV. 917 (1987).

are required to purchase only a limited amount of no-fault coverage (sometimes as little as a few thousand dollars). Therefore, seriously injured accident victims are often in the position of depending on the undependable variables of a tort claim if their loss exceeds no-fault limits.[10]

Moreover, no-fault insurance benefits, especially when relatively generous, often subsidize the litigation of tort claims that no-fault itself was designed to eliminate. When combined with the ability to hire a lawyer on a no-risk, contingent fee, the money paid to accident victims under no-fault policies reduces their incentive, as well as their need, to negotiate a relatively quick and low out-of-court settlement with the other party. With the no-fault payments serving as a sort of cushion, victims can better afford to endure a long, drawn-out battle, often resulting in relatively higher settlements.[11]

What has brought such compromises that undermine no-fault laws? Why has no new state adopted a no-fault bill since 1975, and why have a few other states repealed or rolled back no-fault laws despite a wave of statistics supporting no-fault's advantages?

One answer lies with the nation's trial lawyers, who have banded together to oppose the passage or improvement of no-fault schemes.[12] This group continually emphasizes that Americans must retain their "right to sue."[13]

---

[10] DOT REPORT, *supra* note 3, at 5, 20, 25–37, 40–49; O'Connell & Joost, *supra* note 3, at 84.

[11] O'Connell & Guinivan, *An Irrational Combination: The Relative Expansion of Liability Insurance and Contraction of Loss Insurance*, 49 OHIO ST. L.J. 757 (1988).

[12] J. O'CONNELL & C.B. KELLY, THE BLAME GAME 117–18 (1986). Many in the insurance industry who believe in no-fault insurance (and there has been a variety of opinions in the industry) began to lose heart in the struggle for no-fault in the mid-70's when they saw how easily even well-crafted no-fault bills could be subverted or perverted during the legislative process by trial lawyers, among others. As one insurance industry lawyer explained:

> It doesn't take too much to undo a reasonably good no-fault bill, just the cheerful cooperation of the friends and enemies of no-fault insurance, the friends in raising the benefits without much regard for the threshold [beyond which tort suits can also be brought]; the enemies in lowering the threshold while aiding and abetting the friends in raising the benefits. Then, add a little political grease in the form of a mandated rate reduction and a few other provisions . . . designed only to punish insurers for their support of the concept, and we [in the industry] begin to wonder if the battle is really worth fighting.

Letter to Jeffrey O'Connell (Mar. 22, 1988) (on file with the HARV. J. ON LEGIS.). For a good journalistic discussion of the problems currently plaguing auto insurance, see N.Y. Times, Nov. 6, 1988, at F1, col. 2.

[13] O'CONNELL & KELLY, *supra* note 12, at 119.

Automobile drivers have been exercising this "right to sue" with greater and greater frequency, even in states with no-fault systems. Although no-fault laws were designed to eliminate many tort claims (and have done so),[14] they are doing so with decreasing effectiveness, thus sabotaging no-fault states' attempts to avoid the disadvantages of the tort system. According to a recent insurance industry study conducted by the All-[Insurance] Industry Research Advisory Council (AIRAC):

> Tort thresholds in the no-fault states, designed to reduce cost and legal complexities for the less serious injuries, appear to have eliminated about 21% of the potential liability claims among persons collecting no-fault PIP [personal injury protection] benefits in 1987. This is only half as many as were eliminated in 1977 . . . . This decrease is one measure of how much the effect of the tort thresholds has been eroded by inflation and by [broadening] legal interpretations[15] over the intervening 10 years. The connection between attorney involvement and higher claims costs results in a vicious cycle.[16]

With the erosion of tort thresholds has come an accompanying increase of attorney involvement in accident claims. Overall, attorney involvement in all types of auto personal injury claims has increased forty-two percent since 1977.[17] Interconnecting with increased attorney involvement is the well-known escala-

---

[14] DOT REPORT, *supra* note 3, at 5 and accompanying text.

[15] *See, e.g.*, DiFranco v. Pickard, 427 Mich. 32, 398 N.W.2d 896 (1986) (holding, among other things, that recovery of noneconomic damages under no-fault was by no means confined to catastrophic injuries and that in determining serious impairment of bodily function, one should focus not on the injuries themselves, but on how they affected bodily functions; also holding that determining that whether one is injured seriously enough to sue in tort is a jury question).

[16] AIRAC, *supra* note 8, at 13. The erosion of tort thresholds has been worst in states having monetary versus verbal thresholds. On this distinction, see DOT REPORT, *supra* note 3, at 93–94.

> A surprisingly high percentage of PIP claimants were judged to be eligible for a tort (fault-based) claim in some of the no-fault states, despite the thresholds. In New Jersey, for example, file reviewers indicated that 62.7% of persons receiving PIP benefits could successfully file a [tort bodily injury] BI liability claim under that state's medical expense threshold. More than half (53.5%) of PIP claimants in Massachusetts also were judged to be eligible for a tort claim, as were 49.4% of PIP claimants in Georgia and 41.4% of those in Connecticut . . . . Other states with relatively low percentages of PIP claimants eligible for BI claims included . . . New York (28.7%) and Florida (32.5%).

Id. at 13–14. (Of the above states, only New York and Florida had verbal thresholds.) But for the high cost of even the relatively few surviving tort claims—which after all are the more serious claims—see *infra* notes 28–31 and accompanying text.

[17] *Id.* at 9. "Possible reasons for the general increase in attorney representation include a large increase in the number of attorneys and the fact that the U.S. Supreme Court struck down prohibitions on attorney advertising in 1978." *Id.*

tion between 1977 and 1989 in medical expenses that "accounted for almost 70% of the economic losses reported by accident victims."[18] The Consumer Price Index for medical care items increased about 124% over that ten-year period.[19] High medical bills incurred by victims not only increase costs by increasing the size of claims that insurers have to pay, but they boost the amount of damages paid for pain and suffering.

> Since pain and suffering is an intangible loss that cannot be directly measured, payment typically is treated as a multiple of the tangible economic losses such as medical bills and lost wages. People who were paid under the BI liability coverage collected an average of $211 for every $100 of economic loss they reported, so an increase in medical expense of $100 would have increased the payments by about $211.[20]

At this juncture, one reaches the crucial issue that is the key to controlling auto insurance costs. Claimants with attorneys report much higher than average amounts of economic loss,[21] regardless of the severity of their injuries, mainly because they obtain much more medical treatment.[22] This somewhat inflated view of economic loss translates into higher settlements for both tangible and intangible injuries, thus straining the system and increasing auto insurance premiums.

Proponents of the tort system may argue that lawyer-assisted recoveries are important because Americans should be able to exercise their "right to sue." However, this right is expensive, and the results of suits are unpredictable. Additionally, some would argue that attorney-assisted recoveries ensure that accident victims are amply compensated. But in reality, the tort system promotes windfalls to some victims at a great cost to

---

[18] *Id.* at 8.

[19] *Id.*

[20] *Id.* at 9.

[21] *Id.* at 10. "Fewer than 10% of claimants were represented [by attorneys] when they had reported economic losses of $200 or less, but more than 60% of PIP claimants and more than 80% of BI . . . claimants had attorneys when the report of economic loss exceeded $5,000." *Id.* at 9.

The no-fault PIP claims were much less likely to involve attorneys, but attorney representation of PIP claims has increased 82% over the past 10 years, probably because of the erosion of tort thresholds in the no-fault states. In most instances, however, attorneys become involved in PIP claims because there is an associated tort claim; only 3.8% of PIP claimants were reported to have hired attorneys solely to handle their PIP claims. *Id.* at 10.

[22] The greater use of medical treatment persists even when claims are sorted by type of injury, days of restricted activity, need for hospitalization, and other measures of injury severity. *Id.*

the system, while leaving others with little or no recovery. States that have mandated a fault-based system, and those that have low thresholds for fault-based claims in no-fault systems, have institutionalized the "right to sue." Perhaps instead of providing such rigid protection for this "right," however, states should protect consumers' "right to choose" between no-fault and fault-based insurance.

As things now stand, the consumer is allowed little discretion in deciding what type of auto insurance he will purchase; his state government largely decides it for him. If he lives in a typical fault-oriented state, he is under legislative mandate to buy a certain amount of liability insurance, but he cannot opt for no-fault coverage at the price of eschewing tort claims. In the so-called no-fault states, residents must purchase some of each, again without provisions for such a choice.

At this time, however, provision of a choice between fault and no-fault would not work under either of the aforementioned state systems. The problem is that auto insurance is interactive—one consumer's insurance selection affects another's ability to get compensation for an injury. If one driver negligently causes serious injury to another, the victim's possibilities for a courtroom recovery will likely depend on how large a liability insurance policy the person at fault holds. Therefore, most states require all drivers to have some minimum coverage.

This sort of interaction would seem likely to cripple any plan permitting a basic choice on types of insurance. If two drivers with no-fault policies collided there is no problem: each would collect from his own insurer. If each carried only third-party insurance, likewise, no difficulties arise: each would claim against the other, as presumably desired. But what if a no-fault motorist ran into one holding liability insurance? Here we find the stumbling block. Mr. No-Fault would have forsaken his right to sue, as a condition of his policy, regardless of the other's negligence. Mr. Fault, however, would still be counting on a fault-based claim against the other motorist for his compensation. If the law allows drivers with liability insurance to bring an action, the no-fault-insured would have to insure both against his own injuries and against any he might cause Mr. Fault, too. What result? No-fault insurance might become far more expensive than fault coverage, and despite other advantages, the built-in cost disincentive would undoubtedly discourage motorists

from making the switch.[23] In addition, there is the inherent unfairness of Mr. No-Fault being unable to sue Mr. Fault while Mr. Fault retains the right to sue Mr. No-Fault, with the attendant windfall to Mr. Fault.

A relatively minor adjustment to the available systems could allow for choice. A "choice" scheme could become workable with the creation of a mechanism to cover collisions between fault and no-fault cars.[24] In fact, a version of that mechanism already exists. Insurance companies routinely are required to offer policies to protect drivers against accidents with uninsured motorists and against damage done by hit-and-run vehicles. Such coverage compensates insureds to the same extent they would have been had the uninsured motorist been at fault and insured.[25]

UM insurance could be extended to treat no-fault insureds as, in effect, uninsured motorists, thus allowing recovery by Mr. Fault if he were to be injured in a collision with a no-fault driver. By the same token, Mr. No-Fault, covered for his medical and

---

[23] O'Connell & Joost, *supra* note 3, at 77–78. This kind of provision was added to the District of Columbia's no-fault statute under the dubious title of "reform." *See D.C. Scraps Fault Requirement*, Wash. Post, Nov. 20, 1985, at 1, col. 4. *But see infra* note 37, concerning a Kentucky "choice" law.

[24] O'Connell & Joost, *supra* note 3, at 78–89. For a crisp description of this "choice" plan, see Passell, *Selling No-Fault Auto Insurance*, N.Y. Times, Nov. 23, 1988, at D2, col. 1. For a recent change in New Jersey's no-fault law allowing a choice between "add-on" no-fault insurance barring no tort suits and no-fault insurance with a high (New York-type) "no-lawsuit" threshold barring most tort suits, see N.J. STAT. ANN. § 39:6A-8 (West 1988). (On the difference between "add-on" and "no-lawsuit" thresholds, see O'Connell & Joost, *supra* note 3, at 63–64.) The New Jersey choice bill is projected to save around 50% in bodily injury premiums for those choosing no-fault. Cummins, *What's Driving Auto Insurance Up*, Wall St. J., Jan. 5, 1989, at A10, col. 3. The choice bill herein proposed should produce much greater savings. Those choosing fault-based coverage will not have to purchase no-fault coverage, and the no-fault option avoids both New Jersey's cumbersome, bureaucratic redistribution of the benefits of tort waivers (O'Connell & Joost, *supra* note 3, at 67), and preservation of serious tort cases, which is very expensive in New York. *See infra* notes 28–31 and accompanying text.

[25] *See generally* A. WIDISS, UNINSURED AND UNDERINSURED MOTORIST INSURANCE (2d ed. 1987). Uninsured Motorist coverage (UM) pays when the insured or others in the insured vehicle are injured by an uninsured or hit-and-run driver. The insured's own insurer pays what the injured person was eligible to recover in tort from the other uninsured at-fault driver. Underinsured Motorist coverage (UIM) similarly pays the insured and other occupants of his vehicle when the at-fault other driver is insured with low limits. For a cogent criticism of some features of these coverages, see Schwartz, *A Proposal for Tort Reform: Reformulating Uninsured Motorist Plans*, 48 OHIO ST. L.J. 419 (1987). However, these inadequacies are mostly a function of (a) the shortcomings of fault-based coverage, and (b) the attempts by the courts to expansively (and expensively) provide more payouts under such coverage. Under the reform suggested here, an alternative is provided to fault-based coverage, including UM coverage itself.

wage losses, could neither sue nor be sued based on fault. While it is true that Mr. Fault's UM coverage costs would rise as a result of the larger class of "uninsured" drivers, his liability premiums would correspondingly decrease because none of the no-fault policy holders would be able to bring suit against him.

The greatest advantage of such a framework is that it would allow free choice to be made in the insurance marketplace. The driving public's collective choice would emerge in a referendum of sorts, a tally not so easily refuted by the losing side.

Moreover, this scheme would avoid pricing problems by insulating the costs of fault and no-fault policies from each other; no longer would the no-fault motorist find himself subsidizing his more litigious counterparts.[26] Consequently, many more benefits per premium dollar could be provided under a choice system. Even in states having high levels of no-fault benefits, substantial amounts in total premiums applicable to bodily injury could be saved by a choice of no-fault over fault.[27] For example, in New York, which has a high no-fault benefit level of up to $50,000, premiums were thirty percent lower in 1987 than they would have been had a no-fault law not been enacted;[28] however, much more money could have been saved if a choice system were in place. The high level of no-fault benefits in New York costs relatively little as a percentage of the total premium. The larger percentage of that state's premium is collected to cover suits that may be brought above a threshold designated as "significant bodily impairment," a threshold that precludes about seventy-five percent[29] of all claims for bodily injury. And yet no-fault benefits contributed only thirty-six percent of the total pure premium.[30] In short, the relatively few tort claims preserved by New York's high threshold accounted for a disproportionate amount of total costs.

---

[26] O'Connell & Joost, *supra* note 3, at 70–72, 83–84.

[27] This discussion is not concerned with property damage, which, except in Michigan, is not covered by no-fault insurance.

[28] In Table I, below, col. 3, line 3 ($198.48) minus col. 4, line 3 ($138.42) = $60.38 ÷ $198.48 = .30 (col. 7, line 3). The table is based on work done by Brian A. Smith, research manager, Alliance of American Insurers. A version of the table appears in Smith, *Reexamining the Cost Benefit of No-Fault*, CHARTERED PROPERTY AND CASUALTY UNDERWRITERS JOURNAL, Mar. 1989, at 30. *See also* letter from Brian Smith to Jeffrey O'Connell (Mar. 22, 1989) (on file with the author and with the HARV. J. ON LEGIS.).

[29] Submission from State Farm Insurance Company to Ontario Automobile Insurance Bureau (Apr. 24, 1989) (on file with the HARV. J. ON LEGIS.).

[30] In Table I, $49.78 (col. 5, line 3) ÷ $138.72 (col. 4, line 3) = .36.

Table I: *A Comparison of Estimated 1987 Tort Liability Pure Premiums[a] Without a No-Fault (N-F) Law to Actual Pure Premiums in N-F States*

	1987 Threshold	*1987 N-F Benefit Limits*	*Estimated 1987 Tort Pure Premium*	*1987 BI[b], UM[c] & N-F[d] Pure Premium*	*1987 N-F Pure Premium*	*1987 N-F as % of BI, UM, N-F*	*Change in Injury Coverage Costs in N-F States 1987*	*1982*
VERBAL THRESHOLD								
Florida	Verbal	$10,000	$187.32	$157.45	$41.11	26.1%	−16%[f]	−21%
Michigan	Verbal	Unlimited	171.67	116.57	74.93	64.3	−32	−17
New York	Verbal	50,000	198.48	138.12	49.78	36.0	−30	− 6
THRESHOLD $1,000 OR MORE								
Hawaii	$5,600[e]	15,000	141.49	147.82	62.74	42.4	4	37
Minnesota	4,000	40,000	138.97	112.59	43.88	39.0	−19	− 2
Utah	3,000[e]	14,100	82.22	85.00	22.87	26.9	3	−13
Colorado	2,500[e]	129,925	90.70	131.86	64.22	48.7	45	15
North Dakota	2,500[e]	30,000	66.11	49.81	17.17	34.5	−25	−19
Kentucky	1,000	10,000	93.96	75.06	28.79	38.4	−20	−29
THRESHOLD LESS THAN $1,000								
Georgia	500	5,000	91.32	107.24	39.97	36.3	17	15
Kansas	500[e]	28,925	74.90	58.87	14.64	24.9	−21	−9
Massachusetts	500	2,000	231.70	173.99	21.32	12.3	−25	−33
Connecticut	400	5,000	162.54	170.92	24.02	14.1	5	14
New Jersey	200[e]	unlimited	183.59	226.77	91.17	40.2	24	65
ADD-ON STATES								
Oregon	None	18,380	113.62	110.01	25.26	23.0	− 3	− 8
Delaware	None	10,000	108.56	173.13	53.53	30.9	59	17
Maryland	None	2,500	134.63	170.10	38.73	22.8	26	26
Pennsylvania	None[e]	15,000	118.61	162.78	61.78	38.0	37	53

[a] Pure Premium is that portion of premium used to pay losses, thereby excluding an insurer's expenses in marketing and administrative costs as well as legal defense costs.

[b] BI means tort liability coverage for bodily injury, thereby excluding property damage.

[c] UM means uninsured motorist coverage. See *supra* note 25 and accompanying text.

[d] N-F means no-fault coverage

[e] Threshold was raised between 1982 and 1987. Colorado raised its threshold from $500 to $2,500, effective Jan. 1, 1985. North Dakota raised its threshold from $1,000 to $2,500, effective July 1, 1985. Hawaii's threshold was $1,500 in 1982. Since 1982, the state's tort threshold has been raised several times. Utah increased its threshold from $500 to $3,000, effective July 1, 1986. Pennsylvania eliminated its $750 tort threshold, effective Oct. 1, 1984. New Jersey adopted an optional $1,700 tort threshold, effective July 1, 1984. Kansas raised its threshold to $2,000, effective Jan. 1, 1988.

New York could realize even greater savings if it were to eliminate the tort component of its no-fault coverage and allow motorists to choose between true no-fault insurance and fault-based insurance. Thus, New York could keep its high no-fault benefit level of $50,000 and reduce premiums in the vicinity of seventy-five percent because the portion of the no-fault premium that goes toward claims above the threshold would be eliminated.[31]

Opponents of the no-fault system could argue that providing only $50,000 in no-fault benefits, along with eliminating all tort claims, would unjustifiably lessen the protection provided under current law. They would argue that since some motorists carry more than $50,000 in liability insurance,[32] victims could receive larger recoveries through tort claims than through no-fault benefits. However, this argument ignores the fact that many motorists carry lower levels of liability insurance. In fact, like most states, New York requires that motorists carry only $20,000 in tort liability coverage for bodily injury. Moreover, even with such low limits, approximately twenty percent of motorists in the United States go uninsured, with the percentage greatly in

[31] In Table I, $49.78 (col. 5, line 3) ÷ $198.48 (col. 3, line 3) = .25. This calculation does not include savings stemming from disproportionately higher pure premiums under no-fault than under tort liability since legal defense fees are far lower under no-fault. O'Connell & Joost, *supra* note 3, at 72–75.

All these figures assume that no-fault replaces not only tort liability but Uninsured and Underinsured Motorist coverage. Once insured at limits one deems adequate, one is arguably no longer dependent on coverage from third parties. Only those dependent on payment from third parties need coverage for uninsured motorists. (Admittedly, however, this applies only to economic loss. Noneconomic loss, or "pain and suffering," is not payable by no-fault.)

[32]

Table II: *Distribution of Policy Limits*

Policy Limits (000)	*Bodily Injury*	*Uninsured Motorist**	*Underinsured Motorist**
10/20	4.7%	6.0%	19.0%
15/30	11.2	25.9	10.1
20/40	4.3	8.0	4.5
25/50	17.8	23.0	18.5
50/100	17.6	10.2	13.4
100/300	27.6	13.8	23.2
Over 100,000 Per Person	4.5	1.4	2.5
Other Split Limits	3.6	8.2	4.2
Single Limit	8.6	3.5	4.5
	100%	100%	100%
TOTAL	21,081	3074	357

**See* WIDISS, *supra* note 25 and accompanying text. This table is reprinted from AIRAC, *supra* note 8, at 26.

excess of fifty percent in many inner cities. Payments from whatever policy a motorist carries under a fault-based system are fortuitous and dilatory, depending on the variables of proving tort liability.[33] Thus, a motorist wanting more protection than $50,000 would be better off buying more than $50,000 in no-fault benefits than depending on suing in tort above a threshold.

Still other opponents might object to a choice system because, in using the "inverse liability"[34] device derived from uninsured motorist coverage, a victim of a tortfeasor who has chosen the tort system loses his right to claim against the tortfeasor himself. There are several answers to this. First, any tort recovery is almost always paid not by the tortfeasor himself but from a large pool of impersonal insurance dollars. Second, even under the tort system, claims for motoring accidents are increasingly made not against (or only against) a tortfeasor, but against one's own insurer through Uninsured (or Underinsured) Motorist coverage.[35] Third, under a sophisticated jurisprudential theory, the goal of tort law is served regardless of whether or not damages are paid by a tortfeasor.[36] Finally, empirical evidence suggests that motorists making personal injury claims do not harbor resentments against tortious drivers, nor find assuagement from tort payment.[37]

The availability of devices whereby the motorist can choose, but not be compelled, to buy no-fault insurance arguably changes the terms of the whole no-fault debate.[38] For some

---

[33] The AIRAC study indicates that almost 40% of those collecting no-fault benefits in New York would not have qualified for payment under tort criteria. AIRAC, *supra* note 8, at 150.

[34] *See infra* note 41.

[35] *See* WIDISS, *supra* note 25.

[36] Coleman, *The Structure of Tort Law*, 97 YALE L. J. 1233, 1242, 1250, and *passim* (1988).

[37] O'Connell & Simon, *Payment for Pain and Suffering: Who Wants What, When & Why?* 1972 U. ILL. L. F. 1. If one finds these arguments unconvincing, a possible alternative to the use of the inverse liability device would be to provide, as does the Kentucky no-fault auto statute, that those electing tort insurance never lose their tort rights. Under the Kentucky law, when two no-fault insureds collide, tort rights disappear, but when a fault and no-fault motorist collide, each remains potentially liable to the other in tort (as do, of course, two fault insureds). KY. REV. STAT. ANN. § 304.39-060 (Michie/Bobbs-Merrill 1988). Such an arrangement has the disadvantage of forcing those who elect no-fault coverage to buy considerably more tort liability coverage than under the inverse liability device. However, to the extent that no-fault costs much less than tort (*supra* notes 27–31 and accompanying text), the number of tort liability insureds may be very few.

[38] Versions of this choice bill have been introduced in 1989 in Arizona, H. 2059; California, S. 1232; Maryland, H. 1434; Pennsylvania, H. 1165; and Nevada, S. 520.

twenty-five years legislatures have been baffled by the contradictory claims—actuarial and otherwise—of advocates and opponents of no-fault insurance. The result has been that legislatures have either failed to enact no-fault laws, or they have enacted bastardized, faulty versions that do not further the goals of a pure no-fault regime. A law allowing consumers to choose between no-fault and tort liability means legislators need not bet exclusively on one or the other of the competing coverages and claims therefor. Instead, a legislature can undertake the much less daunting task of letting the two coverages compete one with the other. (This will provide healthy pressure to keep the no-fault system operating effectively. For example, it may help guarantee that no-fault benefits track inflation as do tort verdicts.) It is one thing, after all, for the trial bar and other opponents of no-fault to assert that no-fault is so bad that motorists ought not be *forced* to buy it; it is quite another to assert that it is so bad that motorists ought not be *allowed* to buy it even if they want to. In light of the enviable record of even imperfect no-fault laws, such an extreme contention seems manifestly untenable.

A draft bill follows:

II. A DRAFT BILL (9-19-89)

House (Senate) Bill No. ____[39]

An act, which may be cited as the "Consumer Alternatives Auto Insurance Act."

---

[39] This bill began as a redraft of H.R. 353, 34th Assembly, 1987 Delaware. While it purports to be thorough in those matters it deals with, the bill does not purport to be exhaustive in what might be covered, such as the following:

As to the arguable windfall under no-fault insurance to owners of trucks and other commercial vehicles following, for example, truck-car collisions, that is dealt with effectively in NATIONAL CONFERENCE OF COMMISSIONERS ON UNIFORM STATE LAWS, HANDBOOK OF THE NATIONAL CONFERENCE OF COMMISSIONERS ON UNIFORM STATE LAWS, UNIFORM MOTOR VEHICLE ACCIDENT REPARATIONS ACT, §§ 38, 39 and commentary thereto, at 331–39 (1972) [hereinafter UMVARA]. This problem is handled more simply, if less elegantly, under a proposal made for the state of New York in STATE OF NEW YORK INSURANCE DEPARTMENT, AUTOMOBILE INSURANCE . . . FOR WHOSE BENEFIT? A REPORT TO GOVERNOR NELSON A. ROCKEFELLER 90–91, 115 (1970). An even simpler solution would be just to allow subrogation by private passenger car no-fault payors against trucks and commercial vehicles. *See generally* J. O'CONNELL & R. HENDERSON, TORT LAW, NO-FAULT & BEYOND 402–15 (1975). Note this matter is different from how occupants of trucks and commercial vehicles are to be dealt with. That issue is handled in subparagraphs (1)(c) and (2)(g) below.

In many ways, the reverse of the truck-car collision is presented by the car-motorcycle

BE IT ENACTED BY THE GENERAL ASSEMBLY OF THE STATE OF _____:

Section 1. (a) No owner of a motor vehicle required to be registered in this State shall operate or authorize any other person to operate the vehicle unless the owner has insurance on the motor vehicle providing either the fault insurance alternative under paragraph (1) below or the no-fault insurance alternative under paragraph (2) below.[40] The choice between such coverages is applicable to every motor vehicle of the owner. In the event an owner of more than one vehicle chooses different alternatives, the earliest choice governs, and in the event of simultaneous choices, the choice of the fault insurance alternative governs.

---

collision in that in the latter it is the car that disproportionately inflicts the damage (as opposed to occupants of cars suffering disproportionate personal injury losses in a collision with a truck.) The tremendous exposure of motorcyclists to personal injury (whether in a collision with a car or in any other type of accident) means that switching to first-party coverage (whether fault- or no-fault based) will cause an exponential rise in motorcyclists' personal injury premiums. One solution would be to simply exempt motorcycles from the choice system, *i.e.*, motorcyclists can sue and be sued in tort after collisions with no-fault as well as with fault-based insureds.

[40] Although compelling people to purchase insurance can jeopardize affordable rates and rankle consumers who would prefer to decide for themselves what to do with their money, those who operate the most dangerous machines widely available in our society should arguably not be free to neglect carrying insurance. The long history of compulsory auto insurance throughout the Western world, and its recent extensive adoption in the United States indicate a long-burgeoning awareness of the necessity of requiring drivers to purchase insurance against the possible consequences of operating a motor vehicle. *See* R. KEETON & J. O'CONNELL, BASIC PROTECTION FOR THE TRAFFIC VICTIM, 76–102, 189–217 (1965). *See generally* AMERICAN INSURANCE ASSOCIATION, SUMMARY OF SELECTED STATE LAWS AND REGULATIONS RELATING TO AUTOMOBILE INSURANCE (1983).

And yet, this requirement is not without its drawbacks. Today, insurance rates must not be merely actuarially sound, they must also be affordable. These two objectives, however, are sometimes irreconcilable. Compelling people to act against their own economic wishes simply exacerbates the problem, as consumers resent being forced to divert their resources to compulsory insurance instead of food, clothing, housing or other more tangible necessities. The result may be huge numbers of uninsureds, along with unrealistic, irresponsible demands for the artificial lowering of rates by regulatory authorities, legislatures, or even the Initiative. On the purported drastic reduction in casualty insurance rates by Initiative in California (fueled by rapidly rising rates), see Wall St. J., Nov. 10, 1988, at B1, col. 3. *See also* N.Y. Times, Nov. 6, 1988, at F1, col. 2; for subsequent development, see N.Y. Times, Nov. 23, 1989, at 9, col. 4.

These are powerful points. But one can at least question the arguably regressive step of abandoning a requirement that those who operate dangerous motor vehicles accept the responsibility for insuring against death and damage caused by their operation. For more on the subject of compulsory coverage see *infra* notes 43, 48, 63 and accompanying text.

(1) Fault Insurance Alternative.

a. Fault insurance covers each person who is a named insured under the motor vehicle policy or his spouse, or other relative or dependent residing in the same household as the named insured, in addition to anyone operating the vehicle with permission of the owner, for legal liability for bodily injury, including death, arising out of ownership, maintenance, or use of the vehicle to the limit, exclusive of interest and costs, of at least the limits prescribed by the Financial Responsibility [Compulsory Insurance] Law of this State.

b. Inverse liability[41] insurance provides compensation applicable under the fault insurance alternative to those listed under subparagraph c below if injured in a motor vehicle accident as a consequence of the fault of another person who is insured under the no-fault insurance alternative in paragraph (2) below.

c. The inverse liability coverage provided under subparagraph (1)b above is applicable to:

1. **each person who is a named insured under a policy issued under the fault insurance alternative;**
2. **each person who is the spouse of the named insured, or other relative or dependent of the named insured residing in the same household as the named insured, under a policy issued under the fault insurance alternative unless under the terms of subparagraph (2)s below that person is covered under a policy issued under the no-fault insurance alternative. But when both fault and no-fault alternatives are deemed equally applicable to a person purported to be bound by the choice of another, the applicable alternative is that applicable to the vehicle of the named insured in which the person was injured or, if none, the [no-fault] alternative; and**
3. **each person who is an occupant of a motor vehicle insured under an insurance policy issued under the fault insurance alternative other than a person who is covered under a policy issued under the no-fault insurance alternative.**

d. Priority of inverse liability coverage:

1. **In case of injury to the driver or other occupant of a motor vehicle, if the accident causing the injury occurs while the**

[41] The term "inverse liability" has been substituted for the term "connector" originally used in O'Connell & Joost, *supra* note 3, at 78. "Inverse liability" is a term often applied to a coverage, like Uninsured Motorist coverage, whereby an insurer pays its own insured based on a third party's tort liability. *See* WIDISS, *supra* note 25.

vehicle is being used in the business of transporting persons or property, the inverse liability coverage for payment of damages is the coverage applicable to the vehicle, or, if none, the coverage under which the injured person is insured by his or another's choice under subparagraph c above or (2)s below.

2. In case of injury to an employee, his spouse, or other relative or dependent residing in the same household, if the accident causing the injury occurs while the injured person is driving or occupying a motor vehicle furnished by the employer, the inverse liability coverage for payment of benefits is the coverage applicable to the vehicle, or, if none, the coverage under which the injured person is insured by his or another's choice under subparagraph c above or (2)s below.

3. In all other cases, the inverse liability coverage for payment is the coverage under which the injured person is insured by his or another's choice under subparagraph c above or (2)s below or, if none, the coverage applicable to the vehicle in or by which he is injured.[42]

e. The insurance coverages provided under this paragraph (1) are subject to such terms, conditions, and exclusions as are approved by the [Commissioner] of Insurance.

(2) No-Fault Insurance Alternative.

a. No-fault insurance provides compensation to persons injured in a motor vehicle accident within the United States, its territories or possessions, and Canada for reasonable and necessary expenses or losses [incurred within [2] [10] years][43] **after the accident as a**

[42] Under the inverse liability coverage, then, the coverage follows the family, not the car. For more on this, see *infra* note 51.

[43] The choice of the duration of coverage is obviously tied to its amount. On this point, see UMVARA, *supra* note 39, § 34 and commentary thereto, at 321–23. But even with high limits, some in the insurance industry would urge the use of a two-year coverage limit, which is the time within which almost all medical bills are incurred.

Coverage periods longer than two years pose practical problems for the industry. First, it is expensive to maintain records for long periods of time and, without records, it is hard to deal with late-blooming claims. Second, it is all too easy for doctors to finance today's medical expenses by attributing them to an automobile accident that happened many years ago. From the industry's viewpoint, it is one thing to invite those problems voluntarily, but it is another matter to be compelled by statutory fiat to provide coverage so broad and so potentially troublesome.

On the other hand, auto accidents are unique in their propensity to inflict widespread catastrophic losses. Thus, society may be justified in calling for uniquely extensive coverage. This argument could limit such coverage to paraplegia, quadriplegia, severe brain damage and other defined severe injuries, as those long-term effects are indisputable. Those who want and can afford a longer coverage period for all injuries would presumably be able to purchase it from their agents, if market demands exist.

consequence of the accident, to at least the limits prescribed by this paragraph (2). Compensation is provided for the following items in the following way:

1. Reasonable charges for reasonably needed medical, hospital, dental, surgical, medicine, x-ray, ambulance, emergency medical, rehabilitation,[44] nursing, and funeral services. Compensation for funeral services, including all customary charges and the cost of a burial plot for one person, is not to exceed the sum of [$3000]. Compensation includes reasonable expenses for any non-medical remedial care and treatment rendered in accordance with a recognized religious method of healing. Compensation does not include any charge for a room in any facility engaged in providing any such service to the extent that the charge is in excess of a reasonable and customary charge for a semiprivate room, unless a private room is medically required.
2. Subject to a maximum of [$4000] per month in compensation for lost earnings, compensation for eighty percent of lost earnings, net after taxes, sustained by each such injured person [during the first thirty-six months] after a motor vehicle accident [and, subject to a maximum of [$2000] per month, forty percent of the lost earnings so sustained during the subsequent thirty-six months].[45] "Lost earnings" include lost earnings, net after taxes and expenses, of a self-employed person.
3. If death is proximately and directly caused by a motor vehicle and occurs within one year of the date of the accident, the lost earnings benefit to which the injured person would have been entitled had he or she survived is payable as a survivor's benefit to the heirs or estate of the decedent.
4. Reasonable amounts required to pay for services which would have been performed by the injured person had that person not been injured in a motor vehicle accident, up to $25 per day for the first thirty-six months after such an accident.

b. The term "injury" does not include damage to property; the term "injury" includes death; the term "injured person" includes the personal representative of an estate; a "cause of action for

[44] For more extensive model provisions governing rehabilitation, see UMVARA, *supra* note 39, § 34 and commentary thereto, at 321–23.

[45] The bracketed material reflects the legislative option for more than two years' coverage under subsection (2)a.

injury" means any claim for injury for both economic and noneconomic loss based on another's faulty conduct or product.[46]

c. Except as provided in subparagraph d below, no-fault insurance benefits are payable without regard to fault but are not payable for any noneconomic loss which term includes, but is not necessarily limited to, pain, suffering, inconvenience, physical impairment, grief, and other nonpecuniary damages recoverable under the law in this State in a cause of action for injury.

d. A person intentionally causing or attempting to cause injury to himself or another person is disqualified from no-fault benefits for injury arising from his acts, including benefits otherwise due him as a survivor. If a person dies as a result of intentionally causing or attempting to cause injury to himself, his survivors are not entitled to no-fault benefits for loss arising from his injury. A person intentionally causes or attempts to cause injury if he acts or fails to act for the purpose of causing injury or with knowledge that injury is substantially certain to follow. A person does not intentionally cause or attempt to cause injury (1) merely because his act or failure to act is intentional or done with the realization that it creates a grave risk of causing injury or (2) if the act or omission causing the injury is for the purpose of averting bodily harm to himself or another person.[47]

e. 1. The minimum insurance coverage which satisfies the requirements of the no-fault insurance alternative is a limit of [$15,000] [$100,000][48] for the total of all payments

---

[46] *But see infra* note 58.

[47] UMVARA, *supra* note 39, § 22 and commentary thereto, at 306–07.

[48] On the subject of compulsory no-fault limits see *supra* note 43 and *infra* notes 62, 63 and accompanying text. It is much easier simply to eliminate all tort claims for those choosing no-fault if the no-fault limits are so high that very few accident victims covered by no-fault will suffer higher losses than their coverage. O'Connell & Joost, *supra* note 3, at 72–75. The table *supra* in note 32 indicates that only about five percent of motorists are insured under liability insurance for over $100,000 per person. That would mean that providing $100,000 of no-fault benefits would disadvantage very few injury victims, albeit admittedly the most severely injured. However, the tort system makes it very unlikely that one will be injured by a motorist who is not only at fault but also carrying more than $100,000. Therefore, generally speaking, it is not nearly as favorable to the severely injured as a $100,000 no-fault benefit package.

There is a great deal of resistance from insurers to unlimited exposure under auto no-fault. Yet, a 1978 Michigan Insurance Bureau Report found that there was only a nominal cost involved in providing unlimited no-fault benefits:

> A $100,000 ceiling [on Michigan's no-fault medical benefits, as opposed to the unlimited benefits provided under the law] since the inception of no-fault would have saved only about $6 per car per year, based on claim reserves of the state's six largest insurers. One large insurer estimates the cost of paying claims in excess of $100,000 as no more than the cost of insuring against the theft of CB radios. Providing protection against the catastrophic loss represents the essence of insurance. To attempt to place the burden of catastrophic losses on

which must be made pursuant to that alternative for every injured person so covered.

2. Each insurer authorized to sell no-fault insurance may offer no-fault coverage within monetary or territorial limits in excess of required no-fault coverages. The [Commissioner] of Insurance may adopt rules that such excess no-fault coverages, in amounts specified by the [Commissioner] of Insurance, be offered by insurers writing no-fault insurance.

f. The coverage provided under the no-fault insurance alternative is applicable to:

1. each person who is a named insured under a motor vehicle liability insurance policy issued under the no-fault insurance alternative;

2. each person who is the spouse of the named insured, or other relative or dependent of the named insured residing in the same household as the named insured, under a policy issued under the no-fault insurance alternative, unless under the terms of subparagraph s below that person is covered under a policy issued under the fault insurance alternative, and subject further to the terms of clause (1)c2 above.[49]

g. Priority of no-fault coverages:

1. In the case of injury to the driver or other occupant of a motor vehicle, if the accident causing the injury occurs while the vehicle is being used in the commercial business of transporting persons or property, the coverage for payment of no-fault benefits is the coverage applicable to the vehicle, or, if none, the coverage under which the injured person is

---

the individuals suffering the loss, when its average cost is only a few dollars, is inconsistent with the provision of insurance.

INS. BUREAU, MICH. DEP'T OF COMMERCE, NO-FAULT INSURANCE IN MICHIGAN: CONSUMER ATTITUDES AND PERFORMANCE 76 (1978). But for contrary views on the effect of high no-fault limits, see U.S. DEP'T OF TRANSP., STATE NO-FAULT AUTOMOBILE INSURANCE EXPERIENCE: 1971–77, at 64 (1977); O'Connell & Joost, *supra* note 3, at 74, 80, 85 nn.42, 55, 65. *See also* DOT REPORT, *supra* note 3, at 125–28.

If, as is often indicated (N.Y. Times, *supra* note 2), lowering insurance costs is the name of the insurance reform game, the bill could match the no-fault benefit level with the applicable state compulsory minimum tort liability.

For a means of dealing with serious losses under either high or low no-fault benefits, see *infra* notes 63 - 69 and accompanying text.

[49] In addition to precluding an insured's choice from necessarily binding an adult member of his family, this provision does not permit the insured's choice of coverage to bind those who are not members of a car-owning family, whether they are guests, passengers, or pedestrians injured in or by the insured's car. Any necessary tort liability coverage is provided for such a case in subparagraph (2)h below. *See* O'Connell & Joost, *supra* note 3, at 80–81 n.56.

insured by his or another's choice under subparagraph f above.

2. **In the case of injury to an employee, an employee's spouse, or other relative or dependent residing in the same household, if the accident causing the injury occurs while the injured person is driving or occupying a motor vehicle furnished by the employer, the coverage for payment of no-fault benefits is the coverage applicable to the vehicle, or, if none, the coverage under which the injured person is insured by his or another's choice under subparagraph f above.**[50]
3. **In all other cases, the coverage for payment of no-fault benefits is the coverage under which the injured person is insured by his or another's choice under subparagraph f above, or, if none, the coverage applicable to the vehicle in or by which the person is injured.**[51]

---

[50] This is so regardless of whether the employee is acting within the course of his employment at the time of injury. *See* UMVARA, *supra* note 39, § 4 and commentary thereto, at 277–79.

Under this subparagraph (2)g (as well as under the companion subparagraph (1)d), an attempt is made to internalize the cost of auto accidents to commercial vehicles in the form of common carriers (clause 1) ("carriers") and employer-supplied vehicles (clause 2). Unless internalization is imposed on common carriers, such vehicles covered by no-fault will have a windfall when their passengers are also covered by no-fault. The passengers, being precluded from claiming in tort, would have their no-fault claims paid by their own insurers exempting the carriers from *any* liability to such passengers. But under subparagraph (2)g, no-fault claims are made against the carrier, thus internalizing those costs.

However, the no-fault carrier *will* be able to externalize the losses of fault-insured passengers, since under subparagraph (1)d, such passengers will be claiming under their own inverse liability insurance coverage. The no-fault carrier, therefore, avoids paying tort claims. There is a compensatory mix of advantages and disadvantages when the fault-insured carrier collides with a no-fault insured car. In that situation, the no-fault insured occupant of the fault-insured vehicle collects against his own insurer (otherwise the fault-insured party would be required to pay no-fault benefits), and the fault-insured occupant claims against the liability coverage of the fault-insured car.

There is the same compensatory mix of advantages and disadvantages with reference to claims against an employer-provided car under clauses (1)d2 and (2)g2 in considering the priorities of no-fault and fault-insured occupants' claims against no-fault and fault-insured cars. That is, when a fault-insured employer-supplied car collides with a no-fault-insured car, the no-fault-insured occupant of the fault-insured car will collect against his own no-fault coverage, but the fault-insured occupant will claim against the fault-insured car's inverse liability coverage.

[51] Subparagraph g determines which no-fault coverage will provide compensation to a person qualified to receive benefits. The underlying principle, as set out in clause 3 of subparagraph g, is that a person suffering a loss should make his claim for benefits against his own insurer. In effect, then, under both the inverse liability and no-fault coverages, the insurance follows the driver (and his family), not the car (except for commercial and employer-provided cars under clauses 1 and 2, respectively.) *See* DOT REPORT, *supra* note 3, at 137; *see also* R. KEETON & J. O'CONNELL, *supra* note 40, at 370–79.

There are two sets of trades tied to being covered under a no-fault system versus being covered under a fault system. First, one trades the assurance of payment for

h. The owner of a motor vehicle may elect, but cannot be required, to have the coverage described under the no-fault insurance alternative in subparagraphs a and b above written subject to deductibles, exclusions, waiting periods, sublimits, percentage reductions, excess provisions, and similar reductions approved by the [Commissioner] of Insurance. Except as provided in subparagraph o, such elections apply only to benefits payable to the named insured, the insured's spouse, or other relative or dependent residing in the same household.[52]

i. Timing of payment:

1. **For purposes of this subparagraph i, loss accrues not when injury occurs, but as medical expense, work loss, replacement services loss, or rehabilitation services expense is incurred.**
2. **Expenses under paragraph (2) shall be submitted to the insurer as promptly as practical, in no event more than two years after the loss accrues.**
3. **Payment of expenses submitted under paragraph (2) shall be made monthly as loss accrues. Benefits are overdue if not paid within 30 days after the insurer receives reasonable proof of the fact and amount of the loss realized, unless the insurer elects to accumulate claims for a period not exceeding 31 days and pays them within 15 days after the period of accumulation. If reasonable proof is supplied as to only part of a claim, and the part totals $100 or more, the part is overdue if not paid within the time provided by this clause.[53]**
4. **Overdue payments bear interest at the rate of [18] percent per annum.[54]**

---

economic losses for the chance of higher payments under the tort system. Second, one trades the right to assert claims for the benefit of not being claimed against in tort. Those who are ordinarily or solely passengers in cars (such as young children or others who drive rarely or not at all) benefit from the first, but not the second, trade. They would not be claimed against anyway. To that extent, the trade involved in substituting no-fault for fault, disadvantages, relatively speaking, such minors or other non-drivers. This is true, however, under any no-fault auto scheme regardless of whether it is a choice system. There seems to be no other effective way of achieving no-fault; it would be too difficult to distinguish those whose tort claims are to be preserved from those who are to be precluded from asserting tort claims.

[52] Under this provision, no-fault benefits could be limited to payment in excess of all, or most, collateral sources. *See* O'Connell & Joost, *supra* note 3, at 81 n.57. For a draft of a provision comprehensively deducting collateral sources (even term life insurance), see O'Connell, *A Proposal to Abolish Contributory and Comparative Fault with Compensatory Savings by Also Abolishing the Collateral Source Rule*, 1979 U. ILL. L. F. 591, 601–02. For a more restricted provision, see UMVARA, *supra* note 39, § 11 and commentary thereto, at 291–93.

[53] UMVARA, *supra* note 39, § 23 and commentary thereto, at 307–09.

[54] *Id.*

5. Benefits for medical expenses and rehabilitation services expenses may, at the election of the insurer, be paid directly to the persons supplying such services to the claimant.
6. A claim for expenses under paragraph (2) shall be paid without deduction for benefits payable by collateral sources under subparagraph h above, notwithstanding specific provisions of the laws of this State which require the amount of such benefits to be subtracted from the insurer's obligation, if these benefits have not been paid to the claimant before the benefits are overdue or the claim is paid. The insurer is entitled to reimbursement from the person obligated to make the payments or from the claimant who actually receives the payments.
7. An insurer may bring an action to recover benefits which are not payable, but are in fact paid, because of an intentional misrepresentation of a material fact upon which the insurer relied, by the insured or by a person providing medical or rehabilitation services. When medical or rehabilitation services are at issue, the action may be brought only against the person providing the services, unless the insured intentionally misrepresented the facts or knew of the misrepresentation. An insurer may offset amounts it is entitled to recover from the insured under this subsection against any benefits otherwise due.
8. An insurer who rejects a claim for benefits under paragraph (2) shall give to the claimant prompt written notice of the rejection, specifying the reason.[55]

j. The coverage provided under the no-fault insurance alternative is applicable to accidents involving a motor vehicle in any state of the United States, its territories or possessions, or Canada.[56]

k. The coverage provided under the no-fault insurance alternative includes liability coverage for payment up to at least (i) [$15,000] for bodily injury to any person and [$10,000] for property damage, as a result of a motor vehicle accident occurring in any other state of the United States, its territories or possessions, or Canada, if caused by the fault of an insured under the no-fault policy or one for whom he is vicariously liable;

---

[55] *See id.* at 96–98.

[56] *See* subparagraph (2)u *infra*. Concerning the possibility of covering accident victims anywhere in the world, see UMVARA, *supra* note 39, § 2 and commentary thereto, at 275–77.

(ii) [$10,000] for property damage as a result of a motor vehicle accident within this state, or outside the state as defined in (i) above, if caused by the fault of an insured under the no-fault policy or one for whom he is vicariously liable; and (iii) at the insured's option, [$15,000] for bodily injury to any person as a result of a motor vehicle accident within this state if the victim is eligible to claim liability for bodily injury caused by the fault of an insured under the no-fault policy or one for whom he is vicariously liable. The minimum limits for bodily injury in this subparagraph k are further subject to a limit of [$30,000] for all damages arising out of one accident.[57]

l. Any person eligible at the time of an accident for benefits under the no-fault insurance alternative may neither claim against, nor be claimed against by, any other person insured, or required to be insured under either the fault or no-fault insurance alternative, including any other person for whom such insured is vicariously liable, for liability based on fault for an accident occurring within this state arising out of the ownership, maintenance or use of a motor vehicle, except as provided in subparagraph k above or subparagraphs m, n, and o below, or for harm intentionally caused as defined by subparagraph d above.[58]

m. A victim of an accident has a cause of action for injury against any party convicted of driving under the influence of alcohol or illegal drugs or guilty of intentionally causing harm as defined in subparagraph d above, when such violation or conduct caused or substantially contributed to the accident.[59]

1. If the party thus convicted is insured under no-fault insurance, the victim has a right, in the alternative, to claim against the convicted driver for no-fault benefits up to the amount specified in the policy applicable to the convicted driver (irrespective of whether no-fault benefits are payable

---

[57] This subparagraph provides for (a) compulsory liability insurance for (i) personal injury and property damage inflicted out-of-state and (ii) property damage in or out-of-state, (neither (i) nor (ii) being affected by no-fault coverage), and (b) voluntary liability insurance for personal injury inflicted in-state. On this latter lack of compulsory liability insurance, see *infra* note 63 and accompanying text.

[58] As to the definition of intentional injury, see *supra* note 47 and accompanying text. It should be noted also that no-fault motorists lose only their right to claim against other motorists, not against non-motorists causing injury, such as car manufacturers producing defective cars or railroads negligently colliding with cars.

[59] This provision preserves the right of victims of particularly egregious conduct to claim in tort.

to the convicted driver himself), plus a reasonable attorney's fee as provided under subparagraph p.[60]

2. Any party providing other coverage to the victim, including inverse liability coverage, has a right of subrogation for a claim under this subparagraph m.

n. A victim of an accident caused by a party insured under no-fault insurance, in which the victim is uninsured for either fault or no-fault insurance solely by virtue of not being a member of a car-owning family, may claim against the party covered by no-fault insurance for either no-fault benefits or tort liability damages.[61]

o. A victim of an accident not alleging a cause of action for injury under subparagraph m, where the injury is caused by a party insured under no-fault insurance, has the right to bring a cause of action for injury irrespective of subparagraph l above for damages in excess of the applicable inverse liability or no-fault coverage available to the victim.[62] But when a claim is made under this subparagraph o, if the party insured under no-fault insurance provides the victim, within ninety days after either the accident or the filing of the victim's claim, with a written tender to pay the equivalent of no-fault benefits covering economic loss in excess of any other coverage available to the victim, plus a reasonable attorney's fee as provided under subparagraph p, the victim is foreclosed from pursuing the claim any further.[63] However, the

---

[60] A victim eligible to claim in tort, but not wanting to undergo prolonged tort litigation, can claim for more prompt, but lesser, no-fault benefits.

[61] *See* O'Connell & Joost, *supra* note 3, at 80–81 n.56.

[62] To the extent that a large majority of motorists elect no-fault insurance with limits of $100,000 (*see supra* note 48 and *infra* note 63 and accompanying texts), this provision will rarely be applicable. The provision allows members of a motoring family whose motorist provided either no or low amounts of no-fault or tort coverage to recover. This would supply redress, for example, to a poor child seriously injured by tortious conduct of an affluent no-fault insured.

[63] This provision is designed to deal with the problem of larger losses above either inverse liability or no-fault coverages. For the origins of this "early offers" approach, whereby full-scale tort liability can be avoided, see O'Connell, *Offers That Can't Be Refused: Foreclosure of Personal Injury Claims By Defendant's Prompt Tender of Claimants' Net Economic Losses*, 77 NW. U.L. REV. 589 (1982). Under this approach, defendants are encouraged, but not required, to provide expeditiously an offer to pay benefits covering only net economic loss above collateral sources, rather than spending precious resources litigating fault and the value of noneconomic loss.

Note that there is no limit on the amount of benefits for economic loss which must be offered under this provision. This would mean that when a defendant-motorist's insurance limits are relatively low compared to his victim's net economic loss, an insurer will not be inclined to make "an early offer" covering economic loss. However, in the large number of cases where losses are substantially less than the defendant's coverage, an early offer becomes feasible.

Another possibility would be to authorize an offer of the equivalent of whatever no-

victim may pursue the claim if it can be proven beyond a reasonable doubt that the party insured under no-fault insurance, either

---

fault benefits the offeror carries, subject to a minimum of $100,000, to the victim to cover economic loss in excess of collateral sources. Such an offer, though not unlimited in amount, would foreclose pursuit of an ordinary tort claim. This coverage, being in excess of collateral sources, means that payment of collateral sources will not count towards exhausting the $100,000 limit. This change would be accomplished by changing the bill to read "tender to pay the equivalent of no-fault benefits with benefits of $100,000, or any higher no-fault limits applicable to the name insured, covering economic loss in excess of any other coverage."

An "early offers" approach is deemed better, for example, than just allowing an injury victim to claim in tort for only economic loss (and not for pain and suffering) above his no-fault or inverse liability coverage. Under the latter approach, a defending insurer is under a strong incentive to resist, and delay payment of a tort claim for economic loss, knowing that its exposure is thus limited. This is a common complaint under tort claims for property damage against less responsible insurers when they similarly face no exposure to payment of noneconomic loss. Under the "early offers" approach, an insurer must *earn* the right to pay a tort claimant only economic loss, by promptly (within 90 days) offering to do so. On the other hand, a defendant with either no liability or very doubtful liability—or no or low tort liability insurance—would not be inclined to make an early offer to evade full-scale tort liability.

Under the bill, motorists electing no-fault insurance are not required to carry liability insurance. Thus a poorer person with no or few assets to protect can buy only no-fault insurance protecting himself and his family for their medical bills and any wage loss. This does this not overly disadvantage those injured by the poor since the poor are so likely to be either un- or underinsured anyway.

In order to avoid requiring the poor to buy more insurance than they can readily afford (and thus perversely encouraging them to go completely uninsured), the bill could allow motorists to buy either $15,000 *or* $100,000 in no-fault benefits and thereby meet the mandatory minimum insurance requirement. One could then further provide that those buying only a $15,000 limit remain liable in tort for losses above any applicable coverage to those whom they tortiously injure, without the right to limit their exposure to tort liability by making early offers to pay net economic loss. Recall, however, there is no requirement that tort liability insurance be carried. This would mean the poor, prompted to buy only $15,000 in no-fault coverage in order to save money, need not worry any more than they do now about no or low tort liability coverage to protect their limited or nonexistent assets. But the more affluent *would* be prompted to buy $100,000 in no-fault coverage to get such no-fault coverage plus the benefit of making early offers covering economic losses of those they injure, with a concomitant lowering of their cost of excess liability coverage.

If a no-fault insured buying full coverage of $100,000 is liable for losses in excess of inverse liability, no-fault or other coverage (subject to being able to make early offers), the person buying $15,000 of either liability (including inverse liability) or no-fault coverage, or the dependents of those buying no insurance at all, will be able to assert claims against no-fault insureds more often than those buying fuller no-fault coverage. The justification for this situation is simply one justifying income redistribution generally. *See* O'Connell, *A Proposal to Abolish Defendants' Payment for Pain and Suffering in Return for Payment of Claimants' Attorney's Fees*, 1981 U. ILL. L. REV. 333, 356–58. Those buying only minimum coverage (under either tort, including inverse liability, or no-fault coverage) will be the poor. Under liability insurance, the poor, with either no or low liability insurance, are thereby in a position to draw more from the pool of liability insurance dollars than they pay in. However, the reluctance of the poor to invoke the legal process, combined with the relatively smaller wage loss they suffer, results in comparatively exiguous payment to them from liability insurance. Whereas the proposed plan preserves the theoretical advantage of the poor to claim in tort, it also improves the condition of the affluent by guaranteeing those who buy no-fault insurance that they will be covered up to high limits. In addition, the tort exposure of those buying $100,000 of no-fault will most often be only for net economic loss payable

himself or vicariously, was guilty of wanton conduct in causing the accident. But if such tender is not made, the victim not only retains his right to pursue a claim for injury by proof of lack of ordinary care by a preponderance of evidence,[64] but shall not have his claim [defeated] [diminished] by reason of [contributory] [comparative] negligence or assumption of risk except where the victim intentionally caused or intended to cause injury to himself or to another person for injury arising from his acts, intent to injure being defined as in subparagraph (2)d above. But evidence of the victim's failure to exercise reasonable care for his own safety may be relevant to the issue of whether the defendant failed to exercise reasonable care.[65]

1. Denial of defenses based on a victim's conduct does not apply to any damages accruing to any person who unreasonably refuses, either himself or through one legally empowered to act on his behalf, to accept medical care, rehabilitation, rehabilitative occupational, or other medical treatment and care if the procedure, treatment or training is reasonable and appropriate for the particular case and its cost is reasonable in relation to its probable beneficial effects.[66]

2. No owner of a motor vehicle failing to have insurance in effect in violation of subsection 1(a) is eligible to claim under this subparagraph o.[67]

p. No-fault benefits under subparagraphs m, n, and o include reasonable expenses incurred by the victim in collecting such

---

periodically, because of their capacity to make early offers. Thus, the temptation of tort claimants under normal tort liability to pad smaller claims will largely disappear. *See* AIRAC, *supra* note 701, at 13. For more on this phenomenon of "padding" of claims, see O'Connell, *supra*, 1981 U. ILL. L. REV. 333, 334–40.

[64] This provision lessens the absolute foreclosure of tort claims and allows an accident victim to pursue a tort claim for egregious conduct other than intentional conduct or driving under the influence of alcohol or drugs pursuant to subparagraph m. *See supra* note 48 and accompanying text.

[65] *See, e.g.*, Lorenzo v. Wirth, 170 Mass. 596, 49 N.E. 1010 (1898); Davis v. Consolidated Rail Corp., 788 F.2d 1260 (1986).

[66] Under this provision, a *quid pro quo* for allowing a defendant to make an early offer to pay only economic loss, thereby cutting off a claim for more, is that the defendant loses the defense of the claimant's fault. For the justification of this exchange and further exegesis concerning it, see O'Connell, *A Proposal to Abolish Contributory and Comparative Fault, with Compensatory Savings by Also Abolishing the Collateral Source Rule* 1979 U. ILL. L. F. 591–600. The precedent for ignoring a claimant's faulty conduct in his personal injury claim can be traced to the Employer's Liability Acts passed as a precursor to workers' compensation acts. PROSSER & KEETON, TORTS § 80, at 568–80 (5th ed. 1984).

[67] The statute might provide that a motorist failing to get either fault or no-fault insurance could claim like others but would be subject to a deductible amounting to the minimum coverage foregone, *i.e.*, $15,000.

benefits, including a reasonable attorney's fee. Such expenses may be deducted from the amount of no-fault benefits otherwise provided, if any significant part of the claim is fraudulent or so excessive as to have no reasonable foundation.[68]

q. No subtraction is made against no-fault benefits due because of the value of a cause of action for injury, including under inverse liability coverage, in a claim under subparagraph o, except that after recovery is realized under such a cause of action, including under inverse liability coverage, a subtraction is made to the extent of the net recovery, exclusive of reasonable attorneys' fees and other reasonable expenses incurred in effecting the recovery. If no-fault benefits have already been received, the claimant shall repay to the insurer paying no-fault benefits out of the recovery a sum equal to the benefits received but not more than the realized net recovery, and the insurer shall have a lien on the recovery to this extent. Any remainder of the net recovery from such a cause of action, including under inverse liability coverage, applies periodically against loss as it accrues, until an amount equal to the net recovery under such a cause of action, including under inverse liability coverage, has been subtracted.[69]

r. Rights over:

1. Whenever a person who receives or is entitled to receive benefits under the no-fault insurance alternative has a claim or cause of action against any other person for breach of an obligation or duty causing injury, the insurer of the applicable no-fault benefits is subrogated to the rights of the claimant and has a claim for relief or cause of action separate from that of the claimant, to the extent that (i) elements of damage compensated by no-fault insurance are recoverable and (ii) the insurer of the applicable no-fault benefits has paid or become obligated to pay accrued or future no-fault insurance benefits.

2. An insurer of applicable no-fault insurance has a right of indemnity against a person who has converted a motor vehicle involved in an accident, or a person who has intentionally caused injury to a person, for no-fault insurance benefits paid to other persons for the injury caused by the conduct of that person, for the cost of processing a claim for those

---

[68] *See* UMVARA, *supra* note 39, § 24 and commentary thereto, at 309–10.

[69] This provision coordinates tort and no-fault benefits in cases covered by subparagraph o. For its origins, see R. KEETON & J. O'CONNELL, *supra* note 40, at 307, 402–04.

benefits, and for reasonable attorneys' fees and other expenses for enforcing the right of indemnity. For purposes of this clause, a person is not a converter if he uses the motor vehicle in the good faith belief he is legally entitled to do so.[70]

s. Elections:

1. Each insurer prior to the initial issuance of a motor vehicle policy after the effective date of this Act shall notify in writing the owner of the vehicle of the availability of the two alternatives of fault insurance and no-fault insurance, pursuant to a form and manner consistent with regulations of the [Commissioner] of Insurance. These regulations shall set forth the written terms of the document under which a party elects to bind himself, and anyone he has a right to bind by his election, to the fault or no-fault insurance alternative.

2. A spouse, relative or dependent of a named insured under the no-fault alternative who is *sui juris* is under a duty to file a written notification with the company issuing the motor vehicle liability policy, in a form and manner consistent with reasonable regulations of the [Commissioner] of Insurance, within 30 days of the issuance of the policy, that he elects to be covered under a coverage of the type specified under the fault insurance alternative, in order to be so covered. Such election takes effect 10 days after its receipt by the company. A company shall provide such fault insurance for those making such an election.

3. Any party signing, or otherwise bound by, a document containing terms of an election between fault and no-fault alternatives approved by the [Commissioner] of Insurance is held to the terms of such election and precluded from claiming liability of any party based on being inadequately informed in making the election between the fault or no-fault alternatives.[71] Any such election continues in force for at least the term of the policy and as to subsequent renewal or replacement policies unless the company issuing the policy receives notice to the contrary consistent with the above terms of this subparagraph s.

4. If any person fails to indicate, prior to a motor vehicle accident, whether he wishes to maintain the fault or no-fault

---

[70] UMVARA, *supra* note 39, § 6, and commentary thereto, at 284–86. Clause (2)r1 would apply, for example, when a no-fault insured is injured out-of-state where the tort exemption under subparagraph (2)l does not apply.

[71] *See* O'Connell & Joost, *supra* note 3, at 87 n.69.

alternative, he, and anyone he is empowered by this Act to bind by his choice, is conclusively presumed to have chosen the fault [no-fault] alternative.[72]

t. Unless specifically stated, nothing in this Act changes provisions of [this State's] law requiring liability insurance covering property damage arising out of the ownership, maintenance, or use of a motor vehicle.

u. Insurance on Out-of-State Motor Vehicles in In-State Accidents

1. Notwithstanding any contrary provision in it, every contract of liability insurance for injury, wherever issued, covering ownership, maintenance, or use of a motor vehicle, includes coverage such that a person eligible for no-fault benefits under this Act can neither sue nor be sued in accordance with the provisions of this Act, and persons not eligible for no-fault benefits who would otherwise have a claim against a person eligible for no-fault benefits can claim against inverse liability coverage in accordance with the provisions of this Act.
2. An insurer authorized to transact or transacting business in this State may not exclude in any contract of liability insurance for injury, wherever issued, covering ownership, maintenance, or use of a motor vehicle, provisions implementing the terms of clause u1 above.[73]

v. An insurer of the no-fault alternative may require an injured person submitting a claim to submit to any reasonable examination or re-examination, at times and places convenient to the injured person, by medical experts selected by the injured person from a list provided by the State medical society, on such terms and conditions as are approved by the [Commissioner] of Insurance. The cost of any such examination shall be paid by the

---

[72] [W]hat to do with insureds who, despite clear requests from their insurers, simply fail to make a choice[?] To solve this problem, the statute could provide that an insured will be deemed to have selected a specified coverage; the statute should designate this as either traditional or no-fault coverage. Arguably, because no-fault is the newer and more novel coverage, the statute should provide that a motorist will receive traditional coverage if he does not choose. [This is what the above provision does.] On the other hand, if [the legislature believes, as does this author, that] . . . no-fault coverage is the more socially beneficial coverage . . . no-fault should be the applicable coverage, absent a designated choice.

Id.

[73] UMVARA, *supra* note 39, § 9 and commentary thereto, at 289–90.

insurer, and a copy of a written report of any such examination or re-examination shall, upon request, be furnished to the injured person.

w. The [Commissioner] of Insurance may adopt rules to provide effective administration of this Act that are consistent with its purposes and are fair and equitable.

Section 2. This Act becomes effective on January 1, 198X.

NOTE

EXAMINATION OF GOVERNMENTAL DECENTRALIZATION IN NEW YORK CITY AND A NEW MODEL FOR IMPLEMENTATION

JAMES W. LOWE*

Since the middle of the nineteenth century the power of American city and town governments over their own populance has shifted to the state and federal level. Local authority, once exemplified by the New England town meeting, has lost the strength that once so impressed observers like Tocqueville. At the same time, the citizenry has lost interest in its democratic institutions, as reflected by the declining rate of participation in elections.

In this Note, Mr. Lowe argues that Americans need to explore ways to reinvigorate local governments in order to stimulate the citizenry into participating in the democratic process. He discusses the theoretical underpinnings of participatory models of democracy and the lessons to be learned from those models for restructuring local governments. Mr. Lowe traces the history of decentralization attempts in New York City and draws on this history to present his own model of decentralization in New York City as an answer to the need for greater citizen participation.

> A Nation may establish a free government, but without municipal institutions, it cannot have the spirit of liberty.[1]

On November 7, 1989, the people of the City of New York voted by a five-to-four margin to approve a series of changes in the City's governing Charter.[2] The Charter revision was mandated, in part, by the United States Supreme Court's decision in *Board of Estimate v. Morris*,[3] which held that the structure of the Board of Estimate, the City's most powerful governing

---

* Law Clerk to the Hon. Peter K. Leisure, United States District Court for the Southern District of New York. B.A. Yale University, 1985; J.D. Harvard Law School, 1989; New York City Urban Fellow, 1985–86. The author would like to thank Professor Lance Liebman for his encouragement, supervision, and advice; Linda Gibbs of the New York City Charter Revision Commission for her assistance in identifying and providing background material; and Natalie M. Hanlon for her support during the completion of this project.

[1] A. DE TOCQUEVILLE, DEMOCRACY IN AMERICA 61 (P. Bradley trans. 1946).

[2] *New York Adopts Broad Revision of City Charter, Voter Poll Shows*, N.Y. Times, Nov. 8, 1989, at A1, col. 4; *A Overhaul of New York City Charter is Approved, Polls Show*, N.Y. Times, Nov. 8, 1989, at B1, col. 2.

[3] 109 S. Ct. 1433 (1989). The Supreme Court affirmed the decision of the Court of Appeals, Morris v. Board of Estimate, 831 F.2d 384 (2d Cir. 1987), which affirmed the decision of the district court, Morris v. Board of Estimate, 647 F. Supp. 1463 (E.D.N.Y. 1986).

body, was unconstitutional.[4] The Court's ruling forced the City to restructure its government so that it could meet the constitutional requirements of one person, one vote while at the same time addressing the complexities of governing the Nation's most populous municipality.

The decision of the Supreme Court was not unexpected, and the City was already in a position to meet the challenge of restructuring imposed by the ruling. Shortly after the district court ruling in late 1986 finding the Board of Estimate unconstitutional,[5] Mayor Edward I. Koch appointed a commission to study and propose changes in the structure of City government that would meet the objections raised by the court.[6] After the Supreme Court's decision, the New York City Charter Revision Commission ("the Commission"), which had halted its work pending the outcome of the Supreme Court's decision, managed to complete its task within a few months, approving its final recommendations on time for submission to the voters on the November ballot.[7]

The Commission was given a daunting task: to restructure the City's government in a way that would gain approval from the Court, the Justice Department,[8] and, most importantly, the people of New York City who had to vote to approve any change in the Charter.[9] The Commission was faced with enormous po-

---

[4] The Board of Estimate consisted of the Mayor, the City Council President, the Comptroller (all of whom are elected citywide), and the five Borough Presidents who represent jurisdictions ranging from 350,000 to 2.2 million persons. Each Borough President had one vote on the Board of Estimate; the Supreme Court in *Board of Estimate*, 109 S. Ct. 1433, found this allocation of votes to be a violation of the principle of one person, one vote. *See* Reynolds v. Sims, 377 U.S. 533 (1964) (declaring the principle of one person, one vote to be mandated by the Constitution); Avery v. Midland County, 390 U.S. 474 (1968) (applying one person, one vote to local governments). For a concise description of the powers of the Board of Estimate, see *From Budget to Land Use: The Powers of the Board of Estimate*, N.Y. Times, Mar. 30, 1989, at B1, col. 2.

[5] *Morris*, 647 F. Supp. 1463.

[6] NEW YORK CITY CHARTER REVISION COMMISSION, THE VOTER'S HANDBOOK ON CHARTER CHANGE 4 (1988). The Mayor has the authority under state law to appoint a charter revision commission. N.Y. MUN. HOME RULE LAW § 36(4) (McKinney 1969).

[7] NEW YORK CITY CHARTER REVISION COMMISSION, SUMMARY OF FINAL PROPOSALS 1 (August 1989) [hereinafter FINAL PROPOSALS].

[8] The Justice Department must review any proposed restructuring to assure compliance with the Voting Rights Act. *Planners Must Satisfy Voting Law in Shifting of Municipal Power*, N.Y. Times, Mar. 25, 1989, at 31, col. 3. The Commission made its submission to the Justice Department on August 11, 1989. *See* Letter from Frederick A.O. Schwarz, Jr. to Barry H. Weinberg, Esq. 9 (August 11, 1989) (on file with the HARV. J. ON LEGIS.).

[9] N.Y. MUN. HOME RULE LAW § 36(5)(b) (McKinney 1969). On the ballot, the Charter changes were reduced to two questions. *On Ballot, a Charter is Distilled to 276 Words*, N.Y. Times, Sept. 10, 1989, § 1, at 41.

litical pressure from those empowered by the current structure.[10] Party leaders, elected officials, and those owing their careers or other political debts to the current elected officials pressured the Commission to make as few changes as possible.[11]

While the New York City government will certainly look quite different after the new Charter is implemented, the actual functioning of the City on a day-to-day basis will change only modestly. Government in New York City will continue to be controlled almost entirely from City Hall, although that control will be distributed differently within that small building.[12] The City's neighborhoods will be recognized by an increase in the number of City Council seats from thirty-five to fifty-one, thus reducing district size approximately thirty percent.[13]

The new Charter, in the end, will most likely accomplish the narrow purposes established for it by the Commission: the elimination of the Board of Estimate without paralyzing City government. Power is shifted in ways that allow for the greatest continuity of authority with the least possible radical restructuring. Not a single elected office is eliminated; the Borough Presidents, the Council President, and the Comptroller, all or part of whose power lay in their role on the Board of Estimate, have survived in the new Charter. However, their roles have been reduced, and, in the case of the Council President and Borough Presidents, become rather nebulous.[14] A step has been

---

[10] *See, e.g., 3 Borough Leaders Seek Strategy to Save Board*, N.Y. Times, Mar. 28, 1989, at B1, col. 4.

[11] *See* K. BRADBURY, A. DOWNS & K. SMALL, URBAN DECLINE AND THE FUTURE OF AMERICAN CITIES 295 (1982). "Nearly all U.S. metropolitan areas have failed to reform those institutions and practices that perpetuate or aggravate the problems of urban decline for one reason: too many people benefit from the existing arrangements. That is why radical restructuring of urban areas seems so unlikely." *Id.* For a description of the coalition mounted by beneficiaries of the structure to try to defeat the Charter revision proposal, see *Coalition Opposing Charter Revision Starts Its Campaigning*, N.Y. Times, Sept. 28, 1989, at B1, col. 2.

[12] Indeed, one of the major criticisms of the new Charter is that it puts too much control in the hands of the Mayor and his appointees—power that was formerly restrained by the Board of Estimate. *Panel Finishes Plan to Revise New York City's Government*, N.Y. Times, Aug. 3, 1989, at A1, col. 6. *See also* NEW YORK CITY CHARTER REVISION COMMISSION, DISSENTING REPORT BY COMMISSIONER W. BERNARD RICHLAND (August 1989).

[13] FINAL PROPOSALS, *supra* note 7, at 9.

[14] The Council President becomes the vice-president of New York City. He continues to be first in line of succession to the Mayoralty and is now assigned the responsibility of acting as a "Public Advocate" for citizen complaints. He has no major policy role; his only substantive activities are to chair a new commission on public information and to vote in the Council in case of a tie—an unlikely event in a 51-member Council. FINAL PROPOSALS, *supra* note 7, at 19. The Borough Presidents retain substantial appointment power. However, their input into the City's budget power, quite substantial on the

taken toward improving the anemic Community Boards by improving the selection of members and by increasing their input into land use decisions, but the Boards remain devoid of real authority over service delivery and social policy within their districts.[15]

This Note will argue that while the Commission had little choice but to focus on rearranging the existing structure, the opportunity to look more deeply at the entire governmental structure, opened by the need for restructuring, should not be missed. Indeed, a thorough re-evaluation of the structure of local government, not just in New York City, but throughout the country, is needed.[16] While this Note does not pretend to be such an exhaustive review of American municipal governments, it is intended to spur thoughts of such re-evaluations in other cities. This can be accomplished by probing the reasoning behind reorganization of local government through decentralization. In addition, this Note will examine the history of decentralization in New York City and present a model of a tiered, small unit government for that vast city.

American political society has lost its sense of the role of local government in its political structure, and, more clearly, Americans have lost their sense of the potential of local government to energize their democracy. Tocqueville wrote over 150 years ago,

> I believe that provincial institutions are useful to all nations, but nowhere do they appear to me to be more necessary than among a democratic people . . . . *How can a populace unaccustomed to freedom in small concerns learn to use it temperately in great affairs*? What resistance can be offered to tyranny in a country where each individual is weak and where the citizens are not united by any common interest? Those who dread the license of the mob and those who fear absolute power ought alike to desire the gradual development of provincial liberties . . . . The only nations which deny the utility of provincial liberties are those which have fewest of them . . . .[17]

Board of Estimate, is reduced to a mere recommendation for five percent of the expense and capital budgets. *Id.* at 13.

[15] *Id.* at 47.

[16] This Note is not a criticism of the work of the New York City Charter Revision Commission or of the United States Supreme Court ruling in Board of Estimate v. Morris, 109 S. Ct. 1433 (1989).

[17] A. DE TOCQUEVILLE, *supra* note 1, at 95, 97 (emphasis added).

American government and democracy can be tested against Tocqueville's ideal; some indicators may give cause for concern. Only about fifty percent of eligible voters went to the polls in November 1988.[18] Only once, in 1924, has turnout for a presidential election been lower.[19] Voter participation has declined in almost every election since 1960.[20] Moreover, the United States has the lowest turnout rate of any of the world's democracies.[21] Meanwhile, more and more authority formerly exercised at the local level has shifted to the state and federal governments. This shift of power is not a new phenomenon. Since the middle of the nineteenth century, power has shifted away from localities.[22] More recently, the increasing dependence of cities on state and federal aid,[23] and the removal of fiscal authority from some cities in the wake of financial crises in the 1970's[24] have further diminished municipal authority.

---

[18] *Ask the Globe*, The Boston Globe, Jan. 2, 1989, at 24, col. 1.

[19] *Id.*

[20] *Voting: . . . for the many, by the few*, Christian Sci. Monitor, Nov. 9, 1988, at 1. There are many possible explanations for the decline in voter participation. While a sense of the meaninglessness of an individaul vote is a compelling explanation, there is also a powerful argument that the voter registration system is to blame for low turnout, though not for declining turnout. A vigorous recent attack on registration requirements called them potentially as restrictive as the poll tax. *See* Note, *Voter Registration: A Restriction on the Fundamental Right to Vote*, 96 YALE L.J. 1615 (1987). *See also* Carlson, *Personal Registration Systems Discourage Voter Participation*, 60 NAT'L CIVIC REV. 597 (1971). Currently 35 to 40% of eligible Americans are not registered to vote. Note, *supra*, at 1615. For proposals to reduce the burden of registration, see Note, *Providing Access to Voter Registration: A Model State Statute*, 24 HARV. J. ON LEGIS. 479 (1987); Comment, *A Model Voter Registration System*, 73 NAT'L CIVIC REV. 104 (1984).

[21] CITY OF NEW YORK, NEW YORK CITY VOTER ASSISTANCE PROGRAM 1988 ANNUAL REPORT 1 (1988).

[22] For a discussion of the legal aspects of city power, see Frug, *City as a Legal Concept*, 93 HARV. L. REV. 1057, 1095–1120 (1980). *See also* Williams, *City Status in American Law*, 1986 WIS. L. REV. 83 (1986) (arguing, in part, that municipalities are vulnerable to the whims of courts and commentators because cities are not explicitly enumerated in our constitutional framework). For a discussion of the 19th century power shift in New York City, see H. HARTOG, PUBLIC PROPERTY AND PRIVATE POWER (1983); Frug, *Property and Power: Hartog on the Legal History of New York City*, 1984 AM. B. FOUND. RES. J. 673 (1984).

[23] In 1974, federal aid accounted for 10.6% of the general expenditures of cities over 300,000 in population. D. CARALEY, CITY GOVERNMENTS AND URBAN PROBLEMS 135 (1977). In 1979, federal aid to New York City was $2.837 billion, 21.8% of total City revenue. Vitullo-Martin & Nathan, *Intergovernmental Aid*, in C. BRECHER & R. HORTON, SETTING MUNICIPAL PRIORITIES 1981, at 52 (1980). In 1984, state and federal aid paid for $6.3 billion of New York City's expenditures. M. SHEFTER, POLITICAL CRISIS, FISCAL CRISIS 137 (1987).

[24] The most famous of these crises was the near default of New York City in the period 1975 to 1978. Many of the legal steps taken by the state to stem the crisis involved removing authority from the City over its own finances. For a contemporary description of these measures, see Comment, *New York—A City in Crisis: Fiscal Emergency Legislation and the Constitutional Attacks*, 6 FORD. URB. L.J. 65 (1977). The best descrip-

Despite these obstacles, city government is the government closest to the people, and it was the strength of our local governments that so impressed Tocqueville 150 years ago. In an effort to energize the citizenry, the United States needs to explore ways to reinvigorate local governments. Thus, this Note is an argument and a proposal with one foot firmly planted in the republican tradition. Increasing the role of individual citizens in the governing process is the underlying theme.[25]

This Note is divided into three sections. The first explores further the need to look at the arguments for participatory models of democracy and recognize the lessons to be learned from those models for future attempts to restructure local governments. Arguing that an empowered electorate can be an energetic and interested electorate, this section proposes that the local level is the appropriate place to empower the electorate by granting to neighborhoods power over local service delivery and land use decisions, thereby giving them greater influence over the central city government.[26] The second section, a discussion of the history of the New York City government, focuses on prior efforts to decentralize government and increase citizen authority. The final section presents a blueprint of a decentralized structure for New York City. The model presented is not intended as a republican or communitarian ideal, but draws on past studies, experiences, and criticism. Focusing on the realities of managing a major city in the 1990's, the model is intended as a pragmatic approach to a perceived need for greater citizen involvement in the operation of government. It is hoped that the model will serve as an incentive for others both to re-examine their local government through new lenses and to look for the possibility of increasing citizen control in their communities.

---

tions of the fiscal crisis as a whole are K. AULETTA, THE STREETS WERE PAVED WITH GOLD (1979); C. MORRIS, THE COST OF GOOD INTENTIONS (1980); M. SHEFTER, *supra* note 23. Shefter provides one of the best descriptions of the political effects of the fiscal crisis and their impact on City authority and the City electorate. *Id.* at 149–216.

[25] The scope of the decentralization proposed in this Note is limited. *See* Section III, *infra*.

[26] This Note will not examine other means of increasing citizen involvement in government such as the initiative and referendum. For a discussion of these governmental techniques, see T. CRONIN, DIRECT DEMOCRACY, THE POLITICS OF INITIATIVE, REFERENDUM AND RECALL (1989); Gillette, *Plebiscites, Participation, and Collective Action in Local Government Law*, 86 MICH. L. REV. 930 (1988).

I. THE NEED FOR AND STRENGTH OF PARTICIPATION

As the nation grows larger, government pulls further away from the governed as individual legislators represent more and more people.[27] Local participation was once possible through political clubs and organized wards which functioned on the neighborhood level. However, the ward system became dominated by corrupt and racist politicians and thus became synonymous with bossism and corruption growing out of the excesses of Tammany Hall in New York[28] and the Democratic Machine in Chicago,[29] to mention two of many. The ward system with its cadre of precinct captains had a good side: it got people to participate who otherwise might not have done so, and it gave many people, particularly the very poor, a political voice.[30] And these people voted in recognition of those who had provided services to them, their neighbor, or their block.[31]

The point is not to argue for a return to the ward system, for the corruption, racism, and favoritism endemic in the political machines outweighed any advantages.[32] The ward system did, however, represent government at its most mundane, and thus, Tocqueville argues, at its most important and energizing level.[33] Local citizens had a local unelected representative who, more often than not, would be responsive to their desires and needs. Yet local systems need to be approached with care. The parochialism of the ward system and the corruption that grew up

---

[27] Congressional districts now include over 500,000 people each. In 1830, Tocqueville's time, congressional districts included only about 48,000 people each.

[28] For a description of the political machine in New York, see R. PEEL, THE POLITICAL CLUBS OF NEW YORK CITY (1935). The most entertaining, if self-serving, account of the workings of Tammany Hall comes from a member of the organization itself, George Washington Plunkitt, who is best known for describing the system of patronage and petty bribery rampant at the time as "honest graft." W. RIORDON, PLUNKITT OF TAMMANY HALL (1963).

[29] *See* Meyerson & Banfield, *A Machine at Work*, in URBAN GOVERNMENT 135–44 (E. Banfield ed. 1961).

[30] *See* Fitch, *The People*, 30 PUB. ADMIN. REV. 481, 484 (1970).

[31] A. MACDONALD, AMERICAN CITY GOVERNMENT AND ADMINISTRATION 293–95 (6th ed. 1956); T. REED, MUNICIPAL GOVERNMENT IN THE UNITED STATES 111 (1934). *See also* C. KNEIER, CITY GOVERNMENT IN THE UNITED STATES 416, 417 (3rd ed. 1957) (arguing that the precinct system links the poorer classes to city government).

[32] T. REED, *supra* note 31, at 112–14; *see also* A. MACDONALD, *supra* note 31, at 301–02.

[33] [T]he township, at the center of the ordinary relations of life, serves as a field for the desire of public esteem, the want of exciting interest, and the taste for authority and popularity; and the passions that commonly embroil society change their character when they find a vent so near the domestic hearth and the family circle.

A. DE TOCQUEVILLE, *supra* note 1, at 67.

around it rightly soured those interested in good municipal administration and neighborhood control.[34] Add to this experience the intolerable history of racism and exclusion in many small and some large local communities,[35] and the combination leads to an understandable distrust of local government units.[36] This distrust is fed by contemporary revelations of deeply ingrained corruption in existing local government structures.[37] These legitimate concerns about parochialism and corruption based on a distressing and, in many ways, unacceptable urban history should not lead to the conclusion that local governments are simply a necessary evil to be controlled.[38] To reach such a conclusion would be to downgrade the importance of our most

---

[34] The corruption of the city machines led one observer to write in 1888, "There is no denying that the government of cities is the one conspicuous failure of the United States." T. REED, *supra* note 31, at 115.

[35] One of the better-known examples of the racist application of local authority is the decision of the town of Shaw, Mississippi to pave only those streets passing through White neighborhoods and to leave those in the predominantly Black neighborhoods unpaved and lined with open sewers. The United States Court of Appeals for the Fifth Circuit, after extensive litigation, found an equal protection violation and ordered the town to provide services regardless of race. Hawkins v. Town of Shaw, 437 F.2d 1286 (5th Cir. 1971), *aff'd on rehearing en banc*, 461 F.2d 1171 (5th Cir. 1972). For more on this case, see C. HAAR & D. FESSLER, THE WRONG SIDE OF THE TRACKS 11–55 (1986).

[36] There may be other, less acceptable, reasons for the distrust of local governments by businessmen, leaders of state and national government, and academics. There are personal links among the leaders of business, higher governmental bodies, and academia in this country, captured most convincingly in C. MILLS, THE POWER ELITE (1956), and reinforced in G. DOMHOFF, WHO RULES AMERICA? (1967). However, city government leaders, indeed the leaders of most localities, do not fit within Mills's and Domhoff's matrix of the American ruling class. G. DOMHOFF, *supra*, at 132. In general, localities are run by people who fall outside any definition of a cohesive American power structure, though there have been exceptions such as the Lindsay Administration in New York. The suspicion that those in the ruling elite—judges, members of Congress, businessmen, constitutional lawyers—feel toward those who have risen from local political organizations or small local businesses to run localities may well account, in part, for the disempowerment of localities through federal control, *see, e.g.*, Garcia v. San Antonio Metro. Transit Auth., 469 U.S. 528 (1985), *reh'g denied*, 471 U.S. 1049 (1985); South Dakota v. Dole, 483 U.S. 203 (1987); G. FRUG, LOCAL GOVERNMENT LAW 252–97 (1988), narrow interpretations of home rule powers, *see, e.g.*, City of LaGrande v. Public Employees Retirement Bd., 281 Or. 137, 576 P.2d 1204 (1978), and limitations on the ability of localities to raise sufficient operating funds without state approval, *see, e.g.*, Amador Valley Joint Union High School Dist. v. State Bd. of Equalization, 22 Cal. 3rd 208, 583 P.2d 1281, 149 Cal. Rptr. 239 (1978). For a discussion of how the ruling elite works to control municipalities in which the government positions are held by those outside the power elite structure, see G. DOMHOFF, WHO REALLY RULES? (1978), especially chapter five. For an historical approach to this issue, see M. WHITE & L. WHITE, THE INTELLECTUAL VERSUS THE CITY (1962).

[37] *See, e.g.*, J. NEWFIELD & W. BARRETT, CITY FOR SALE: ED KOCH AND THE BETRAYAL OF NEW YORK (1988).

[38] For a description of some of the attacks on local authority, see Williams, *supra* note 22.

democratic governmental institution.[39] As Professor Gerald Frug has stated,

> Many people, both on the right and the left of the American political spectrum, argue that decentralization of power is an essential—and increasingly threatened—ingredient of political freedom. Genuine democratic self-government, they claim, is possible only on the local level. Only local government is close enough to its constituents to permit popular participation in the decisionmaking that affects people's lives; only a local government can tailor its policies to the needs and desires of particular community.[40]

Frug's is a careful condensation and simplification of an extremely complex and sometimes contradictory set of rationales for increased decentralization of local government. The arguments for decentralization, as Frug points out, come from a broad range of positions in the political spectrum including the far left,[41] communitarians,[42] and libertarians.[43] Decentralization literature flourished in the late 1960's[44] to the point where serious proposals were presented for widespread decentralization of major localities.[45] Some localities attempted decentralization and citizen control in isolated programs, with mixed results.[46]

The disappointment with the ineffectiveness of a number of the efforts to increase community participation, particularly those mandated by federal urban programs such as Model Cities,

---

[39] John Stuart Mill wrote on this issue:

> Except by the part they may take as jurymen in the administration of justice, the mass of the population have very little opportunity of sharing personally in the conduct of the general affairs of the community . . . . But in the case of local bodies, besides the function of electing, many citizens in turn have the chance of being elected, and many, either by selection or by rotation, fill one or other of the numerous local executive offices. In these positions they have to act, for public interests, as well as to think and to speak, and the thinking cannot all be done by proxy.

J.S. MILL, *Considerations on Representative Government*, in THREE ESSAYS 365 (1975).

[40] G. FRUG, *supra* note 36, at xv.

[41] *See, e.g.*, D. KRAMER, PARTICIPATORY DEMOCRACY (1972).

[42] *See, e.g.*, Frug, *The City as a Legal Concept*, *supra* note 22.

[43] *See, e.g.*, R. NOZICK, ANARCHY, STATE AND UTOPIA 320–31 (1974).

[44] K. GRAHAM, THE BATTLE OF DEMOCRACY 149 (1986).

[45] For a description of proposals for New York City, see *infra* notes 220–252 and accompanying text.

[46] *See* H. HALLMAN, NEIGHBORHOOD CONTROL OF PUBLIC PROGRAMS 185–86, 202–04 (1970). One of the areas where local participation was encouraged was in the federal urban renewal programs of the 1960's. For a contemporary critique of the effectiveness of the participation elements in those programs, see J. BELLUSH & M. HAUSKNECHT, URBAN RENEWAL: PEOPLE, POLITICS AND PLANNING 274–311 (1967); Strange, *Citizen Participation in Community Action and Model Cities Programs*, 32 PUB. ADMIN. REV. 655 (1972). For a discussion of the success of the New York City public school decentralization, see *infra* notes 275–280 and accompanying text.

diminished the power of those calling for decentralization.[47] The call for decentralization was further weakened by the fiscal crises suffered by many cities during the mid-1970's[48] which moved the policy focus of many local governments away from increasing citizen participation and toward increasing efficiency.[49] There was not necessarily a sacrifice of participation for efficiency, but simply a shift in priorities.

Decentralization of this nation's myriad of local governments[50] should be re-examined. This call for reconsideration of government organization comes from both a sense that the current governmental structures have failed[51] and a belief that the general health of the nation and the success and vitality of our localities, particularly larger cities, can be improved by moving the center of governmental power closer to the gov-

---

[47] The attacks on federal urban renewal policies were powerful and inevitably damaged the reputation of all parts of the programs, including those aimed at participation. The general critique has been that, even with participatory elements, the urban renewal programs failed the neighborhoods. Three years after the end of the Model Cities program, Senator William Proxmire (D-Wis.) stated, "You probably would have better neighborhoods today if there had been no federal programs at all." *Neighborhood Preservation: Hearings on the Cause of Neighborhood Decline and the Impact, Positive or Negative, of Existing Programs, Policies and Laws on Existing Neighborhoods Before the Senate Comm. on Banking, Housing and Urban Affairs*, 94th Cong., 2d Sess. 59 (1976), *quoted in* McClaughry, *Recycling Declining Neighborhoods: Give the People A Chance*, 10 URB. LAW. 318 (1978). *See also* C. GLAAB & A. BROWN, A HISTORY OF URBAN AMERICA 289 (1976). A more radical critique of participation requirements in federal programs argues that these programs were implemented to co-opt the demands of protest movements and radical community leaders. *See* Fainstein & Fainstein, *Economic Change, National Policy and the System of Cities*, in S. FAINSTEIN, N. FAINSTEIN, R. HILL, D. JUDD & M. SMITH, RESTRUCTURING THE CITY 21 (1983).

[48] The decline in the focus of political scientists and urban planners on decentralization can be seen in the indexes to the PUBLIC ADMINISTRATION REVIEW. For 1970, the REVIEW had 29 entries in its index under the heading "Decentralization," five of which referred to book reviews. 30 PUB. ADMIN. REV. 677 (1970). In 1985, the yearly index contained no mention at all of "Decentralization" and had only two items listed under the heading "Citizen Participation and Groups," one of which was a book review. 45 PUB. ADMIN. REV. 889 (1985).

[49] An example of this change in emphasis was the official slogan of Ed Koch's successful 1977 Mayoral campaign attacking his predecessors, including Lindsay, who had emphasized participation: "After eight years of charisma [Lindsay], and four years on the clubhouse [Beame], let's try competence." E. KOCH & W. RAUCH, MAYOR 30 (1984).

[50] In 1972 there were 78,218 local governments, including public authorities and special districts, in the United States. UNITED STATES ADVISORY COMMISSION ON INTERGOVERNMENTAL RELATIONS, IMPROVING URBAN AMERICA: A CHALLENGE TO FEDERALISM 145 (1976).

[51] In their failure to govern democratically, our current structures have failed to act as a government "of the people, by the people." "[T]he drift has been away from institutions that put a premium on public debate toward those that operate in a managerial mode . . . ." S. ELKIN, CITY AND REGIME IN THE AMERICAN REPUBLIC 143 (1987).

erned.[52] Decentralization, then, is not an end in itself; it is a means of increasing the participation of the governed in their governance and thus invigorating the government.[53] This vision is not one of a nation governed by Tocquevillian New England town meetings.[54] Instead, it is a vision of increased citizen involvement in government through increased opportunities to share in those governmental decisions which directly affect their communities.

The scope of the decentralization discussed in this Note is limited. First, central city governments have very limited powers to devolve to neighborhood units.[55] Second, as the history of the ward system illustrates,[56] there is a need for checks and limitations on truly local units to avoid excess parochialism and possible discrimination. Thus, the decentralization considered here consists of substantially increased neighborhood control over delivery of basic services and increased ability to influence land use decisions within the neighborhood. Decentralized units, like other governmental units, would be subject to all constitutional requirements. Completely eliminating supervision of decentralized units by central governments and the courts would open the door both to discrimination and tyranny within the unit and to inequality and discrimination between units.[57]

The belief in the importance of citizens' involvement in their government can be viewed as an example of "republican" thought,[58] drawing as it does from the writings of Toc-

---

[52] This belief has often been expressed before. *See, e.g.*, D. YATES, NEIGHBORHOOD DEMOCRACY 25 (1973).

[53] There is no difficulty in showing that the ideally best form of government is that in which the sovereignty . . . is vested in the entire aggregate of the community; every citizen not only having a voice in the exercise of that ultimate sovereignty, but being, at least occasionally, called on to take an actual part in the government . . . .

J.S. MILL, *supra* note 39, at 186.

[54] There may be places, however, where such gatherings remain an appropriate form of governance, particularly in small towns or rural communities where the number of citizens is sufficiently small to make such gatherings manageable.

[55] *See infra* notes 75–84 and accompanying text.

[56] *See supra* notes 28–34 and accompanying text.

[57] Ravitch, *The Rhetoric of Decentralization*, N.Y. AFFAIRS, Summer 1974, at 103, 110; Dixon, *Rebuilding The Urban Political System: Some Heresies Concerning Citizen Participation, Community Action, Metros, and One Man-One Vote*, 58 GEO. L.J. 955, 966 (1970).

[58] "Civic republicanism" has seen an intellectual renaissance in legal circles in recent years, thanks in part to the writings of Frank Michelman and Cass Sunstein. Michelman, *The Supreme Court, 1985 Term—Foreword: Traces of Self-Government*, 100 HARV. L. REV. 4, 17–55 (1986); Michelman, *Law's Republic*, 97 YALE L.J. 1493 (1988); Sunstein, *Beyond the Republican Revival*, 97 YALE L.J. 1539 (1988). *See also* Rose, The Ancient Constitution vs. the Federalist Empire: Antifederalism From the Attack on "Monarch-

queville[59] and Mill,[60] among others. But the sources are broader. The belief in the potential of increased participation which drives this Note also grew out of the writings of the proponents of a pluralist theory of democracy,[61] the communitarian proposals and hopes for society,[62] and finally, by an unrooted determination that citizens should be given the broadest possible avenues to participate in their own governing.[63]

Despite the glorified image of the United States presented by Tocqueville, the widespread participation that he documented in New England is not generally part of the American tradition. The "We the People" who declared the American Constitution their blueprint for government were, in reality, only a portion of the population to be governed.[64] As late as the 1960's even the right to vote, the most visible if least participatory element of a popular government, was restricted by law through the poll tax,[65] residency requirements,[66] and even requirements of property ownership.[67]

The rhetoric of participation, however, permeates American political traditions. Thomas Jefferson wrote,

> In government, as well as in every other business of life, it is by division and subdivision of duties alone, that all matters, great and small, can be managed to perfection . . . . And the whole is cemented by giving to every citizen, personally, a part in the administration of public affairs.[68]

---

ism" to Modern Localism (August 1988) (unpublished manuscript). *See generally Symposium: The Republican Civic Tradition*, 97 YALE L.J. 1493 (1988).

[59] A. DE TOCQUEVILLE, *supra* note 1.

[60] J.S. MILL, *supra* note 39.

[61] For this Note, the most important of the theorists who see American politics and democracy as a contest between interest groups is Robert Dahl. *See* R. DAHL, WHO GOVERNS? (1961); R. DAHL, DILEMMAS OF PLURALIST DEMOCRACY (1982).

[62] *See, e.g.*, Pitkin & Schumer, *On Participation*, 2 DEMOCRACY 43 (1982); R. KANTER, COMMITMENT AND COMMUNITY (1972).

[63] This sense of the innate importance of providing an opportunity to participate has been taken to its logical end by the United Nations, which has declared the right to participate in one's own government to be a human right. *Universal Declaration of Human Rights of 1948*, G.A. Res. 217A (III), U.N. Doc. A/810 (1948) and *International Covenant on Civil and Political Rights*, G.A. Res. 2200, 21 U.N. GAOR Supp. (No. 16) 52, U.N. Doc. A/6316 (1966), *quoted in* Steiner, *Political Participation as a Human Right*, 1 HARV. HUM. RTS. Y.B. 77, 86–87 (1988).

[64] Bell & Bansal, *The Republican Revival and Racial Politics*, 97 YALE L.J. 1609, 1610 (1988).

[65] *See* Harper v. Virginia State Bd. of Elections, 383 U.S. 663 (1966).

[66] *See* Dunn v. Blumstein, 405 U.S. 330 (1972).

[67] *See* Kramer v. Union Free School Dist. No. 15, 395 U.S. 621 (1969).

[68] 11 T. JEFFERSON, THE WORKS OF THOMAS JEFFERSON 347 (P. Ford ed. 1905), *quoted in* Williams, *supra* note 22, at 106.

Later Judge Thomas Cooley of the Michigan Supreme Court, an influential constitutional scholar of the late nineteenth century, took Jeffersonian rhetoric and republican thought and applied them to the problem of government organization and power distribution.[69] Cooley argued that the power of government came from the people, and that only a portion of that power had been delegated to the state through the Constitution. He believed that local self-government, free from the controls of state authorities, was an "absolute right."[70]

While Cooley's theory did not become law or reality,[71] it remains a powerful vision.[72] Two parts of that vision are of particular importance to contemporary municipal government. First, Cooley believed that local governments should be masters of their own realm, that localities should be able to pass their own laws and control their own destinies absent interference from states and the federal government.[73] Second, he believed that this notion of "non-interference" was required by the right of all people to govern themselves.[74]

Dillon's Rule attacked Cooley's notion of autonomous self-government most directly. Local governments are, simply, creatures of their states.[75] At the turn of the century, urban reformers began the "Home Rule Movement" which resulted in the inclusion of amendments in the constitutions of many states giving municipalities a certain measure of autonomy.[76] In fact, the area

---

[69] T. COOLEY, A TREATISE ON THE CONSTITUTIONAL LIMITATIONS WHICH REST UPON THE LEGISLATIVE POWER OF THE STATES OF THE AMERICAN UNION (1868).

[70] People *ex rel.* Le Roy v. Hurlbut, 24 Mich. 44, 108 (1871) (Cooley, J., concurring). For a further discussion of Cooley's views, see Frug, *The City as a Legal Concept*, *supra* note 22, at 1113; Williams, *supra* note 22, at 88.

[71] What did become law was "Dillon's Rule": local governments contain only as much authority as the state will give them. In essence, all power flows from above and is granted at the will of the sovereign authorities, in this case, the state and the federal governments. Dillon's Rule is named after John Dillon, whose 1872 treatise on local government law set out the parameters of local authority which, in general, remain law today. J. DILLON, TREATISE ON THE LAW OF MUNICIPAL CORPORATIONS (1872). For a discussion of Dillon's Rule, see Gere, *Dillon's Rule and the Cooley Doctrine: Reflections of the Political Culture*, 8 J. URB. HIST. 271 (1982); G. FRUG, *supra* note 36, at 56–59; Frug, *The City as a Legal Concept*, *supra* note 22, at 1109–13; F. MICHELMAN & T. SANDALOW, MATERIALS ON GOVERNMENT IN URBAN AREAS 252–56 (1970); Williams, *supra* note 22, at 84, 90–100.

[72] It was a powerful vision at the time it was written as well, selling more copies than any other book on American law in the period. Williams, *supra* note 22, at 145.

[73] People *ex rel.* Le Roy v. Hurlbut, 24 Mich. at 97–99, 108.

[74] *Id.* at 98, 107–08.

[75] *See* Hunter v. City of Pittsburgh, 207 U.S. 161 (1907).

[76] Some form of municipal home rule exists in 35 states, county home rule in 17 states. J. FORDHAM, LOCAL GOVERNMENT LAW 73 (1986). For examples of home rule provisions, see G. FRUG, *supra* note 36, at 88–91. A thorough examination of home rule

of autonomy granted by these provisions was quite narrow and has been further restricted by courts adopting an increasingly restricted definition of local matters.[77] Further, a local government can be preempted by state law covering the same issue as that regulated or controlled by local initiative.[78] City actions can also be preempted by federal law.[79]

Not all restrictions on local governments are objectionable. The application of anti-discrimination laws to local actions is essential to avoid abuses of parochialism and racism. Some state and federal actions designed to avoid discrimination among local governments are also necessary to avoid the creation of class divisions between local areas.[80] But even when seemingly protecting individual rights, courts and state and federal officials can take actions which, when balanced against local governmental interests, are more destructive than protective.[81]

State and federal controls beyond those essential to eliminating base discrimination against racial and social minorities weaken local government in at least two related ways. First, and most obvious, outside control of local governmental policy, particularly as it relates to land use and service delivery issues, means that an individual locality does not control much of the governmental authority nominally within its jurisdiction.[82] Not only may states and the federal government control much of the governmental activity in localities, the localities also fear that

---

in theory and in practice, written at the time the Home Rule Movement was still powerful, is J. McGOLDRICK, LAW AND PRACTICE OF MUNICIPAL HOME RULE, 1916–1930 (1933).

[77] *See* Libonati, *Restructuring Local Government*, 19 URB. LAW. 645, 646 (1987); Frug, *The City as a Legal Concept*, *supra* note 22, at 1117.

[78] G. FRUG, *supra* note 36, at 203–21.

[79] *Id.* at 297–306.

[80] An example of a state action to avoid discrimination among localities is the line of cases requiring Mt. Laurel and similar wealthy suburbs in New Jersey to bear their fair share of the obligation to house middle and low income individuals. *See, e.g.*, Southern Burlington County N.A.A.C.P. v. Township of Mt. Laurel, 67 N.J. 151, 336 A.2d 713 (1975), *appeal dismissed*, 423 U.S. 808 (1975); Southern Burlington County N.A.A.C.P. v. Township of Mt. Laurel, 92 N.J. 158, 456 A.2d 390 (1983); Hills Dev. Co. v. Township of Bernards, 103 N.J. 1, 510 A.2d 621 (1986).

[81] The United States Supreme Court has, in the past, indicated some willingness to give some weight to local governmental interests in certain limited cases. *See* White v. Massachusetts Council of Constr. Employees, 460 U.S. 204 (1983) (holding that a residency requirement for workers on city contracts does not violate the Commerce Clause). However, the reach of this balancing is apparently quite limited. *See* United Bldg. & Constr. Trades Council of Camden and Vicinity v. Mayor of Camden, 465 U.S. 208 (1984) (holding that a residency requirement very similar to that in *White* may violate the Privileges and Immunities Clause).

[82] For a concise statement of the current legal status of local governments, see Frug, *The City as a Legal Concept*, *supra* note 22, at 1062–67.

any action they might take, though non-discriminatory and widely acceptable to those within the locality, will be struck down as outside the scope of local authority.[83] The result is that American local governments are restricted in their ability to shape their own budgets, formulate their own policies, and organize their own polity.[84]

Outside control of even service and land use decisions within areas controlled by local governments can lead to the demoralization and delegitimation of local government. Even the operation of basic local services is often directed from afar, be it from a state capital or from Washington, D.C. Since these distant policymakers often do not understand or know the political, physical, and social reality[85] of the area on which their decisions will be imposed, the resulting policies on local service and land use issues often poorly fit the needs of the individual locality or neighborhood.[86]

The delegitimation of local governments leads to a second problem with outside control of local affairs: the disenfranchisement and alienation of localities, their officials, and their citi-

---

[83] Macchiarola, *Local Government Home Rule and The Judiciary*, 48 J. URB. L. 335, 336 (1971). An example of a state court reducing local policymaking authority is Anderson v. City of Boston, 376 Mass. 178, 380 N.E.2d 628 (1978), *appeal dismissed*, 439 U.S. 1060 (1979) (holding that the City of Boston could not disseminate material supporting a proposed state constitutional amendment before the voters on referendum). It should be noted that while cities may not use taxpayers' money to indicate support for specific legislation they believe could help localities, no restrictions may be placed on corporations wishing to influence legislators or voters on proposals which would affect those corporations. First Nat'l Bank of Boston v. Belotti, 435 U.S. 765 (1978).

[84] While the decision in *White*, 460 U.S. 204, may have indicated further willingness on the Supreme Court's part to favor local interests, Board of Estimate v. Morris, 109 S. Ct. 1433 (1989), implies that the area of local discretion, particularly in the area of governmental organization, is extremely limited. In *Board of Estimate* the Court refused to consider the appropriateness of a locality developing unique government structures to meet local needs. While it is unlikely that the New York City Board of Estimate could survive any strict scrutiny of its structure's relation to any perceived strong local need—in fact it failed such an examination in the district court, Morris v. Board of Estimate, 647 F. Supp. 1463 (E.D.N.Y. 1986)—the complete refusal of the Court even to entertain the possibility that uniqueness or extraordinary local circumstances could justify any deviation from constitutional norms, 109 S. Ct. at 1442 n.10, suggests a rejection of the notion of independent local government legitimacy.

[85] For an excellent illustration of the harm that well-meaning policymakers can potentially cause when they do not understand the physical and social realities of the area their work will affect, see J. JACOBS, THE DEATH AND LIFE OF GREAT AMERICAN CITIES 8–13 (1961).

[86] The problem is neatly summed up in a poem from an anonymous social scientist:

> The reason why cities are ugly and sad,
> Is not that the people who live in them are bad
> It's just that the people who really decide
> What goes on in the city live somewhere outside.

R. FRIEDLAND, POWER AND CRISIS IN THE CITY xvii (1982).

zens.[87] Because there is little political authority to distribute, the powerlessness of local governments makes meaningful decentralization of local *governing* functions difficult. Local governments provide services, often in a way mandated by the state or federal governments, and local governments determine the use of land in their jurisdiction. Local governments, then, manage; they do not govern. Citizens who try to interact with their centralized local governments are often frustrated by this lack of authority, and good public servants are often driven away from local government because of the inability of that level of government to affect any substantive policy.

The elimination of Dillon's Rule would involve a dramatic shift in the American understanding of intergovernmental relations.[88] Such a significant reorganization is not essential to improving local governments. Some reduction in the subjugation of local governments under Dillon's Rule, particularly in the areas of service delivery and land use control, would increase the governing power and authority of local governments. The difficulty of accomplishing even this change in institutional arrangements does not mean that decentralization is a fruitless task. If anything, the difficulty of the task acts as an impetus to move more determinedly. By giving some authority to small groups and allowing those groups an opportunity to exercise that power, it may be possible to show that local authority can, in fact, be a more effective means of governance. If small, participatory units are successful in providing civic governance, such success may provide the empirical underpinning for a challenge to the regime of Dillon's Rule.[89]

---

[87] COMMISSION ON THE CITIES IN THE '70's, THE STATE OF THE CITIES 10–11 (1972).

[88] Frug argues that federalism, as now understood, cannot support a notion of non-interdependent, autonomous localities. Frug, *Empowering Cities in a Federal System*, 19 URB. LAW. 553, 567–68 (1987). Such a radical reconception of American political philosophy as envisioned by Frug may not be necessary to support a theory of autonomous or semi-autonomous localities. Certainly there is room in the liberal tradition for local autonomy. *See* Libonati, *supra* note 77, at 649. Great Britain has had, in the past, significantly more local autonomy and decentralization than the United States. H. HANHAM, THE NINETEENTH CENTURY CONSTITUTION 373 (1969); Hill, *Local Government in Present-Day England*, 4 URB. LAW. 463 (1972).

[89] There is little empirical evidence on the success or failure of attempts to decentralize, other than a series of anecdotal studies, perhaps because it is difficult to develop objective criteria to judge the effectiveness of government reorganization. Boyle, *Reorganization Reconsidered: An Empirical Approach to the Decentralization Problem*, 39 PUB. ADMIN. REV. 458 (1979). *But see* Fitch, *supra* note 30, at 485 (advocates of decentralization have no evidence that the system would work, and the example of the decentralization of London is not encouraging).

Decentralization, then, can be a powerful opportunity for exploring intergovernmental arrangements. More importantly, decentralization provides an opportunity for reversing the alienation and political disempowerment of urban dwellers by providing governmental units of sufficiently small size so as to allow individual citizens to have a voice regardless of their race, sex, or economic status.[90] This does not mean creating a government of the whole, structured around a series of New England town meetings.[91] The idea is to create local governmental institutions of a size and scale that will allow citizens a real opportunity to participate in their own governing—to govern themselves.[92] These local units[93] should not be totally autonomous.[94] The constitutional notion of checks and balances is as essential at the local level as at the national to avoid the Madisonian nightmare of factional discrimination.[95]

Too much control from above in areas of local authority, however, will result in the symbolic disempowerment of these local units and the corresponding loss of interest in them by

---

[90] This is particularly true in very large urban areas. New York is the most extreme example, but there are an increasing number of American cities whose size has reached the point at which the central city government cannot be considered, in terms of citizen participation, a local government.

[91] *But see* Frug, *supra* note 88, at 559, 563, 565. Frug, among others, views this communal governing as the goal of participation theory.

[92] Kramer would divide the city into units of 10,000. D. KRAMER, *supra* note 41, at 143. Dahl believes that the appropriate population per unit of government for successful democracy in a contemporary American city is between 50,000 and 200,000. Dahl, *The City in the Future of Democracy*, 61 AM. POL. SCI. REV. 953 (1967). Such a size allows at once a sense of meaningful community, Dahl says, while allowing sufficient diversity to avoid some of the parochialism inherent in many smaller communities. *Id.* at 967. Such parochialism has often been the focus of the critics of decentralization. Fesler, *Approaches to the Understanding of Decentralization*, 27 J. POL. 536, 542–45 (1965). *See also supra* notes 30–34 and accompanying text.

[93] The term "local unit" is used here rather than the traditional terms "neighborhood" or "community." While both "neighborhood" and "community" are warm, comforting terms, defining a neighborhood or a community is often difficult. Claiming to design local structures on historic, undefined neighborhood or community boundaries may create border disputes when lines are actually drawn. Also, it may be undesirable in many cases to develop units based on historic boundaries since some historic neighborhoods and communities are traditionally so xenophobic that to give them autonomy would reinforce their isolationism. Similarly, defining a local unit within an existing troubled neighborhood may result in an immediate demoralization of that unit as it views itself in comparison to other local units. For a discussion of the advantages and pitfalls of place as a basis of collective action, see Logan, *Growth, Politics, and the Stratification of Places*, in R. LAKE, READINGS IN URBAN ANALYSIS 73–83 (1983).

[94] In fact, their direct powers should be limited to service delivery and land use. *See infra* notes 308–319 and accompanying text.

[95] THE FEDERALIST No. 10 (J. Madison). *See also supra* note 35 and accompanying text.

their communities.[96] Instead, these local units should, through their political effectiveness, be able to significantly influence the upper, geographically larger, levels of government.[97] Those who manage and govern the larger community—the city and the state—should be more responsive and closely connected to the individual local units and their citizens.[98] This involves decentralization of two types, administrative and political.[99] First, as has been discussed above, political authority would devolve toward the local unit by creation of governmental structures within those units and of direct connections between the local unit governments and the central city government, perhaps through overlapping officials elected at the local unit level. The second element involves changing the service management structure of the larger governmental unit, the city, so as to make the delivery of services more responsive[100] to the needs of the individual local units.[101] This is the central feature of local unit government.

It is essential that both administrative and political power devolve to the local units, for without both, the local units will be unsatisfactory. Simply creating governmental units on a local level, which pass only on a few policy issues and otherwise have only advisory powers, will be of little value. Thanks to

---

[96] This disempowerment brought about by excessive control from above explains, in part, the failure of the local school boards and Community Boards in New York City. *See infra* notes 275–280 and accompanying text.

[97] For a radical presentation of the need for government from the bottom, see P. GREEN, RETRIEVING DEMOCRACY 182–84 (1985).

[98] There is some empirical evidence that central governments are more responsive when there are strong local units. For an analysis of the impact of the system of powerful neighborhood organizations in Cincinnati, see J. THOMAS, BETWEEN CITIZEN AND CITY (1986).

[99] An example of political and administrative decentralization which reaches to the neighborhood level is the "governmental model" of decentralization, or, more broadly, community control. *See* Nordlinger & Hardy, *Urban Decentralization: An Evaluation of Four Models*, 20 PUB. POL'Y 359, 372 (1972).

[100] Unequal service delivery has been a subtle form of *de jure* discrimination. Boles, *Urban Equality: Definitions and Demands*, in THE EGALITARIAN CITY 9–10 (J. Boles ed. 1986). *See also supra* note 35.

[101] There is a substantial literature on such administrative decentralization. *See, e.g.*, Kaufman, *Administrative Decentralization and Political Power*, 29 PUB. ADMIN. REV. 3 (1969); *Curriculum Essays on Citizens, Politics and Administration in Urban Neighborhoods*, 32 PUB. ADMIN. REV. 565 (1972); D. YATES & R. YIN, STREET-LEVEL GOVERNMENTS: ASSESSING DECENTRALIZATION AND URBAN SERVICES (1975). Perhaps the most thorough attempt to apply the theory of administrative decentralization and measure its feasibility in relation to major government services is MCKINSEY & COMPANY, INC., THE IMPACT OF COTERMINAL SERVICE DISTRICTS ON THE DELIVERY OF MUNICIPAL SERVICES (1973), prepared for the New York State Charter Revision Commission.

Dillon's Rule, the political power of the central city government, from which the authority of the local unit began, was thin in the first place.

In order for these local units to translate into a successful means of enlivening American democracy, they must prove to be more than an additional layer of government. This is not an easy task. American political culture has come to view political participation for the vast majority of citizens to be limited to periodic—and increasingly infrequent[102]—trips to the voting booth, creating a polity devoid of a sense of responsibility for political decision. There are other problems as well. Those in the existing power structure will resist attempts to dislodge them from what they might see as their hard-earned place. One example crucial to local government decentralization is the reaction of municipal unions to administrative decentralization. These unions may well feel threatened by new political units with the power to rearrange service delivery patterns. It will be necessary, over time, to accommodate the unions.[103]

The positive impact on citizenship that comes from close contact with government can be reinforced in the structure of the local units. First, the local unit governments must be located in their communities. The local units must have representative deliberative bodies and all of the units' authority should rest in these deliberative bodies.[104] The deliberative bodies should meet frequently and at times convenient to the vast majority of the population of the local unit.[105] Public comment should be encouraged on every issue. Members of the deliberative body not present for public comment on an issue should be barred from participating in the decision on that issue. Finally, local unit deliberative bodies should be severely limited in their freedom

---

[102] *See supra* notes 18–21 and accompanying text.

[103] *See infra* notes 311, 328.

[104] In essence, this would create a governmental unit similar to the familiar city manager-council structure, with the manager simply implementing the decisions of the deliberative body. For a description of the manager-council form of government, see D. CARALEY, *supra* note 23, at 86–89. It is important that the arrangement be structured so that the manager cannot usurp the authority of the council or come to rule it. *Id.* at 226–47. For a discussion of the development of the council-manager system, see Lockard, *The City Manager, Administrative Theory and Political Power*, in URBAN POLITICS AND PROBLEMS 74 (H. Manhood & E. Angus ed. 1969).

[105] This will involve adjustment for each individual unit, but will undoubtedly involve meetings on weekends and evenings for many local units. It may even be desirable for the deliberative bodies to hold at least two sessions at different times during the week on each issue to allow maximum opportunity for participation.

to hold executive sessions,[106] and must vote on every issue in open session after a period of discussion and public comment.[107]

The procedures outlined above provide an opportunity for participation on issues directly affecting those who live in the local unit. These citizens may vote for representatives to the deliberative body and can either contact the individual representatives—who will be more accessible than current elected officials as they represent fewer individuals—or they can participate in the deliberative process of the local units more personally and with greater ease than they can participate in the central city government. However, citizens will not necessarily choose to participate, despite the structural, procedural, and political incentives.[108] This problem raises the second aspect necessary for successful participatory government: education.

Participatory government and civic education[109] work together in a dialectic. Civic education alone is a simple, if not uncontroversial,[110] idea. It involves teaching such "old-fashioned" no-

---

[106] The use of executive sessions to reach decisions prior to the public comment period has been one of the major criticisms of the functioning of the New York City Board of Estimate. For a description of this practice, see WHK COMMUNICATIONS ASSOCIATES, INC., THE STRUCTURE, POWERS AND FUNCTIONS OF NEW YORK CITY'S BOARD OF ESTIMATE 20–26 (1973) [hereinafter KRAMARSKY REPORT] (prepared for the New York State Charter Revision Commission).

[107] Most, if not all, of the actual attempts at governmental decentralization in the late 1960's and 1970's involved either appointed local bodies, or the selection of pre-existing local community groups to act as representatives of the central city administration in communities. For a description of the workings of the last type of decentralization in Boston, see E. NORDLINGER, DECENTRALIZING THE CITY, A STUDY OF BOSTON'S LITTLE CITY HALLS (1972); in New York, see NEW YORK STATE CHARTER REVISION COMMISSION & BUREAU OF APPLIED SOCIAL RESEARCH, COLUMBIA UNIVERSITY, OFFICE OF NEIGHBORHOOD GOVERNMENT (1973–74).

[108] *See* D. YATES & R. YIN, *supra* note 101, at 173–80 (critiques of governmental decentralization). *See also* S. DAVID & P. PETERSON, URBAN POLITICS AND PUBLIC POLICY 9–11 (1973) (arguing that decentralization leads to apathy toward a multitude of political offices and confusion over which officials and which organizations are responsible for which interests; and that lack of information, time and resources restricts involvement to organized groups, who, in turn, discourage participation by individuals).

[109] "Civic education" here means education in primary and secondary schools in the value of participation, deliberation, toleration, and choice in public affairs.

[110] A major source of controversy is the content of a civic education. Is it possible to have a socially and politically neutral educational system which teaches the values of civic involvement without taint? The answer may seem easy until the values of isolationist communities such as the Amish are taken into account. Professor Amy Gutmann has developed a theory of "democratic education" which urges the teaching of civic values and which admits the needs for the community to make choices among the values to be taught. A. GUTMANN, DEMOCRATIC EDUCATION (1987). There are some troubling, unanswered questions in Professor Gutmann's work, the most important of which is who will choose the values to be taught. Nonetheless, the work provides an excellent consideration of the purposes and practicality of civic education. For an earlier version of how civic education could work, see R. CLEARY, POLITICAL EDUCATION IN THE AMERICAN DEMOCRACY (1971).

tions as good citizenship and community values, partly through instruction and partly through example.[111]

Civic education through participation in local units is far more complex and nebulous. Tocqueville and his successors believe that through participation in their own government, members of a community learn civic virtue.[112] The notion of the educational value of participation is captured by Hanna Pitkin:

> Drawn into public life by personal need, fear, ambition or interest, we are there forced to acknowledge the power of others and appeal to their standards, even as we try to get them to acknowledge our power and standards. We are forced to find or create a common language of purposes and aspirations, not merely to clothe our private outlook in public disguise, but to become aware ourselves of its public meaning . . . . In the process, we learn to think about the standards themselves, about our stake in the existence of standards, of justice, of our community, even of our opponents and enemies in the community; so that afterwards we are changed. Economic man becomes citizen.[113]

Individuals will learn to be citizens engaged in debate, discussion, and argument about the shape of their society. Discussion will lead to understanding and interest. This is an evolution to occur over time, but one which, if successful, will invigorate the American polity and reinforce a fragile democracy.[114]

---

[111] Example can come in two forms: first, participatory student governments with faculty advisers to instruct young leaders in the functioning of a participatory institution; second, frequent use of model political bodies such as model United Nations or mock constitutional conventions or legislatures.

[112] Dahl, *supra* note 92, at 953; Brest, *Further Beyond the Republican Revival: Toward Radical Republicanism*, 97 YALE L.J. 1623, 1624 (1988); S. ELKIN, *supra* note 51, at 148–53. *Cf.* J. MANSBRIDGE, BEYOND ADVERSARY DEMOCRACY 300 (1980) (By keeping unitary governments small in size, citizens can participate in the process of government and learn to adopt "democratic procedures for dealing with common and conflicting interests").

[113] Pitkin, *Justice: On Relating Private and Public*, 9 POL. THEORY 347 (1981), *quoted in* S. ELKIN, *supra* note 51, at 149.

[114] Government is not the only method of encouraging participation in the life, health, and governing of the community. Voluntary associations and community corporations may also provide mechanisms for participation. Community corporations may, in fact, be one of the best ways to promote economic growth within a local unit area. For innovative suggestions for structuring such corporations, see Miller, *Community Capitalism and the Community Self-Determination Act*, 6 HARV. J. ON LEGIS. 413 (1969). The problems of parochialism and exclusion are likely to be more troublesome in such organizations, just as they are in homeowner associations, and need to be guarded against. *See* Note, *The Rule of Law in Residential Associations*, 99 HARV. L. REV. 472 (1985).

The task of modeling such a participation-oriented government for New York remains.[115] While there is no doubt that New York City is unique among American urban areas, the problems of size addressed by the model presented in Part III are present in all large cities and the model can be adapted to fit local realities.[116] Before presenting a model of local unit structure for New York City, it is necessary to examine the creation and structure of the current government and to explore prior efforts to decentralize the City's government.

II. THE HISTORY AND STRUCTURE OF NEW YORK CITY GOVERNMENT

A. *The History of the City Government*

The City of New York in its present geographic and governmental form came into being on January 1, 1898.[117] The govern-

[115] The newly adopted Charter Revision makes only limited attempts to increase participation. The Community Boards remain virtually unchanged; they have been granted only a slightly enhanced role in the land use review process. FINAL PROPOSALS, *supra* note 7, at 47–48. The Community Boards have not been significantly empowered, nor has any other local unit been created to increase local participation.

[116] While this Note does not deal specifically with problems of suburban government, as the suburbs continue to grow, they have or will develop many of the same governmental problems of distance and disempowerment as cities. *See generally* C. HAAR, THE END OF INNOCENCE: A SUBURBAN READER (1972) (detailing the problematic transformations caused by the explosive growth of the American suburbs).

[117] Until 1874 New York City consisted simply of Manhattan Island. In that year, as a first step toward consolidation of the "Greater City," three western townships—Morrisania, West Farms, and Kingsbridge—in what is now the Bronx were annexed to the City. 1874 N.Y. Laws ch. 329. In 1895, the remainder of what is now the Bronx—the villages of Wakefield, Eastchester, Williamsbridge, the town of Westchester, and portions of the towns of Eastchester and Pelham—was annexed. 1895 N.Y. Laws ch. 934 (Vol. II, Part II). The final consolidation added to the City the counties of Kings and Richmond and the towns and one city now covering the area of Queens County. 1894 N.Y. Laws ch. 64, § 1 (Vol. I); 1896 N.Y. Laws ch. 448, § 1 (Vol. II); 1897 N.Y. Laws ch. 378, § 1 (Vol. III). For an account of the movement to consolidation, see W. SAYRE & H. KAUFMAN, GOVERNING NEW YORK CITY 11–14 (1965). The consolidation both created the City of New York and accomplished the merger of the governments of the independent towns and cities of Long Island and Brooklyn. 1896 N.Y. Laws ch. 488, § 1 (Vol II); 1897 N.Y. Laws ch. 378, § 1 (Vol. III). The consolidation was approved by the State Legislature in 1896. 1896 N.Y. Laws ch. 488 (Vol. II). The first Charter for the City was not approved until May 1897, W. SAYRE & H. KAUFMAN, *supra*, at 13, and it went into effect on January 1, 1898. 1897 N.Y. Laws ch. 378, § 1611 (Vol. III). The consolidation ended a fierce political battle that had begun in earnest seven years earlier with the creation of the Municipal Consolidation Commission. This body had been formed "to inquire into the expediency of consolidating the various municipalities in the state of New York, occupying the several islands in the Harbor of New York." 1890 N.Y. Laws ch. 311 (Vol. III).

ment of this unprecedentedly large municipality[118] was laid out in the extensive, though hastily prepared,[119] Greater New York Charter.[120] The difficulty inherent in creating any governmental structure was increased in this case because the new city government had to satisfy the residents of what had been three separate cities and a number of independent towns.[121] Furthermore, the mandate for consolidation from the voters of the various municipalities involved had not been overwhelming, particularly in Brooklyn.[122] The municipalities now consolidated could not be left to feel completely disempowered, or consolidation might fail.[123] The solution was the creation of a two-tiered government. The first, higher tier consisted of the central government for the consolidated City. The second tier involved the division of the City into five geographical sections called boroughs.[124] Each borough had its own governmental structure headed by an elected president.[125]

Additional decentralization was instituted through the two-tiered governmental structure. The Charter created twenty-two Local Improvement Districts,[126] each with a board consisting of the borough president and the members of the Municipal As-

---

[118] In 1900, the consolidated City of New York contained almost three and a half million people. W. SAYRE & H. KAUFMAN, *supra* note 117, at 11.

[119] The Charter contained 1620 sections. 1897 N.Y. Laws ch. 378. The Charter Commission had only a few months to complete its work. W. SAYRE & H. KAUFMAN, *supra* note 117, at 13–14; A. MACMAHON, STATUTORY SOURCES OF NEW YORK CITY GOVERNMENT 15 (1923); F. SHAW, THE HISTORY OF THE NEW YORK CITY LEGISLATURE 8 (1954). One of the members of the Charter Commission, Seth Low—who served at various times as mayor of Brooklyn, mayor of the City of New York, and president of Columbia University—admitted that time pressures on the Commission may have caused some inadequacies in the final product. A. MACMAHON, *supra*, at 16. For a description of the care—or lack of it—with which the Charter Commission acted, see Pryor, *The Greater New York Charter*, 10 ANNALS AM. ACAD. POL. & SOC. SCI. 1, 25–28 (1897).

[120] On May 4, 1897, the New York State Legislature adopted the new charter, which was officially titled the Greater New York Charter. GREATER N.Y. CHARTER § 1, at 1897 N.Y. Laws ch. 378, [hereinafter GREATER N.Y. CHARTER].

[121] *Report Accompanying the Proposed Greater New York Charter*, in GREATER N.Y. CHARTER iii, ix–x [hereinafter *1897 Commission Report*].

[122] In the 1894 referendum on consolidation, the voters of the City of Brooklyn had approved consolidation by only 277 votes. W. SAYRE & H. KAUFMAN, *supra* note 117, at 12.

[123] The Charter Commission tacitly recognized this problem in its report. "Manifestly one of the most difficult problems that the Commission has had to meet has been to determine to what extent and how the interests of the different localities in the great city could be provided for . . . ." *1897 Commission Report*, *supra* note 121, at xxvi.

[124] GREATER N.Y. CHARTER, *supra* note 120, § 2. The boroughs are Manhattan, the Bronx, Richmond (Staten Island), Queens, and Brooklyn. *Id.*

[125] *Id.* § 382.

[126] *Id.* § 390.

sembly from that district.[127] The boards had limited power both to recommend capital improvements within their districts and to act as local ombudsmen.[128] Despite the appearance of an equitable distribution of power, however, the reality was that almost all authority rested with the central City government.[129] The central government was run by a popularly elected mayor and a bicameral legislature known as the Municipal Assembly.[130] Preparation of the City budget was left to the Board of Estimate and Apportionment.[131] A Board of Public Improvements regulated capital expenditures.[132]

The drafting of the original Charter, and almost all attempted and successful revisions of it, were conducted under the auspices of the state legislature. The legislature set up the Charter Commission, outlined its duties, and oversaw its work. Until Home Rule allowed local input into Charter revisions,[133] the Charter and any changes in it were completely controlled by the state legislature.

The effects of the hasty drafting of the Charter quickly began to show.[134] The Municipal Assembly was tightly restricted in its powers,[135] cumbersome in its operations, and generally ineffective.[136] The outer boroughs also chafed at their lack of power.[137] These structural problems, combined with a political uprising

---

[127] *Id.* § 391.

[128] *Id.* § 393.

[129] NEW YORK CITY CHARTER REVISION COMMISSION, REPORT OF THE CHARTER REVISION COMMISSION OF 1907 27 (1907) [hereinafter 1907 COMMISSION REPORT].

[130] The Assembly consisted of a 60-member Board of Aldermen (the lower house) and a 29-member City Council (the upper house). GREATER N.Y. CHARTER, *supra* note 120, §§ 18, 19, 24; NATIONAL MUNICIPAL LEAGUE, THE GOVERNMENT OF METROPOLITAN AREAS IN THE UNITED STATES 345 (1930).

[131] The Board of Estimate and Apportionment consisted of the Mayor, the Comptroller, the President of the City Council, the President of the Department of Taxes and Assessment, and the Corporation Counsel. GREATER N.Y. CHARTER, *supra* note 120, § 226. Both the Comptroller and the President of the City Council were, and still are, popularly elected city-wide. *Id.* §§ 18, 149.

[132] *Id.* §§ 410–426. The Borough Presidents sat on the Board of Public Improvements but could only vote on matters relating to their own boroughs. *Id.* § 410.

[133] *See supra* notes 9, 76–78 and accompanying text; *infra* notes 152–155 and accompanying text.

[134] F. SHAW, *supra* note 119, at 9–12.

[135] The 1897 Charter Commission had intentionally restricted the Municipal Assembly, citing the history of poor experience with local legislative bodies. *1897 Commission Report*, *supra* note 121, at xii.

[136] F. SHAW, *supra* note 119, at 10–11; Goodnow, *The Charter of the City of New York*, 17 POL. SCI. Q. 1, 7–8 (1902).

[137] Goodnow, *supra* note 136, at 16.

by those forces excluded from the first City government,[138] led to the creation of a new commission to revise the less than two-year-old Charter.[139] The revisions made by that commission created a governmental structure that has survived in its basic form to this day.[140] The central government was reconstituted in three governing bodies: the Mayor, the Board of Aldermen, and the Board of Estimate and Apportionment ("Board of Estimate").[141]

The new Charter also took steps towards increasing the authority of the Borough Presidents.[142] First, the Board of Estimate was restructured to contain three city-wide officials[143] and the five Borough Presidents.[144] The three city-wide officials could out-vote the Borough Presidents.[145] Nonetheless, the presence of the Borough Presidents on the Board gave these formerly powerless officials a substantially increased role in City government, particularly since under the new Charter the Board of Estimate was the most important of the City's governing bodies.[146] Furthermore, the Borough Presidents were members of the Board of Aldermen.[147]

Actual governmental decentralization was instituted by the new Charter. The Charter devolved to the Borough Presidents some of the administrative control over City services exercised by the Mayor under the 1898 Charter. Some services were divided into five departments—one for each borough—with each department reporting to its respective Borough President.[148] In

---

[138] Charo, *Designing Mathematical Models to Describe One-Person, One-Vote Compliance By Unique Governmental Structures: The Case of the New York City Board of Estimate*, 53 FORD. L. REV. 735, 743 (1985); W. SAYRE & H. KAUFMAN, *supra* note 117, at 15–16.

[139] This commission was again created by state legislative action. 1900 N.Y. Laws ch. 465.

[140] The revised Charter was passed by the state legislature and approved by the Governor on April 22, 1901. 1901 N.Y. Laws ch. 466.

[141] *See* NATIONAL MUNICIPAL LEAGUE, *supra* note 130, at 351–54.

[142] *See* Goodnow, *supra* note 136, at 14–15.

[143] The Mayor, the Comptroller, and the President of the Board of Aldermen.

[144] M. ASH & W. ASH, THE GREATER NEW YORK CHARTER § 226 (1901) [hereinafter 1901 CHARTER].

[145] The three city-wide officials each had three votes, for a total of nine. The Borough Presidents had a total of seven votes: the Presidents of Manhattan and Brooklyn had two votes apiece while the remaining three Borough Presidents each had one vote. KRAMARSKY REPORT, *supra* note 106, at 5.

[146] The Board of Estimate retained its former budgetary powers and also inherited the powers of the Board of Public Improvements, which was eliminated by the new Charter. *Id.* at 5–6; NATIONAL MUNICIPAL LEAGUE, *supra* note 130, at 352.

[147] 1901 CHARTER § 18.

[148] NATIONAL MUNICIPAL LEAGUE, *supra* note 130, at 354.

addition, six city-wide departments were abolished and their functions transferred to the Borough Presidents.[149]

After 1901 there was no substantial revision of the Charter for over thirty years.[150] The powers of the Board of Estimate continued to grow,[151] however, correspondingly increasing the power of the Borough Presidents. In 1923, New York State added a Home Rule amendment to its constitution,[152] and the legislature passed a corresponding Home Rule law in 1924.[153] The Home Rule law officially designated the Board of Estimate as the upper house of the New York City legislature.[154]

This act legally recognized the Board of Estimate's growing importance in City government and gave it a traditional place in the governmental structure.[155] The power of the Board of Estimate was further increased by the near collapse of the Board of Aldermen as an effective governing body.[156] As the Board of

[149] The departments of highways, sewers, building, public buildings, lighting, and supplies were abolished. *Id.* For a discussion of the reasons for this shift of control to the boroughs, see Goodnow, *supra* note 136, at 16–18.

[150] There was no shortage of attempts at revision, however. Revision commissions presented proposed charters in 1907, 1907 COMMISSION REPORT, *supra* note 129, followed by other attempts in 1909, 1911, and 1923. NATIONAL MUNICIPAL LEAGUE, *supra* note 130, at 358. Almost all of these revisions called for the recentralization of governmental powers. *Id.* at 359.

[151] The new Charter fueled this growth by giving the Board of Estimate all residual municipal powers. Charo, *supra* note 138, at 744. In 1905, franchising power was shifted to the Board of Estimate due to a scandal involving the Board of Aldermen's granting of a terminal franchise to the Pennsylvania Railroad. 1905 N.Y. Laws ch. 629. In 1911, the power to authorize improvements was assigned to the Board of Estimate. 1911 N.Y. Laws ch. 679. The Board gained authority in 1916 to create zoning regulations. 1916 N.Y. Laws ch. 497. *See generally* KRAMARSKY REPORT, *supra* note 106, at 6–7. Moreover, a failed attempt was made in 1915 to merge the still-existing county governments within the City into the City government. If it had been successful, this consolidation would have further increased the power of the Board of Estimate. *See* H. BRUERE & L. WALLSTEIN, STUDY OF COUNTY GOVERNMENT WITHIN THE CITY OF NEW YORK AND A PLAN FOR ITS REORGANIZATION (1915).

[152] N.Y. CONST. art. XII (1923)

[153] 1924 N.Y. Laws ch. 363.

[154] *Id.*

[155] At the time of their passage, and for a period afterwards, the Home Rule amendments were considered a major step forward for the City. NATIONAL MUNICIPAL LEAGUE, *supra* note 130, at 360. The actual impact of Home Rule has already been discussed. *See supra* notes 76–78 and accompanying text. In 1942, the New York Court of Appeals explained, in definite terms, the lack of authority granted by Home Rule: "[A] city is not sovereign, as are the federal government and the states. 'A municipal corporation is, so far as its purely municipal relations are concerned, simply an agency of the state for conducting the affairs of government . . . .'" LaGuardia v. Smith, 288 N.Y. 1, 7, 41 N.E.2d 153, 155 (1942) (quoting Williams v. Eggleston, 170 U.S. 304, 310 (1898)).

[156] Critics constantly derided the Board of Aldermen for its corruption and incompetence and often referred to the body as the "Boodle Board" or "Forty Thieves." W. SAYRE & H. KAUFMAN, *supra* note 117, at 617. For a comprehensive look at the state of the City's legislative branch during this period, see F. SHAW, *supra* note 119, at 15–109.

Aldermen weakened, the Board of Estimate assumed more authority simply by taking advantage of the Aldermen's inaction.[157]

The 1933 election of Fiorello LaGuardia as Mayor on a reform-fusion ticket, after the scandals of Jimmy Walker's Democratic administration, created an atmosphere ripe for restructuring.[158] In 1934, LaGuardia obtained from the state legislature the authority[159] to appoint a charter revision commission. This body became known as the Thacher Commission, after its chairman.[160] The Thacher Commission completed its report in 1936, and despite opposition from four Borough Presidents and the Democratic organization,[161] the voters overwhelmingly adopted the Commission's recommendations.[162]

The revisions made in the new Charter were the most radical in the City's history—before or since. The Board of Estimate was no longer designated the upper house of the municipal legislature.[163] Instead, under the new Charter, the Board of Estimate was made the City's chief administrative body, complete with substantial budgetary authority.[164] The administrative power of the Borough Presidents was reduced as some services over which they had been granted authority under the 1901 Charter reverted to agencies of the central City government.[165] The Borough Presidents were also removed from the Board of Aldermen.[166]

The most significant changes affected the Board of Aldermen. That body, by this time completely discredited,[167] was replaced

---

[157] *See* Charo, *supra* note 138, at 745.

[158] While scandal has often been the impetus for Charter reform, the organizational arrangements outlined in the Charter cannot thoroughly address the underlying causes of corruption. Structural changes can eliminate from office those who have most recently plundered the public till, but such reform, in general, cannot prevent new corrupt officials from gaining power. Public vigilance, which can be enhanced by moving government closer to the governed, may be one of the best ways to reduce the opportunities for corruption.

[159] 1934 N.Y. Laws ch. 689. For more on the creation and powers of this commission, see L. TANZER, THE NEW YORK CITY CHARTER 1–5 (1937).

[160] KRAMARSKY REPORT, *supra* note 106, at 7–8.

[161] *Id.* at 10.

[162] *Id.*; *see also* Hallet, *The New York Victory*, 25 NAT'L MUN. REV. 745, 746 (1936).

[163] KRAMARSKY REPORT, *supra* note 106, at 9.

[164] *Id.* at 10–11.

[165] *Id.* at 9–10.

[166] L. TANZER, *supra* note 159, at 27.

[167] Henry Curran, former majority leader of the Board of Aldermen, told the following story about his own organization: "A New Yorker was showing the sights to his friend from out of town. As they entered the aldermanic chamber, while the meeting was going on, the stranger stared in amazement. 'Are these the aldermen?' he asked. 'They are.' 'But they're all fast asleep—oughtn't we do something about it?' 'Leave them be, leave

by a City Council. Though major changes were not made in the duties of the Council, the process of selecting the Council changed radically. The voters approved a proposal by the Thacher Commission to elect councilmen by proportional representation.[168] Under the new plan, voters were to select among councilmanic candidates from their boroughs, electing one council member for every 75,000 valid ballots cast.[169]

The proportional representation system[170] was intended to break the stranglehold the Democratic organization traditionally had on the Board of Aldermen.[171] In that respect, proportional representation worked well.[172] Over the next ten years, encompassing five councilmanic elections, the composition of the council changed to include previously excluded groups, which resulted in a more lively, active Council which generated substantially increased citizen interest.[173] However, proportional representation was able to survive for only ten years in New York City. It succumbed to its own success and to changes in the external political atmosphere.[174] The resulting return to plu-

---

them be,' counselled the New Yorker. 'While they sleep, the city's safe.'" F. SHAW, *supra* note 119, at 56.

[168] Proportional representation had been a *cause célèbre* among government reformers. One of the leaders of the movement, George H. Hallet, Jr., edited a section on proportional representation in virtually every issue of the NATIONAL MUNICIPAL REVIEW in the 1930's. Hallet hailed the passage of the new Charter in 1936, Hallet, *supra* note 162, and continued his quest by producing a primer whose title bespoke his message. *See* G. HALLET, PROPORTIONAL REPRESENTATION, THE KEY TO DEMOCRACY (1937).

[169] L. TANZER, *supra* note 159, at 32.

[170] Although there are a number of proportional representation schemes, New York chose the Hare system. Under the Hare system as implemented in New York, the ballot contained the names of all the candidates from that borough, and a voter would select candidates in order of preference. If a voter's first choice candidate had already received sufficient votes for election or had already been clearly defeated, the ballot would be counted to the candidate the voter placed second, and so on down the preference list. The results should closely track the voting allegiances of the electorate, so that in an election where 40% of the voters are Democratic, 40% Republican, 10% Liberal, and 10% Labor, a ten-member council should contain four Democrats, four Republicans, a Liberal, and a Laborite. *See* Charo, *supra* note 138, at 746 n.64.

[171] W. SAYRE & H. KAUFMAN, *supra* note 117, at 176; Charo, *supra* note 138, at 746; F. SHAW, *supra* note 119, at 127, 193.

[172] Not only was minority party representation increased, but the percentage of seats won matched almost exactly the percentage of votes cast for each party, just as the Hare system had intended. *See* Zeller & Bone, *The Repeal of P.R. in New York City—Ten Years in Retrospect*, 42 AM. POL. SCI. REV. 1127, 1132 (1948). For an explanation of the Hare system, *see supra* note 170.

[173] McCaffrey, *Proportional Representation in New York City*, 33 AM. POL. SCI. REV. 841, 849–50 (1939).

[174] The party organizations had been opposed to proportional representation from the start and had fought it constantly from its inception in 1936. Two external forces united shortly after World War II to give the anti-proportional representation forces sufficient ammunition to defeat the system. First, proportional representation had been used in

rality voting in a district system brought one-party rule back to the City.

The return to district voting also eliminated the additional legitimacy the boroughs had enjoyed from being the basic unit for the proportional representation elections. Hence, it was only through the Board of Estimate that the Borough Presidents had any real power.[175] This weakness was exacerbated by the next revision of the Charter, in 1961, which removed from the Borough Presidents all public works authority and gave these powers to city-wide agencies.[176]

B. *The Movement Toward Decentralization*

The 1961 Charter did lay the groundwork for City government decentralization. In 1951, Manhattan Borough President Robert F. Wagner established twelve Community Planning Councils as local outlets for a citizen voice on public decisions.[177] The 1961 Charter revision extended this concept of local councils to all the boroughs by renaming them Community Planning Boards and placing them under the jurisdiction of the City Planning Commission and the Borough Presidents.[178] Members of the Community Planning Boards, whose authority was purely advisory, were appointed by their respective Borough Presidents.[179] While the impact of the Community Planning Boards on City government was minimal,[180] the Boards did serve as a

---

Germany during the Weimar Republic, and arguably had allowed the Nazis to come to power. Second, the post-war "Red Scare" allowed the party organizations to capitalize on the fact that Communist Party members had been elected to the Council in three of the five elections held under proportional representation. For detailed discussions of the rise and fall of proportional representation in New York City, see F. SHAW, *supra* note 119, at 188–210; Zeller & Bone, *supra* note 172.

[175] W. SAYRE & H. KAUFMAN, *supra* note 117, at 638.

[176] T. SMITH, GUIDE TO THE MUNICIPAL GOVERNMENT OF THE CITY OF NEW YORK 56 (1973). The powers of the Borough Presidents are described at N.Y.C. CHARTER § 82, *reprinted in* NEW YORK STATE DEPARTMENT OF STATE, LOCAL LAWS OF THE CITIES, COUNTIES AND VILLAGES 249 (1962) [hereinafter 1961 CHARTER].

[177] NEW YORK STATE CHARTER REVISION COMMISSION, COMMUNITY BOARDS 18 (1974) [hereinafter COMMUNITY BOARDS].

[178] 1961 CHARTER § 84.

[179] *Id.* This Charter provision was criticized as providing inadequate power, and thus § 84 was reenacted in 1968 by New York City Local Law 39, changing the name of the bodies to "Community Boards" and broadening their mission. Their powers remained advisory and their membership appointed.

[180] *See* D. YATES, *supra* note 52, at 39–43 (documenting the Boards' inability to make decisions, take action, or follow through on their actions). *But see* COMMUNITY BOARDS, *supra* note 177 (strongly supporting the potential of the Community Boards as a force for local control while recognizing their limited powers and inadequacies).

focal point for a growing governmental decentralization movement. This crusade engendered substantial intellectual interest, if limited success, under the Lindsay administration.[181]

The Lindsay administration came in on the wave of a national spirit of reform energized, in part, by the Civil Rights Movement. The movement for greater citizen participation was an element of this national call for reform.[182] Lindsay's first effort to decentralize occurred within a few months of his inauguration when the new Mayor proposed creation of "Little City Halls" in neighborhoods that would act as multi-service/ombudsman centers.[183] The idea was met with skepticism from those who believed the Mayor was trying to create political power bases,[184] and Lindsay was forced to turn to private funding sources to establish such centers in a few neighborhoods.[185]

While Lindsay's first effort at decentralization was encountering obstacles, the federal government was beginning to expand its own neighborhood programs. The most expansive community participation program came from the Economic Opportunity Act of 1964.[186] Congress stated in the Act that it wanted to create opportunities for "maximum feasible participation" by residents of areas affected by the federal programs covered by the Act.[187] The participatory element of the Act became known as the Community Action Program.[188] At the local level, this participation was accomplished through neighborhood community corporations with large boards which were initially appointed and later elected.[189]

The Model Cities Program was a second federal effort to encourage local participation.[190] Model Cities created local pol-

---

[181] John Lindsay was elected in 1965 and served through 1973.

[182] Citizen participation and empowerment were seen to go hand-in-hand. The Voting Rights Act of 1964 took the first step toward enfranchising minorities. The next logical step was to open opportunities for people previously shut out of the government to participate in their own governing.

[183] C. MORRIS, *supra* note 24, at 28.

[184] *Id.*

[185] Note, *Conflict Resolution in a Politically Decentralized Local Government System*, 11 COLUM. J. L. & SOC. PROBS. 633 n.15 (1975). Little City Halls were more successful in other cities, most notably Boston. E. NORDLINGER, *supra* note 107.

[186] Pub. L. No. 88-452, 78 Stat. 508 (codified as amended in scattered sections of 42 U.S.C. beginning at § 2701 (1973)).

[187] 42 U.S.C.A. § 2781(a)(4) (West 1973).

[188] For a description of the goals of the Community Action Program created by the statute, see NEW YORK STATE CHARTER REVISION COMMISSION, THE COMMUNITY ACTION EXPERIENCE 5–10 (1973).

[189] *Id.* at 30–31.

[190] Demonstration Cities and Metropolitan Development Act of 1966, Pub. L. No. 89-

icy committees composed of both elected and appointed members.[191] These committees, like the community corporations of the Community Action Program, were quite large,[192] and their functions were limited to the reach of the federal program which created them.[193]

Despite its difficulties with the Little City Halls program, the Lindsay administration continued to look for ways to increase neighborhood contact with City government. In June 1970, Lindsay presented a plan which aimed to reduce citizen alienation by increasing the power of the Community Boards. The Mayor proposed giving the Community Boards full-time staff, more connection to City services, and a community cabinet.[194] This proposal was never implemented, but it laid the groundwork for a plan to decentralize the administration of city services, which was put into effect one year later.[195] The program was centered around the new Office of Neighborhood Government in the Mayor's Office. This office coordinated District Service Cabinets which were initially established in eight neighborhoods.[196] These cabinets, consisting of the local district managers of eight city service agencies,[197] were charged with improving the coordination and responsiveness of those services in their district.[198] Although the cabinets were not originally designed for direct citizen participation, in 1973 they were linked to the Community Boards to increase citizen input.[199] The Lindsay administration implemented other, more narrowly tailored efforts at increasing community participation which included

---

754, 80 Stat. 1255 (codified as amended in scattered sections of 12 U.S.C., 15 U.S.C. and 42 U.S.C. (1973)).

[191] NEW YORK STATE CHARTER REVISION COMMISSION, STRUCTURAL ISSUES FOR LOCAL UNITS OF GOVERNMENT: LOCAL COUNCIL AND DISTRICT EXECUTIVE 6–7 (1974) [hereinafter STRUCTURAL ISSUES].

[192] *Id.* at 7. Twenty-six community corporations and three local policy committees were created in New York City. *Id.*

[193] *Id.*

[194] *Id.* at 8. The Mayor's proposal was entitled, "Plan for Neighborhood Government for New York City." *Id.*

[195] This proposal was entitled, "Program for the Decentralized Administration of Municipal Services in New York City Communities." *Id.* at 8–9.

[196] *Id.* at 9.

[197] The agencies were: Police, Environmental Protection, Housing and Development, Health Services, Parks, Recreation and Cultural Affairs, Transportation, and Addiction Services. STRUCTURAL ISSUES, *supra* note 191, app. H, at ii, (Lindsay's "Program for the Decentralized Administration of Municipal Services in New York City Communities").

[198] *Id.* app. H, at iv.

[199] STRUCTURAL ISSUES, *supra* note 191, at 9–10.

establishing Comprehensive Health Planning Districts[200] and community advisory boards to a variety of City agencies.[201]

The most far-reaching[202] and controversial of the Lindsay Administration's decentralization efforts was the creation of the Community School Boards in 1969.[203] The decentralization of the public school system came after almost two decades of constant criticism of the New York City Board of Education. According to its critics, the Board of Education was inefficient, unresponsive, uninnovative, and out of touch with the needs of its pupils, particularly minorities and immigrants.[204]

The sheer size of the school system was one of the most obvious problems.[205] The centralized bureaucracy necessary to run such a system was vast and distant from those it served. The Board of Education was inflexible,[206] unable to respond to the changing conditions in the schools,[207] and to the changing demands of the parents and students. This reproach is similar to the criticism currently aimed at large municipal governments.[208]

In response to the failures of the Board of Education, a steady effort was made to increase parent and community participation in operating the schools. A series of scandals at the Board of Education in 1961 forced the state legislature to re-examine the City's public education system.[209] That review spawned the cre-

---

[200] *Id.* at 10–11.

[201] *Id.* at 12. Community advisory boards were established for many agencies including Police, Health and Hospitals, Addiction Services, and Environmental Protection. *Id.*

[202] "The only major municipal function that has been substantially decentralized in New York City is the public and secondary education system." NEW YORK STATE CHARTER REVISION COMMISSION, SCHOOL DECENTRALIZATION IN NEW YORK CITY 2 (1974) [hereinafter SCHOOL DECENTRALIZATION REPORT].

[203] The Community School Boards were created pursuant to state law. 1969 N.Y. Laws ch. 330. The Community School District law is codified at N.Y. EDUC. LAW §§ 2590–2590n (McKinney 1981).

[204] For a comprehensive review of the ills of the Board of Education in the 1950's and 1960's, see generally D. ROGERS, 110 LIVINGSTON STREET (1968); D. RAVITCH, THE GREAT SCHOOL WARS 251–66 (1974).

[205] The New York City school system is the largest in the nation, serving approximately 940,000 students in over 1100 school buildings. NEW YORK CITY MAYOR'S OFFICE OF OPERATIONS, MAYOR'S MANAGEMENT REPORT 361, 383 (preliminary ed. Feb. 15, 1989).

[206] For a description of the restraints on the central Board of Education in 1960, see W. SAYRE & H. KAUFMAN, *supra* note 117, at 279–85.

[207] SCHOOL DECENTRALIZATION REPORT, *supra* note 202, at 29.

[208] *See supra* notes 50–52 and accompanying text.

[209] *See* Comment, *Decentralization: The Respective Powers of the City Board of Education and the Community School Boards*, 5 FORD. URB. L.J. 239, 243 (1977) (noting that charges of the Board's corruption and malfeasance prompted the convening of a special session of the state legislature).

ation of local school boards intended to increase community involvement in school operations.[210] These local boards were, however, purely advisory and appointed by the Board of Education.[211]

Dissatisfaction with the operation of the school system continued to grow during the 1960's, especially among minority groups.[212] The crisis of confidence in the schools became linked with the national calls for increased neighborhood control of government and created movement toward decentralization. In 1967, the Board of Education announced its commitment to decentralize policymaking and established three demonstration districts in which the local school boards were given increased authority over the schools in their districts.[213] Meanwhile, the state legislature and the City moved toward legislation which would completely restructure the school system and shift much of the authority to community school boards.[214]

Careful consideration of school decentralization plans became impossible when, in the spring of 1968, the experimental school board in Ocean Hill-Brownsville tried to transfer nineteen ineffective teachers out of the district. This lead to a confrontation whose antagonists included the teachers' union, the Board of Education, the local school board, and community groups.[215] The teachers' union, which objected strenuously to the involvement of the community in personnel decisions, went on strike three times in a short period in an attempt to force changes in

---

[210] 1962 N.Y. Laws ch. 615.

[211] Id. § 3.

[212] As in many cities, there was a fierce battle over integration of the school system. D. RAVITCH, *supra* note 204, at 267–79. The most important of these battles took place in East Harlem in 1966 over the status of the newly constructed I.S. 201. The conflict focused on both integration and community control, and that fight helped set the stage for increased school decentralization. *Id.* at 292–311.

[213] M. ZIMET, DECENTRALIZATION AND SCHOOL EFFECTIVENESS 8–9 (1973). The three demonstration districts were in the East Harlem neighborhood surrounding I.S. 201, in Ocean Hill-Brownsville in Brooklyn, and in the Two Bridges neighborhood on the Lower East Side of Manhattan. SCHOOL DECENTRALIZATION REPORT, *supra* note 202, at 37.

[214] M. ZIMET, *supra* note 213, at 9–10.

[215] The conflict escalated when the Board of Education suspended the local board for refusing to reinstate the transferred teachers. Comment, *supra* note 209, at 261. The local board challenged its suspension in court, but the authority of the Board of Education over the local board was upheld, and the local board's suit was dismissed. Ocean Hill-Brownsville Governing Bd. vs. Board of Educ., 301 A.D.2d 447, 294 N.Y.S.2d 134 (2d Dept. 1968), *aff'd*, 23 N.Y.2d 483, 245 N.E.2d 219, 297 N.Y.S.2d 568 (1969). For a thorough discussion of the Ocean Hill-Brownsville battles, see D. RAVITCH, *supra* note 204, at 320–78; N. LEVINE, OCEAN HILL-BROWNSVILLE: SCHOOLS IN CRISIS (1969); M. BERUBE & M. GITTELL, CONFRONTATION AT OCEAN HILL-BROWNSVILLE (1969).

the decentralized system which would ensure teachers' freedom from control by the local boards.[216]

The Board of Education and the state legislature moved quickly to defuse the tension. The bill[217] that came out of the state legislature in 1969 was the result of intense political pressure, quick drafting, and compromise among various groups vying for control of the schools: parents, community activists, the current Board of Education, and particularly and most successfully, the teachers.[218] The new law created elected Community School Boards which had some control over local primary and intermediate schools. Much of the central authority over those schools, however, and all control over high schools was retained by the Board of Education.[219]

C. *Proposals for a Decentralized City Government*

The creation of elected Community School Boards and the various Lindsay Administration experiments with community-based programs caused a surge of interest in broader decentralization of City government.[220] A series of proposals from academics and planners suggested relocation of urban governmental authority in the neighborhoods.[221] When combined with the Lindsay drive for decentralization, these proposals brought forth a number of models of a decentralized New York City government.

Three of the proposals are particularly noteworthy. The first came from two Yale Law School students interning at the City's Bureau of the Budget in the summer of 1969.[222] Like the Lindsay plans of 1970 and 1971, this proposal was in many ways simply

---

[216] *See* D. YATES, THE UNGOVERNABLE CITY 127–28 (1977).

[217] 1969 N.Y. Laws ch. 330, § 4 (codified at N.Y. Ed. Law § 2590 (McKinney 1981)).

[218] M. ZIMET, *supra* note 213, at 11.

[219] *Id.* at 12; D. RAVITCH, *supra* note 204, at 387.

[220] The Lindsay Administration proposals of 1970 and 1971 increased neighborhood participation in government, providing at least some of the momentum. *See supra* notes 194–198 and accompanying text.

[221] Four of the most influential works urging urban decentralization, which acted as backdrops for more specific proposals for New York City, are A. ALTSHULER, COMMUNITY CONTROL: THE BLACK DEMAND FOR PARTICIPATION IN LARGE AMERICAN CITIES (1970); M. KOTLER, NEIGHBORHOOD GOVERNMENT: THE LOCAL FOUNDATIONS OF POLITICAL LIFE (1969); L. MUMFORD, THE URBAN PROSPECT (1968); Babcock & Busselman, *Citizen Participation: A Suburban Suggestion for the Central City*, 32 LAW & CONTEMP. PROBS. 220 (1967).

[222] Danzig & Heineman, *Decentralization in New York City: A Proposal*, 8 HARV. J. ON LEGIS. 407 (1971).

an extension of the existing Community Boards. The authors suggested creating sixty-two neighborhood councils each serving 130,000 people.[223] The councils were to be small and only partially elected,[224] and were to have power to allocate resources for some City services within their districts.[225] The councils, however, seemed destined to have no true connection to the central government,[226] and funding of services in their neighborhoods was left explicitly to the will of the City's budget process—into which the councils had no real input.[227]

The other two major decentralization proposals were closely linked both by personnel and by conclusion. The first, sponsored in 1972 by the New York City Bar Association, presented itself as "A Practical Study of a Radical Proposal for New York City."[228] The second proposal came in the same year from the New York State Study Commission for New York City, known as the Scott Commission.[229] Two of the authors of the Bar Association report, Walter Farr and Jeffrey Wood, served as staff to the Scott Commission, and one of the authors of the Scott Commission report, Edward Costikyan, was a member of the committee overseeing the Bar Association report.

As reflected in its title, the Bar Association's proposal is the more politically radical of the two. The Bar Association report calls for substantial service and budgetary authority to be given to local units of approximately 200,000 people, with elected local councils and an elected district executive as the governing bodies.[230] As the governing experience of the local councils increased, they would be granted more powers, eventually including almost complete control over the service personnel working within their respective local units.[231] The proposal, fo-

---

[223] *Id.* at 422.

[224] *Id.* at 434.

[225] *Id.* at 422–32.

[226] The plan proposed the creation of a Department of Community Affairs to coordinate the activities of the councils with those of the central City government, but this department was clearly a creature of the central government and not of the councils. *Id.* at 438–39. The power of the council was further diminished by the existence of an appeals board designed to mediate between the central government and the councils, but which had the authority to punish the councils if it saw fit. *Id.* at 444–45.

[227] *Id.* at 439.

[228] W. FARR, L. LIEBMAN & J. WOOD, DECENTRALIZING CITY GOVERNMENT: A PRACTICAL STUDY OF A RADICAL PROPOSAL FOR NEW YORK CITY (1972).

[229] E. COSTIKYAN & M. LEHMAN, RE-STRUCTURING THE GOVERNMENT OF NEW YORK CITY: REPORT OF THE SCOTT COMMISSION TASK FORCE ON JURISDICTION AND STRUCTURE (1972).

[230] W. FARR, L. LIEBMAN & J. WOOD, *supra* note 228, at 46–47.

[231] *Id.* at 184.

cusing almost exclusively on the local units, suggests little about the relationship of the local units to the central City government[232] and proposes few changes in the operation and organization of the central City government.[233]

The Scott Commission study is broader in scope than, though a good companion to, the Bar Association proposal.[234] The Scott Commission study focuses on City government as a whole and thus spends less time exploring the viability and structure of the local unit than did the Bar Association proposal. The Scott Commission report is presented as a workable, politically feasible plan for the organization of New York City government.[235] While proposing local units not significantly different in size or operation from those suggested in the Bar Association report, the Scott Commission study also explores and proposes a restructuring of the central City government.[236] The proposal is linked to a second plan by the same authors to shift responsibility for some services to a metropolitan regional governing body.[237] Taken together, these proposals are a carefully crafted overall plan for local government reorganization designed to meet the need both for increased community involvement and for rationalization of service management and delivery.

These proposals were just part of the Scott Commission's work. Governor Nelson Rockefeller had charged the Commission with making an extensive study of the City's governmental operations,[238] and in April 1973 the Commission submitted a final report that found fault with the current City management.[239] One of the report's recommendations was that City government

---

[232] *But see id.* at 198–234 (discussing impact of decentralization on provision of sanitation services and on local units' relationship to central government).

[233] *Id.* at 183.

[234] The Scott Commission proposal is also the most similar to the proposal presented in Section III, *infra*.

[235] The authors' first words are, "This is our plan for a new kind of New York City." E. COSTIKYAN & M. LEHMAN, *supra* note 229, at v.

[236] *Id.* at 43–59.

[237] E. COSTIKYAN & M. LEHMAN, NEW STRATEGIES FOR REGIONAL COOPERATION: A MODEL FOR THE TRI-STATE NEW YORK-NEW JERSEY-CONNECTICUT AREA (1973). The authors recommend shifting control of public transportation, water supply, air quality control, water pollution control, and solid waste disposal to one or more regional governments. This Note does not address issues of regionalism.

[238] The Commission was designed in part to embarrass Lindsay who, by this time, was constantly at odds with the Governor. J. NEWFIELD & P. DU BRUL, THE PERMANENT GOVERNMENT: WHO REALLY RULES NEW YORK? 172 (1981); D. YATES, *supra* note 216, at 161.

[239] NEW YORK STATE STUDY COMMISSION FOR NEW YORK CITY, FINAL REPORT OF THE TEMPORARY STATE COMMISSION TO MAKE A STUDY OF THE GOVERNMENTAL OPERATION OF THE CITY OF NEW YORK 9–10 (1973).

should be partially decentralized in a manner very similar to the Costikyan and Lehman proposal to the Commission.[240]

Before the Scott Commission's final report was complete, the state legislature began the task of restructuring City government, as the Scott Commission suggested, and mandated the Governor's creation of a new charter revision commission,[241] which became known as the Goodman Commission.[242] The legislature's findings provided the new Commission with three explicit objectives for restructuring City government: "(i) encourage genuine citizen participation in local city government, (ii) ensure that local city government is responsive to the needs of its citizens, (iii) achieve for cities effective local self-government . . . ."[243]

For advocates of government decentralization, the moment of implementation seemed at hand. Along with this mandate to increase local control came the appointment of Costikyan as vice-chairman of the Goodman Commission. The Commission's preliminary report, released in 1973, demonstrated an apparently deep-seated interest in decentralization.[244] The Commission undertook an exhaustive review of the state of City government, ultimately compiling thirty-two separate reports on various aspects of New York City governance.[245]

When the Goodman Commission presented its final report in 1975, however, much had changed.[246] The preliminary recommendations showed that the focus of the Commission's work had shifted. The report noted ten themes that were covered in the recommendations, but only four of them had any relation to decentralization.[247] Most telling, the section of the recommendations entitled "City Government in the Community" contained a long dissent from Costikyan, who wrote of the Commission's decentralization proposals,

---

[240] *Id.* at 47–48.

[241] 1972 N.Y. Laws ch. 634, *as amended by* 1973 N.Y. Laws ch. 63.

[242] The commission was chaired by Republican State Senator Roy Goodman of Manhattan.

[243] 1972 N.Y. Laws ch. 634, § 1.

[244] NEW YORK STATE CHARTER REVISION COMMISSION FOR NEW YORK CITY, REVISING THE NEW YORK CITY CHARTER, INTRODUCTORY REPORT 14–15 (1973).

[245] *Id.* at 20–25.

[246] D. YATES, *supra* note 216, at 2–3 (discussing the end in the mid-1970's of the period of optimism and innovation in addressing urban problems).

[247] NEW YORK STATE CHARTER REVISION COMMISSION FOR NEW YORK CITY, PRELIMINARY RECOMMENDATIONS vi (1975).

> [T]he proposed Charter attempts to give the appearance of structural reform without the reality. It reflects a desire for change, counterbalanced by a desire not to disturb the status quo . . . . [T]he majority recognizes the ideal but proposes a process for achieving it which precludes its achievement and passes the buck . . . . [T]he result appears to be not dissimilar to what might have been expected if a dinosaur had been asked to design its successor as king of the beasts, and then to plan the transition.[248]

The Goodman Commission's final proposals, presented to New York City voters on November 4, 1975, were indeed disappointing to advocates of decentralization. The Commission recommended six propositions for Charter revision, three of which nominally, but vaguely, increased community control.[249] Three other proposals, two of which would have mandated that certain powers devolve from the central City government to the community, were placed on the ballot without recommendation from the Commission.[250] The voters, predictably, approved those proposals endorsed by the Commission and rejected the others.[251] The approved revisions institutionalized the Community Boards and, on paper, gave the boards more advisory authority over land use decisions, but despite these changes, the central City government retained the same amount of authority

---

[248] *Id.* at 243.

[249] Those three proposals as presented by the Commission were:

> 4. Shall the City's processes for deciding *planning and land use* issues be made more responsive to the needs and desires of citizens in their local communities consistent with the preservation of City-wide interests?
>
> 5. Shall the City adopt a plan of *administrative decentralization* designed to create common boundary service districts ("coterminality") for agencies within local communities, to strengthen the managerial capacity of agencies at borough and community levels, and to enable citizens to affix responsibility for service delivery?
>
> 6. Shall the City increase opportunities for *citizen participation* in government by strengthening appointed community boards and borough boards in the areas of planning, budgeting, and service evaluation?

NEW YORK STATE CHARTER REVISION COMMISSION FOR NEW YORK CITY, FINAL REPORT 2 (1975) (emphasis in original).

[250] The two unrecommended proposals that would have increased local control were:

> 8. Shall the Borough Presidents be assigned responsibility for the design, construction, maintenance, and repair of local streets and sewers?
>
> 9. Shall eleven- to fifteen-member elected community boards replace the present appointed community boards and be given power to deliver services of a local nature, including local parks and playgrounds, local recreation programs, local neighborhood preservation and related housing rehabilitation programs, and neighborhood inspectional services?

Id.

[251] *First 6 Charter Revisions Backed as the Rest Falter*, N.Y. Times, Nov. 5, 1975, at 22, col. 7.

relative to the neighborhoods that it had had prior to these efforts.[252]

D. *The Collapse of the Decentralization Movement*

New York City's financial crisis dashed what hope there might have been to use the Goodman Commission's report as a start toward more decentralization even before the Commission's recommendations went to the voters. In April 1975, New York City ran out of money,[253] and for four years the City fought to stave off bankruptcy.[254] The payroll was cut drastically, and all capital construction and most capital maintenance ceased.[255] Reform was lost in a desperate effort to rescue the City's finances. The City focused on improving the efficiency of the existing government[256] rather than on creating a new structure with unknown reliability and costs.

For approximately ten years, from the mid-1960's through the issuance of the Scott Commission Report in 1973, New York City seemed to be moving toward a partially decentralized government. Some steps were taken in that direction, but all failed to achieve what was expected of the individual programs specifically or of decentralization in general. Almost three decades after the political and philosophical pressure for decentralization and increased citizen participation began to have an impact, and almost two decades after models of a decentralized New York City received serious consideration, the government of the City of New York is as centralized as ever. In fact, the fiscal crisis led to increased centralization.[257] The two surviving pieces of

---

[252] Charo, *supra* note 138, at 753–54.

[253] C. MORRIS, *supra* note 24, at 11.

[254] *Id.* at 232.

[255] R. STARR, THE RISE AND FALL OF NEW YORK CITY 231 (1985).

[256] Two publications that have appeared since the fiscal crisis underline the emphasis on efficiency. The first, appearing annually since 1981, is entitled SETTING MUNICIPAL PRIORITIES, and is edited by two business professors. It provides assessments of the City's economic resources and proposes policy options for improving City service delivery. The second, the MAYOR'S MANAGEMENT REPORT, is published twice annually by the Mayor's Office of Operations, a unit of the Mayor's Office dedicated solely to monitoring the efficiency of City agencies.

[257] R. BAILEY, THE CRISIS REGIME 137 (1984).

the decentralization movement—the local school boards and the Community Boards—have little power and little respect.[258]

Each step toward decentralization failed to reach the expectation held for it for a different reason. The federal efforts at decentralization were victims of the narrowness of their focus and of the uncertainty of their purpose.[259] The Economic Opportunity Act's Community Action Program and the Model Cities Program both suffered from being somewhat program-specific. Local boards and corporations were created, but they could operate only within the ambit of the programs that created them.[260] In addition, the Community Action Program became entangled in racial and class-based political battles[261] that caused Congress to turn its back on citizen participation for fear of creating more radical power bases.[262] The participation element in Model Cities, coming after the Community Action experience, was quite limited in scope.[263] Finally, corruption and mismanagement plagued the programs and undercut their legitimacy.[264] However, these federal programs did create a base of leadership in minority communities that proved to be a foundation for political power.[265]

Neither Model Cities nor the Community Action Program survived to see the fiscal crisis. Both were victims of Nixon Administration cutbacks,[266] though neither had been sufficiently funded to have a substantial impact on the quality of life of most urban dwellers.[267] Furthermore, by the early 1970's support for the 1960's liberal agenda was weakening, particularly in Wash-

---

[258] The Community School Boards have also been victimized by scandal. *See, e.g.*, Ravitch, *Canarsie and Fuentes: The Limits of School Decentralization*, N.Y. AFFAIRS, Summer 1973, at 88; *Chancellor, Citing Indictments, Suspends Bronx School Board*, N.Y. Times, Mar. 28, 1989, at B1, col. 2.

[259] To some extent the federal programs were premised on a belief that simply giving a lot of money for urban problems would solve those problems. This was, it turned out, overly simplistic. D. YATES, *supra* note 216, at 38.

[260] STRUCTURAL ISSUES, *supra* note 191, at 7.

[261] For a description of the battles about and within the Community Action Program, see D. MOYNIHAN, MAXIMUM FEASIBLE MISUNDERSTANDING (1969).

[262] In 1967 Congress removed the "maximum feasible participation" clause from the Economic Opportunity Act and gave central City governments greater control over anti-poverty programs. K. POLLINGER & A. POLLINGER, COMMUNITY ACTION AND THE POOR 7 (1972).

[263] D. YATES & R. YIN, *supra* note 101, at 23.

[264] *See* J. NEWFIELD & P. DU BRUL, THE ABUSE OF POWER 221–30 (1977).

[265] *See* D. MOYNIHAN, *supra* note 261, at 129; D. YATES & R. YIN, *supra* note 101, at 24.

[266] Model Cities was defunded on January 1, 1973. The Community Action Program was phased out later in 1973.

[267] D. YATES & R. YIN, *supra* note 101, at 177–78.

ington, and it was easy for the Nixon Administration to undermine these programs.

The Lindsay Administration's efforts at decentralization also suffered from a lack of breadth and a dearth of direction. The more sweeping proposals that came out of the Mayor's Office were never implemented,[268] and the broadest program put in place, the Office of Neighborhood Government,[269] involved almost no increased citizen participation. While the Office of Neighborhood Government plan as originally conceived contained participation elements, the program as implemented did not.[270] It accomplished some administrative decentralization, but did not create any mechanism for the citizens being served to participate in or to influence service delivery in their neighborhoods.

The Community Boards as constituted both at first[271] and after the 1975 Charter revisions, suffered from being unelected and from lacking anything but advisory power.[272] Despite some expression in the Charter of intent to continue the administrative decentralization of the Office of Neighborhood Government in the Community Boards,[273] City management remains quite centralized.[274]

The failure of the Community School Boards to operate effectively is the most disappointing aspect of the early attempts at decentralization. One of the reasons for this failure is that the Community School Boards were never granted sufficient authority over the schools in their districts.[275] Almost no significant powers were transferred from the Board of Education to the Community School Boards,[276] and the Community School Boards were given no formal representation on the Board of Education.[277] Because the local boards had no ability to control

---

[268] *See supra* notes 183–185 and accompanying text.
[269] *See supra* notes 195–199 and accompanying text.
[270] Barton, *What Has Been Learned from the New York City Neighborhood Government Experiment?*, in A. BARTON, N. FAINSTEIN, S. FAINSTEIN, N. FRIEDMAN, S. HEGINBOTHAM, J. KOBLENTZ, T. ROGERS, J. BOYLE & R. BRUMBACK, DECENTRALIZING CITY GOVERNMENT 249 (1977).
[271] *See supra* notes 177–181 and accompanying text.
[272] *See* N.Y.C. CHARTER §§ 2800–2801 (1985).
[273] *Id.* §§ 2700–2705.
[274] R. BAILEY, *supra* note 257, at 137.
[275] *See* Comment, *supra* note 209.
[276] Gittell, *School Governance*, in C. BRECHER & R. HORTON, *supra* note 23, at 182.
[277] *Id.*

what went on in the schools,[278] the boards failed to attract able people willing to serve as board members. Those who have been willing to serve tend to do so for political reasons, and this politicization of the local boards has further weakened them.[279] The weakness of the boards has also reduced voter turnout in Community School Board elections[280] since parents and others in the community feel their vote is pointless.

The general problem with past attempts to decentralize New York City government is that they have been too limited, either in scope or in the level of authority delegated to the local unit. Those proposals that had sufficient breadth were never allowed to get started or were never allowed to expand to a point where their practicality could truly be tested. The federal programs never met their early promise.[281] The Community Boards lack both power and, because their members are appointed by the Borough Presidents, any representational legitimacy. The Community Boards could have been a focal point for increased decentralization, but they have never been granted sufficient authority to serve that function. Similarly, the Community School Boards have not been sufficiently empowered to give them any legitimacy as governing bodies.[282] The proposals from the early 1970's for substantial government decentralization were never implemented, leaving New York without any past example to test whether true decentralization can work.

The fiscal crisis severely harmed efforts at decentralization, which had already been weakened as the nation became less interested in community power and the war on poverty. Centralization increased as efficiency and cost-containment became paramount goals which many felt only large governmental units could achieve. Further, the fiscal crisis cut into the City's already fragile Home Rule power,[283] decreasing the amount of

---

[278] The power of the Community School Boards was further weakened by the state courts in their interpretations of the decentralization law. *See, e.g.*, Community School Bd. No. 22 v. Board of Educ., 44 A.D.2d 713, 714, 354 N.Y.S.2d 703, 705 (2d Dept. 1974). *See generally* Comment, *supra* note 209, at 259–75.

[279] D. ROGERS & N. CHUNG, 110 LIVINGSTON STREET REVISITED 216–17 (1983).

[280] *Charges of Corruption Spur School Vote Drive*, N.Y. Times, Mar. 5, 1989, at 42, col. 1.

[281] *See supra* notes 259–265 and accompanying text.

[282] There has recently been renewed interest in school decentralization, and there are currently proposals from the new Schools Chancellor to increase local authority. *Fernandez Wants to Empower Local Schools to Make Decisions*, N.Y. Times, Sept. 22, 1989, at B1, col. 2; *New York's New Schools Chief Has a Super-Decentralization Plan*, N.Y. Times, Sept. 24, 1989, § 4, at 6, col. 1.

[283] R. BAILEY, *supra* note 257, at 150.

control the central City government could give to communities over their own governance. The post-fiscal crisis structure of City government, centralized and increasingly controlled by outside forces, should not be the death-knell for decentralization, however. Instead, it should highlight the need for increasing participation in the workings of local government.

First, the legitimacy of a distant, centralized government rests inevitably on some notion of the government acting in the public interest.[284] In a diverse, ever-changing city, however, the notion of a single, definable public interest seems absurd, and the legitimacy of a government so dependent on homogeneity of interest is tenuous.

Second, centralization carries with it the inevitable possibility of tyranny by the governing faction.[285] It has been suggested that New York City's private power elites created, or at least encouraged, the City's fiscal crisis in an effort to wrest control of the government from the more populist, public spending-oriented power groups of the late 1960's and early 1970's.[286] Regardless of the truth of this conspiratorial theory, the centralization and disempowerment of New York City government between 1974 and 1981 create concerns. First, there is a simple Madisonian fear of faction. While decentralization is often attacked for creating a myriad of potential Madisonian nightmares in each new local unit,[287] the centralized government—run by a small cadre huddled in City Hall, listening almost exclusively to political contributors and the bankers essential to the City's survival—is more dangerous to democracy than any neighborhood council representing only a fraction of the metropolis.

The second concern about the centralization of City government relates back to legitimacy. The social fabric of New York City, often weak, has been shredded in recent years and is dangerously thin.[288] New York, along with other cities, has been described as being in a "slow motion riot"[289] waiting to explode. A government that is distant from those it serves and that is seemingly unresponsive and uncaring only exacerbates social tensions. John Lindsay and his administration firmly believed

---

[284] *Id.* at 190.
[285] *Id.*
[286] *See* E. LICHTEN, CLASS, POWER & AUSTERITY (1986).
[287] Ravitch, *supra* note 57, at 110.
[288] *See generally* R. BAILEY, *supra* note 257, at 188.
[289] Gardells, *The Fire Next Time: Rich vs. Poor*, L.A. Times, Jan. 29, 1987, pt. 2, at 5, col. 4.

that the City government could help change people's lives,[290] and that the way to be most effective in this task was to bring a competent City government to the people and let them be a part of it.[291]

The Lindsay Administration's vision remains a powerful one, particularly for running a fragmented city. The idea is to let the people of the city have a chance to control their own governance and to work together to solve their problems, instead of forcing them to rely on a distant central government. It will take time for citizens to learn to be a part of their own government, and it will require education. But if there is to be a free government for a heterogeneous population, it must be participatory, providing opportunities for all citizens to have real input into governing their communities.

III. A Model of a Decentralized New York City Government

Since the creation of the City of New York almost one hundred years ago, the process of adjusting and restructuring its government has been driven largely by partisan and parochial political considerations.[292] Inevitably, few significant changes are made in the governmental structure except in reaction to scandals.[293] The use of proportional representation was the one radical change made, and it survived the furor of the entrenched power structure for only a decade.[294] The recently completed revision process fits this traditional pattern. The Revision Commission was created in response to a scandal and a crisis—the corruption scandals that broke in 1986, and the federal court rulings ordering the abolition of the existing Board of Estimate structure.[295] While the Commission's research and examination of governmental structures were extremely thorough, the focus of its recommendations was decidedly narrow, steering clear of potentially controversial restructurings.[296] There appears to

[290] J. Lindsay, The City 11–19 (1970).
[291] *Id.* at 116–17.
[292] *See* Section II, *supra*.
[293] The abolition of the Board of Aldermen was such a response to scandal. *See supra* note 167 and accompanying text.
[294] *See supra* notes 168–174 and accompanying text.
[295] *See supra* notes 2–7 and accompanying text.
[296] The restructuring that was accomplished was unavoidable after Board of Estimate v. Morris, 109 S. Ct. 1433 (1989).

have been no serious consideration of the creation of empowered local units or even of the shift of responsibility for some service delivery to more local authorities. Absent these steps, any talk of increased community participation is futile.

The process of creating an effective local unit structure in New York City involves two major changes in current governmental arrangements. The first, to be addressed below,[297] is the creation of a local unit structure within the City, one that provides authority and opportunities for participation sufficient to make the local units viable and useful. The second step, not within the scope of this Note but of substantial importance, is a change in intergovernmental relations that reduces the powerlessness of cities.[298] If cities are not granted more authority over their own destinies, they will have no ability to devolve meaningful authority to local units.[299] Just as cities are now frustrated by their lack of authority, local units can become frustrated if they find that they have little control over the destiny of the areas in their charge.[300] There may well be a symbiotic relationship between effective decentralized government and greater authority for cities: an effective and popular local unit government structure may encourage state and federal officials to loosen their grip on urban policy, allowing the local units greater opportunities to improve the lives of their citizens. No matter the relationship, there is little doubt that for local governments to be more effective, those governments must have more control over the services and policies that are the core of local government.

A. *The Local Unit*

Even within the current intergovernmental structure, it is possible, using the Bar Association and Scott Commission re-

---

[297] *See infra* notes 301–307 and accompanying text.

[298] *See* Bresnick, *The Other Side of Decentralization: Home Rule For New York City*, 64 NAT'L CIVIC REV. 71 (1975).

[299] Frug argues implicitly that simply tampering with existing governmental arrangements, such as Home Rule, will not be enough to give power back to the cities. Frug, *supra* note 88, at 553–54. One way for cities to increase their power is to expand the use of their authority as vast property owners, and indeed, to increase their property ownership. Frug, *The City as a Legal Concept*, *supra* note 22, at 687–91.

[300] The debate about how best to empower localities so as to increase self-governing while retaining sufficient oversight to control parochialism and corruption is a difficult and complex one which cannot be covered within the scope of this Note. For some background on this debate, see G. FRUG, *supra* note 36, at 87–126.

ports as examples, to design a local unit system that will give meaningful opportunities for increased participation and greater government contact with the governed. This proposal has the advantage of two decades of experience with decentralization and of the lessons learned from the retreat from decentralization during the 1970's. The base unit for a decentralized New York City government should be a local area of between 70,000 and 100,000 in population, meaning that there would be between 70 and 100 of these units.[301] The Scott Commission proposed larger units.[302] However, units over 100,000 tend to cover geographical areas so large as to make access to a central location difficult for many in the unit and to decrease any geographical sense of community.

The unit boundaries should initially be drawn by a special city commission consisting of the City Planning Commission, the Borough Presidents, and members of at least two community corporations or their equivalents from each borough. The Planning Commission has the best knowledge of City demographics and land use, the Borough Presidents have the best sense of the political arrangements within their localities, and the community corporations have the best knowledge of the existing community political and activist structure. Final approval of the districting should rest with the City Council and the Mayor in order to provide city-wide political accountability to the process. Boundaries should be reviewed by a similar process after every national census to ensure continued population balance.[303] To the extent possible, units should not cross borough boundaries. Each unit should also be a City Council district and should define a service district for those services that can be decentralized to the local unit level.

A local unit council should govern each local unit. The council should be elected by the citizens of the local unit under a scheme of proportional representation.[304] There should be one council

---

[301] New York City's population is approximately seven million. For discussion of unit size, see *supra* notes 91–93 and accompanying text.

[302] *See supra* notes 229–236 and accompanying text.

[303] Such a population balance is mandated by Avery v. Midland County, 390 U.S. 474 (1968), and Board of Estimate v. Morris, 109 S. Ct. 1433 (1989).

[304] The use of proportional representation is intended to avoid one-party domination of local councils where that one party does not command such a majority in the local unit. Past experience in New York indicates that proportional representation can avoid such domination and in fact results in representation close to that of the party preferences of the voters. *See supra* notes 168–174 and accompanying text. For a powerful argument in favor of proportional representation as a way to increase minority group

member elected for every 3000 valid ballots cast, creating a council of between ten and twenty members in each district, depending on voter turnout. This will create a body that is very representative but that at the same time is not so large as to be unwieldy. The council should elect its chair from amongst its own members. The citizens of the local unit should also elect a single official who will at once act as the unit executive and as the City Council representative for the local unit. This individual should be elected by a system of preferential voting.[305] The position of unit executive/City Council member should be full-time[306] and should be compensated accordingly. A local unit council position should not be considered a full-time job, and members of the local unit council should be paid at a rate that represents the portion of their time spent on council business. This rate should be determined after some experience with council operations. Until that time, the members should be paid on a—rather unsatisfactory—hourly basis.[307]

B. *The Powers of the Local Unit*

The local unit councils should have substantial political and administrative authority. The most important of these powers is the ability to influence service delivery within the local unit. To the extent feasible, city services, particularly uniformed agencies and health and inspectional services, should be divided into

---

representation, see Note, *The Constitutional Imperative of Proportional Representation*, 94 YALE L.J. 163 (1984).

[305] Preferential voting is, in essence, proportional representation for single-office elections. As in proportional representation, *see supra* note 170, the voters indicate their first, second, third, and so on preferences among the slate of candidates. When counting the ballots, if a candidate receives a majority of first choices (not a plurality) he or she is elected. If not, all the second choices are then counted, and so on, until a victor is found. The system allows for the elimination of primaries and of general elections between two candidates who may both be unacceptable to a substantial plurality of voters. *See* C. KNEIER, *supra* note 31, at 365–68. It should also be noted that the use of computerized voting booths would eliminate the counting complexities which often were a problem with both proportional representation and preferential voting when tallying was done by hand.

[306] Persons holding this position should not be allowed to hold any other paid position, public or private.

[307] All local unit and City Council elections should be publicly funded.

service districts coterminal with the local units.[308] The district managers of the decentralized services should have significant operational discretion[309] and should report to the local unit council. The local unit executive should preside at monthly meetings of the district managers of all decentralized services to increase interservice cooperation.[310] The local unit council should hold frequent meetings to allow citizens to express their views about service problems to the service managers. When those problems are serious, the local unit councils could mandate that the service manager take specific appropriate action, such as cleaning up a particular park or patrolling a certain block more frequently.[311]

The local council should have two substantial policy powers. First, the council should recommend to the borough council[312] capital improvements necessary within the local units. These projects should range from the specific (which streets need repaving, in order of priority) to the more general (the unit needs a new school). Second, the local unit council should have the first power of review on any zoning or other land use change in its district now requiring use of the Uniform Land Use Review

---

[308] In 1975 the Goodman Commission found that a number of city services, including the uniformed services, could be divided into coterminal districts—66 was the number of districts used in the study—without significant difficulty. For some services such extensive decentralization does not make sense. Even for those services, however, some decentralization would be possible, perhaps to the borough level. *See* McKINSEY & COMPANY, INC., *supra* note 101, at 3-13 to 3-17, 3-21 to 3-22.

[309] For example, local sanitation managers would be given discretion over pickup schedules and manpower deployment. If a particular unit were willing to have only thrice weekly garbage pickups in exchange for more regular street sweeping, the local manager would have the authority to make that change. Similarly, police precinct captains would have more power over deployment within the local unit, as well as authority to use existing community structures to create greater crime-stopping cooperation with local citizens. Some mild versions of these techniques are already being tried in some cities. *See, e.g.*, *Cities Try Out New Approach In Police Work*, N.Y. Times, Mar. 29, 1989, at A14, col. 1.

[310] For example, police, fire marshals, and building inspectors could work together to drive drug dealers out of vacant or partially occupied buildings within the local units. Or highways, sewer, and water managers could coordinate repairs so as to minimize the number of times a street must be torn up.

[311] Such manipulation of existing service delivery arrangements will require the cooperation of the unions, a lesson learned in the process of school decentralization when the teachers' union objected to the attempt by the Ocean Hill-Brownsville school district to make personnel decisions. *See supra* notes 215–218 and accompanying text. Existing union contracts that include strict work rules which would prohibit substantial local unit impact would have to be renegotiated by the central City government to avoid confrontations between the unions and the local units. This is not impossible if approached appropriately. *See infra* note 328 and accompanying text.

[312] *See infra* notes 321–322 and accompanying text.

Procedure ("ULURP").[313] Disapproval of a project would require a two-thirds vote of the local unit council. Reinstatement of a project after local unit council rejection would require a two-thirds vote of the City Planning Commission sitting at an open meeting held within the affected local unit. The open meeting would at once allow the community clearly to express its opposition and act as a check on the sort of backroom deals that are easier to accomplish when the public is not present.

The granting of land use powers to the local unit councils is probably the most controversial aspect of this model. Granting local units true land use veto authority only encourages the not-in-my-backyard ("NIMBY") syndrome manifesting itself in neighborhood rejection of any and all development proposals.[314] However, local governments have their greatest power in the control of the use of land within their geographic borders;[315] decisions about land use controls often most interest and affect the citizens of a given local area. A local unit government without authority to affect land use decisions within its domain is without one of the basic elements of true local government authority.

Further, the authority over land use granted in this model is not an absolute veto. The local unit council is given, in essence, an enhanced advisory role. The council on its own cannot stop any given development. The City Planning Commission still has the power to override a negative recommendation from the council. The power of the local unit council is in forcing the Planning Commission to reject formally the recommendation of the local unit council, to make that decision in an open meeting in the affected local unit, and to do so by a super-majority vote. The central City government is forced to justify its actions, and to provide amenities, services, or facilities in return for approval of a specific use. There are critics who argue that the receipt of

---

[313] Most major building and improvement projects, and all projects not as-of-right are now subject to ULURP under the auspices of the City Planning Commission. ULURP includes the preparation and submission of extensive technical and environmental studies for approval by the Commission. Based on those studies, the Planning Commission votes on whether to recommend passage of the zoning approval by the Board of Estimate. N.Y.C. CHARTER § 197c. The effect of this local unit power would be to subject most developments to local unit council review.

[314] There is substantial experience with NIMBY, particularly where waste facilities are involved. *See, e.g.*, Davis, *Approaches to the Regulation of Hazardous Waste*, 18 ENVTL. L. 505, 527 (1988).

[315] Frug, *Property and Power: Hartog on the Legal History of New York City*, *supra* note 22, at 687.

amenities by neighborhoods in exchange for their approval is simply blackmail or "horse trading." They claim that in order to place an essential but potentially unpleasant facility such as a jail, detox center, or sanitation transfer station in a particular location, the City is forced to bargain with the community to avoid endless interference from local activists.[316] Under this model, however, if two-thirds of the City Planning Commission believe that the development should go forward despite local unit opposition and without any additional amenities or City services, and the Planning Commission is willing to express this opinion in front of the citizens of the local unit, then the Commission will not need to "horse trade."

The local unit councils should also be involved in the operation of public schools located within the area. The current community school districts should be made coterminous with the local units. The local school board should be separate from, but linked to, the local unit council. The local school board should have eight members, four of whom should be elected members of the local unit council chosen by the full council to sit on the school board. Local unit voters and eligible parent voters[317] should elect the remaining four members by proportional representation[318] at an election held at the same time as the local unit council elections. The local unit executive/City Council member should sit as chair of the local school board and be entitled to vote only in the occasion of a tie. If there is such a tie, the chair may refuse to cast a vote and may instead refer the issue to the local unit council for a final decision. At a vote on school issues, the local unit council and the local school board should sit as a committee of the whole with the non-council school board members granted full parliamentary and voting rights. Even if there is no occasion for such a joint meeting, the local unit council should, at least semi-annually, conduct hearings and, if necessary, investigations concerning the operations and activities of the coterminous local school board.[319]

---

[316] *A Vast City Needs a Strong Mayor*, N.Y. Times, Apr. 18, 1989, at A26, col. 1.

[317] An eligible parent voter is a parent who is not a resident in that local unit but who has a child in a school governed by that local unit. N.Y. EDUC. LAW § 2590-c (McKinney 1981).

[318] The Community School Boards are currently selected by proportional representation. *Id.*

[319] The local school boards should also be granted more authority, including control over the general purpose high schools within the local units, and the local boards should

C. *Borough Government*

The creation of the local units should not result in the elimination of the boroughs within the governmental structure. In fact, borough government should be enlarged to include a borough council to work with the Borough Presidents. The Borough Presidents should be elected by preferential voting,[320] and the borough council should consist of all City Council members from that borough. The borough council should give advice and consent to the Borough President on all borough appointees to central City government agencies and public authorities. The borough council should also, in conjunction with the Borough Presidents, compile the list of capital improvement needs for the borough derived from the various local units. The council and the Borough President should prioritize the capital needs of the borough from the local unit lists. If the Borough President and the council disagree on these needs, the council should determine the priorities subject to a veto by the Borough President. The veto can be overcome by a two-thirds vote of the council. This prioritized list should be submitted to the Mayor and the City Council for use in preparing the City budget.

The borough council should have some authority to affect land use within its borough. After the City Planning Commission has overturned a negative recommendation by a local unit council of a proposed development or land use change, the borough council should have the power to delay implementation of the Planning Commission's decision for sixty days while the full City Council reviews the proposal. The borough council should be permitted to implement this procedure only by petition of the affected local council and within twenty days of the Planning Commission's decision. A unanimous vote of the borough council should be necessary to trigger City Council review. A simple majority of the City Council, however, should sustain the Planning Commission.[321]

---

be allowed to elect at least half the membership on the central Board of Education in order to ensure central Board respect for the decentralized local boards.

[320] *See supra* note 305.

[321] This scheme creates a final appeals process for the local unit councils. While this process may slow land use approval somewhat, it acts as a final check on possible Planning Commission tyranny. The requirement of unanimity of the borough council will result in the infrequent use of this process, since such unanimity will probably only be possible when egregious disagreement exists between the borough council and the Planning Commission.

The borough council will oversee those services which could not be decentralized to the local unit level, such as the Health and Hospitals Corporation. The borough council will work with these services in a manner identical to that used by the local unit councils in dealing with the services decentralized to the local unit level.[322]

Finally, the borough councils should act as arbitrators among the local unit councils within that borough. Disagreements regarding service delivery, land use on unit borders, or competing policy goals should be brought to the borough council. The borough council and the Borough President should work to find an informal solution to any dispute. Failing that, the borough council should sit and hear the competing views of the affected local councils and then should render a decision by majority vote.

The borough councils should meet at least monthly and should, like the local unit councils, sit at a time accessible to as many citizens as possible.

D. *Central City Government*

The central City government should consist of a popularly elected mayor with strong managerial powers, and a City Council. The City Council should consist solely of the local unit executive/City Council members elected in each local unit. The use of these officials as Council representatives will create a fairly large deliberative body of seventy to one hundred members, depending on the number of local units.[323] The Board of Estimate should be abolished, as should the elected positions of Comptroller and City Council President.

The Mayor should be the chief executive officer of the City and should appoint all agency heads, with the advice and consent of the City Council. The Mayor should also appoint at least half the members of all City commissions. However, at least one position on each commission should be available for each Borough President in order to assure both that political views

[322] For a description of that relationship, see *supra* notes 308–311 and accompanying text.

[323] *See supra* note 301 and accompanying text. The new Charter increases the size of the Council to 51 from 35, reducing the size of Council districts by about 30%. FINAL PROPOSALS, *supra* note 7, at 9. These new Council districts will not be coordinated with the Community Boards or any other existing or planned local unit.

other than those of the Mayor are heard and that the whole City is represented. The Mayor should prepare the expense and capital budgets for review by the City Council. Finally, the Mayor should coordinate all City activities with those of regional public authorities, of the State of New York, and of the Federal government.

The City Council should have the authority to pass all city legislation, including taxation changes, subject to mayoral veto.[324] The Council should review the Mayor's budgets and be free to add or subtract items at its discretion, within the limits of the City's balanced budget requirements.[325] The Council should appoint a City Comptroller who should assist the Council in budget review. The Comptroller should also report to the Council frequently on the state of the City's finances and on the operation of its agencies and facilities. The appointment, as opposed to the election, of the Comptroller should reduce the politicization of that office, resulting in the selection of professional auditors rather than politicians. The Comptroller should be allowed to audit the local unit councils and the borough councils. On all issues, including the budget, the Council should have the authority to refer items to the local unit councils for an advisory vote from those bodies. The Council should be given investigatory powers separate from those of the Comptroller, and should be given sufficient staff and resources to accomplish thorough investigations. Finally, all residual powers of the City should rest with the City Council.

Final land use powers should rest with the City Planning Commission, subject to the restraints from the local unit and borough councils.[326] The Commission should consist of fifteen

---

[324] The Mayor should have veto power over all legislation passed by the City Council. A two-thirds vote of the Council will override any veto.

[325] The taxing and budget powers are left with the central City government to avoid inequalities between districts with big tax bases and those with minimal tax bases. This problem has been illustrated in local school funding by the great educational inequalities that have often existed between rich and poor school districts. *See* San Antonio v. Rodriguez, 411 U.S. 1 (1973) (upholding a Texas school funding system that allowed great wealth differentials between school districts). *But see* Serrano v. Priest, 18 Cal. 3d 728, 557 P.2d 929, 135 Cal. Rptr. 345 (1976) (holding that a California funding scheme that allowed inter-district funding differentials violates the state constitution). Some solutions to the problem of financing local unit government at the unit level have been suggested, but the proposals are either extremely complex or unsatisfactory. *See, e.g.*, Church, *A Mechanism for Financing Decentralized Government*, 4 URB. LAW. 557 (1972) (proposing a property tax base equalization levy that would require calculating per capita tax base advantage or disadvantage for each decentralized unit).

[326] *See supra* notes 313–316, 321 and accompanying text.

members, five appointed by the Mayor, five by the City Council, and one by each Borough President. The City Council will only be able to vote on land use issues when a Planning Commission vote has been rejected by a borough council[327] or when a change in the City map is required. All map changes must be approved by the Council.

All City contracting and franchising powers should rest with the Mayor subject to the approval of a commission consisting of the Comptroller, five members appointed by the Mayor, and one member appointed by each Borough President. This commission should act as a check on potential corruption in the contracting process, while also providing a broad-based opportunity for review of the wisdom of the various contracts. The Comptroller should prepare an annual report of all city contracts and franchises approved in the previous year and should audit those contracts and franchises as necessary. The Mayor should negotiate union contracts in conjunction with an advisory Board consisting of the Comptroller and the Borough Presidents. The advisory board should work closely with the local unit governments to develop contracts that increase the flexibility of the work rules to allow sufficient personnel discretion for the local units without undermining the rights of the employees.[328]

E. *Conclusion*

This proposal is not a radical model of town-meeting government in the metropolis. Instead, it is a model of government moved closer to those governed within the practical restraints of operating a City of seven million people with a budget of over $25 billion. The local units have substantial impact, however, even beyond their geographic area of authority. The election of

---

[327] *See supra* note 321 and accompanying text.

[328] Such agreements between unions and governmental agencies designed to improve flexibility and operational control are not unprecedented. In 1985 the New York City Transit Authority won agreement from its unions to change 500–800 positions from union to non-union classification in order to improve supervision and repair depots. *More Transit Supervisors To Join Management Ranks*, N.Y. Times, Jan. 31, 1985, at B1, col. 3.

City Council members from those units and the responsibility of each Council member as a local unit executive will make the City Council more responsive to the interests of the decentralized units. It will also give the local unit councils a voice beyond the already important power they will have over local service delivery and land use.

NOTE

FEDERAL HIGHWAYS AND ENVIRONMENTAL LITIGATION: TOWARD A THEORY OF PUBLIC CHOICE AND ADMINISTRATIVE REACTION

ROGER NOBER*

The construction of interstate highways necessarily disrupts the environment. Formerly, environmental activists and other opponents of the construction of highways, such as residents who are dislocated by the construction of the project, have had little recourse to influence the construction. As a result of society's increased environmental consciousness, two federal statutes were enacted in the 1960's: the Department of Transportaion Act of 1966 and the National Environmental Policy Act of 1969 (NEPA). Section 4(f) of the Department of Transportaion Act prohibits the use of publicly owned parkland if a "prudent and feasible alternative to using that land" exists. NEPA sets out strict procedures which highway planners must follow in order to mitigate any impact a highway may have on the environment. With the aid of these two statutes, opponents of highway projects have been able to cause significant delays and sometimes even the termination of construction plans, despite the fact that the projects often had the support of the general public.

In this Note, Mr. Nober analyzes these two statutes and examines their effects on the public decision-making process. He concludes that the courts are not the proper forum for making such choices. Mr. Nober would prefer that special interest groups did not have what in effect is a veto power over plans for highway construction that the community accepts or welcomes. As an alternative to litigation, he proposes that these statutes be amended to encourage concerned parties to use negotiation and/or mediation before litigation. The result, Mr. Nober argues, would be less fighting over procedural and administrative issues. Rather, the decision-making process would be informed by increased public debate on the merits of a project, thus better reflecting the needs of the whole community.

In the past, large-scale transportation projects usually proceeded rapidly from initial idea to final form.[1] Regional residents

* Law Clerk to the Hon. David N. Edelstein, United States District Court for the Southern District of New York; B.A., Haverford College, 1986; J.D., Harvard Law School, 1989. The author would like to thank Professors Richard Stewart and Lance Liebman for their comments on this Note. In addition, the author extends his appreciation to Ed Kussy and Harold Aikens of the Federal Highway Administration; Jane Catler, Harvard Law School, Class of 1990; and his family for their editorial aid and support.

[1] For the purposes of this Note, transportation will be a generic term referring to the panoply of modes of movement considered under the aegis of the United States Department of Transportation. These include automobile, mass transit (including subways [heavy rail], trolleys [light rail], commuter rail, busses and van pooling), rail, and air modes. The Note will focus primarily on the construction of new roads, notably highways constructed under the Federal Aid Highway Program.

frequently encouraged, supported, or at least tolerated the construction of these new expressways, subways, busroutes and airports, viewing them as progressive additions that would ease congestion. But today, certain improvement plans, especially those in crowded urban areas, ignite storms of protest from local residents who fear the construction's ill effects. Large-scale projects such as interstate highways dislocate large numbers of people, disrupt communities, and thus tend to encounter local opposition.[2] Nevertheless, such projects traditionally have moved ahead as quickly as the less controversial ones, because opponents have had little recourse to stop the construction.[3]

The fortunes of transportation project opponents changed dramatically with the enactment of two federal statutes in the second half of the 1960's. In 1966, Congress passed the Department of Transportation Act,[4] creating the cabinet-level Department of Transportation. Section 4(f) of this Act provides that no highway project can use publicly owned parkland if a feasible or prudent alternative route exists.[5] In 1969, Congress enacted the National Environmental Policy Act (NEPA)[6] which forces highway planners to heed environmental concerns by ordering them to follow strict procedures in the development of construction projects.[7] Together, these statutes require the Department of Transportation (through the pertinent agency) to study environmental impacts and follow proper procedure before approving a transportation project that may expend parkland or substantively impact the environment.[8] With the enactment of the

---

[2] While older urban and rural roads were placed with relatively little disruption or built before accompanying development, large capacity or limited-access highways such as interstate highways are relatively new phenomena. The Interstate Highway program was conceived in 1944, and funded in 1956 under the National System of Interstate and Defense Highways. *See* D. ST. CLAIR, THE MOTORIZATION OF AMERICAN CITIES 149, 160–61 (1986). For a discussion of the interstate highway system, see *infra* notes 49–67 and accompanying text.

[3] "Highway construction is not exactly a democratic process" was Louisiana highway board member Edward Lennox's admonition to recourseless highway opponents, expressing the predominate attitude of planners at that time. R. BAUMBACH & W. BORAH, THE SECOND BATTLE OF NEW ORLEANS 117 (1981).

[4] Department of Transportation Act, Pub. L. No. 89-670, 80 Stat. 931 (1966) (current version at 49 U.S.C. § 101 (1982 and Supp. 1987)).

[5] For the relevant text of section 4(f), see *infra* note 93.

[6] National Environmental Policy Act of 1969, Pub. L. No. 91-190, 83 Stat. 852 (1970) (current version at 42 U.S.C. § 4321 (1982)).

[7] The Act provides that a report detailing the environmental and social impacts of the proposed project be filed for all "major Federal actions significantly affecting the quality of the human environment . . . ." 42 U.S.C. § 4332(2)(C) (1982).

[8] *See* 42 U.S.C. §§ 4321–4347 (1982). For regulations guiding the implementation of NEPA, see 40 C.F.R. §§ 1500–1508 (1988). For the full text of section 4(f) of the

two statutes,[9] opponents of transportation projects gained the ability to stop construction by filing suit, claiming the agency incorrectly assessed environmental harms and/or failed to follow proper procedure in developing the projects.[10]

The NEPA and section 4(f) provisions apply to all actions by the Department of Transportation, but in practice their effects are felt most acutely in the construction of federally-funded highway projects.[11] Federal highways, particularly interstate expressways, often traverse heavily populated or environmentally sensitive areas as planners seek to maximize an expressway's convenience for users. Much of the interstate system was completed or was well under construction when the first environmental statutes were passed, but many of the most complex and controversial segments of the system—necessarily put off until the end of each project—were only in the planning stages when section 4(f) and NEPA were enacted. These controversial seg-

---

Department of Transportation Act, see 49 U.S.C. § 303 (1982 and Supp. 1987). For regulations promulgated under section 4(f), see 23 C.F.R. § 771.135 (1988). For a discussion of the legislative and statutory history behind NEPA, see *infra* notes 68–89 and accompanying text; and for section 4(f), see *infra* notes 90–110 and accompanying text.

[9] To be sure, current legal challenges to transportation improvements may be made under a host of different statutes. But the vast majority of transportation cases involve the Federal Highway Administration (FHWA), and 39 of the 40 cases involving FHWA decided between 1985 and 1988 included either a NEPA or section 4(f) claim, or both. UNITED STATES DEPARTMENT OF TRANSPORTATION, FEDERAL HIGHWAY ADMINISTRATION, AN ANALYSIS: THE EFFECTS OF ENVIRONMENTAL LAWSUITS ON PROGRAM OPERATIONS, 8–9 (1988) [hereinafter cited as FHWA ENVIRONMENTAL LAWSUIT STUDY].

[10] Reviewing courts have interpreted these environmental statutes to confer broad grants of standing that enable anyone affected by the construction to sue. Standing under NEPA is guided by the same general doctrine dictating standing parameters under other federal statutes. Few challenges to standing are successful. *See* F. ANDERSON, NEPA IN THE COURTS 26 (1973). Currently, parties without an economic interest in the project, but with a demonstrated public concern, may be granted standing. *See* Like, *Foreword* to M. SIVE, ENVIRONMENTAL LEGISLATION at vii (1976).

[11] There are 3,879,538 total miles of roads in the United States, of which 847,268 miles, or about 21.8%, were funded with federal money through the Federal-Aid Highway Program. UNITED STATES DEPARTMENT OF TRANSPORTATION, FEDERAL HIGHWAY ADMINISTRATION, HIGHWAY STATISTICS 1986 116 (Table HM-16) (1987) [hereinafter cited as FEDERAL HIGHWAY STATISTICS 1986]. *See also* Kussy, *Wetland and Floodplain Protection and the Federal-Aid Highway Program* 13 ENVTL. L. 164 n.5 (1982) (giving slightly different figures).

It is important to note that these statutes only apply to highway projects constructed at least in part with federal money. Projects paid for solely with state money are subject to review only under the environmental protection statute (SEPA) of that state. For a discussion of SEPA statutes, see M. SIVE, *supra* note 10, at 94–107; *see also* 2 F. GRAD, TREATISE ON ENVIRONMENTAL LAW § 9.08 (1987). SEPA statutes, which may provide a degree of protection greater than, less than, or equal to NEPA, are outside the scope of this Note.

ments gave rise to an opposition that could and did take advantage of the new environmental legislation.

In constructing any major highway project, the community must exercise a public choice and apportion benefits and burdens among local citizens.[12] A highway enriches a large segment of an area's population by reducing congestion, redistributing traffic from local streets, increasing automotive efficiency, and lowering transit costs such as commuting time.[13] In addition to enhancing local transit efficiency, highway construction and usage also benefit the large automotive and trucking segments of the economy.[14] The onus of the project, however, falls on the relative few who live in the right-of-way. This minority, while compensated for the quantifiable costs, must endure incalculable externalities such as dislocation, hardship, and the loss of neighborhood and community bonds so that society as a whole may benefit.[15]

Before the passage of NEPA and section 4(f), the process of weighing the costs and benefits of constructing a highway was not a matter of public debate.[16] The philosophy that saw urban renewal as progress made unfettered highway construction a social imperative. So except in rare instances, public choice and

---

[12] The costs and benefits of a proposed project are to be laid out in the NEPA-mandated environmental impact statement. The environmental impact statement for a highway project chronicles the socio-economic benefits of the new road. *See infra* notes 70–71 and accompanying text. Usually an urban highway not only saves driver time by reducing congestion, but also is safer, due to fewer accidents, and improves local air quality by allowing cars to operate more efficiently. For an example of such analysis, see FEDERAL HIGHWAY ADMINISTRATION & PENNSYLVANIA DEPARTMENT OF TRANSPORTATION, MID-COUNTY EXPRESSWAY: FINAL SUPPLEMENTAL ENVIRONMENTAL IMPACT STATEMENT/SECTION 4(F) EVALUATION (1984) [hereinafter BLUE ROUTE EIS]. In addition, a highway may spur new development. *See, e.g.*, UNITED STATES DEPARTMENT OF TRANSPORTATION AND UNITED STATES DEPARTMENT OF HOUSING AND URBAN DEVELOPMENT, THE LAND USE AND URBAN DEVELOPMENT IMPACTS OF BELTWAYS: CASE STUDIES (1980) [hereinafter cited as BELTWAY STUDY].

[13] For one method of measuring the benefits of highway projects, see generally H. MOHRING & M. HARWITZ, HIGHWAY BENEFITS: AN ANALYTICAL FRAMEWORK (1962).

[14] The size and voracity of the highway lobby cannot be underestimated. As many as one out of six Americans in 1978 were employed either in highway construction and maintenance or in automotive production and maintenance. *See* THE END OF THE ROAD: A CITIZEN'S GUIDE TO TRANSPORTATION PROBLEMSOLVING 36 (R. Golten, O. Houck, & R. Munson eds. 1977) [hereinafter THE END OF THE ROAD].

[15] *See* A. LUPO, RITES OF WAY (1971) (describing how community opposition from fear of dislocation forced Massachusetts Governor Francis W. Sargent to cancel proposed inner-belt and southwest expressways); R. CARO, THE POWER BROKER 850–94 (1974) (describing community suffering as a result of Cross-Bronx Expressway construction).

[16] *See* A. LUPO, *supra* note 15; R. CARO, *supra* note 15; H. LEAVITT, SUPERHIGHWAY-SUPERHOAX (1970).

apportioning burdens lay in the hands of planners alone.[17] Opponents' only means of stopping a project was to attempt to sway public opinion and ultimately induce politicians to cancel the project.[18]

The advent of NEPA and section 4(f) empowered previously impotent highway opponents by giving them the opportunity to bring federal suits against highway planners. As a result of the Supreme Court holding that section 4(f) must be broadly followed,[19] opponents have filed suit on procedural and substantive grounds and delayed projects for years. While the statutes do not allow a court to rule on the actual merits of a highway, opponents can litigate and cause delay; they then use the delay to generate publicity about the plans[20] and perhaps erode political support for the highway.[21] The current laws essentially give highway opponents a "neighborhood veto" over projects, through the power to delay a project by miring it in litigation.[22] As a result, the political system rarely has a chance to pass judgment on the merits of a controversial highway project.

It is the view of this Note that completing new sections of interstate highway under NEPA and section 4(f) should require extraordinary planning and accommodation by highway plan-

---

[17] A poignant example of the futility of local citizens fighting a political establishment determined to build a highway over local opposition was the construction of the Cross-Bronx Expressway through the Bronx in 1952–55. There, in order to connect the George Washington Bridge with the New England Thruway, Robert Moses displaced thousands of Bronx residents and, according to Moses biographer Robert Caro, accelerated the deterioration of the South Bronx. There was a consensus on the need for the highway, but factions argued over its path. *See* R. CARO, *supra* note 15.

[18] Two examples of grassroots victories come from Boston and New Orleans. In Boston, intense local opposition to the construction of an inner beltway through Cambridge, Somerville, and Boston to link with a radial expressway southwest persuaded Governor Sargent to halt all highway construction within the Boston beltway (Route 128, now Interstate 95). *See generally* A. LUPO, *supra* note 15. In New Orleans, citizens revolted against plans to construct an eight-lane raised expressway along the Mississippi River through the periphery of the historic Vieux-Carre (French Quarter). *See generally* R. BAUMBACH & W. BORAH, *supra* note 3.

[19] *See* Citizens to Preserve Overton Park v. Volpe, 401 U.S. 402 (1971).

[20] Generating publicity is a key element of highway opponents' strategy. For a primer on using the media for such ends, see THE END OF THE ROAD, *supra* note 14, at 108–12.

[21] Politicians often support urban highway projects, complicating the task of stopping the project. Political support, however, may be premised on the desire to see large amounts of federal funds be expended in the locality, which, given the multiplier effect, may result in a significant jolt to the local economy. Highway opponents recognize the importance of providing alternate plans that allow for the federal money to be spent in the region on mass transit or other highway projects. *See id.* at 154.

[22] But those who oppose a highway may not actually be affected by it. Liberal standing rules allow groups marginally or tangentially impacted to litigate against construction. *See generally* 4 K. DAVIS, ADMINISTRATIVE LAW TREATISE § 24 (2d ed. 1983).

ners and engineers to generate the least intrusive projects possible. Under the threat of environmental litigation, opponents may force the Department of Transportation, through the Federal Highway Administration (FHWA) to modify, amend, delay, or cancel a project. At the very least, transportation planners may augment their plans by complying with the requirements of holding public hearings, soliciting comments, and consolidating the public reaction into a project's plans. Such local input should result in a more sensitive and prudent finished product.

But committed opponents, those with the energy and resources to pursue every litigation option to its end, may extensively delay projects while their claims are litigated. In some instances, such as the Westway dispute discussed later,[23] delay may be fatal. In others, the prospect of delay may force the Department of Transportation to negotiate a settlement, altering the project to mitigate the environmental impacts. In all cases, opponents challenging highway construction under NEPA or section 4(f) may use litigation over the substantive and procedural requirements of these statutes as a substitute for a public debate over the true merits of a large-scale highway improvement plan.[24] NEPA, intended to guarantee that federal agencies follow proper procedure and adequately quantify environmental impacts, instead allows opponents a procedural veto over contentious projects. Courts that find violations of NEPA procedure remand the dispute back to the agencies for years of further study. Section 4(f)'s bar against taking parkland provides opponents with a tool to challenge the substantive decision that no alternative to taking parkland exists. In essence, the actual effect of NEPA and section 4(f) is similar—under both statutes, opponents may obtain judicial impediments to highway construction.

Against such a background, this Note will examine how NEPA and section 4(f) have influenced the process of making

---

[23] *See* Sierra Club v. Corps of Eng'rs, 772 F.2d 1043 (2d Cir. 1985) (reversing in part District Court (S.D.N.Y.), but leaving intact District Court's ruling that landfill permit was improperly granted). For a discussion of the Westway case, see *infra* notes 121–137 and accompanying text.

[24] A federal district court aptly summarized the issue in 1972:

> This court is fully aware that environmental laws may be misused by those who would like to see not merely compliance by government officials with the law, but the disruption, delay, and destruction of highway projects in general. Whether highways are good or bad as a general rule is not of concern to this court; that is a legislative matter . . . .

Lathan v. Volpe, 350 F. Supp. 262, 269 (W.D. Wa. 1972).

public choices with regard to highway construction projects. The Note will focus on developing a theory of proper public decision-making and then test this model through an exploration of the current situation, in which the use of environmental statutes has functionally allowed litigation to supplant public debate as a determinant of highway construction. Through an examination of the current statutes and the outcomes of litigated situations, this Note concludes that litigation, whether it results in a procedural or substantive ruling from a court, is a poor method for making effective public policy choices on issues such as highways. Litigation should not substitute for public debate on major public issues.

Part I of this Note will develop a model process for making public choices concerning highway projects, including a means of analyzing and choosing among public policy alternatives. Particular attention will be paid to the integration of unquantifiable environmental concerns into the decision-making calculus.

Part II will describe the emergence of litigation under NEPA and section 4(f), which has in part supplanted the interactive processes discussed in Part I. Following a brief political and economic history of limited-access highway development, the Note will focus in turn upon these two major environmental statutes, and examine their scope as defined by both their language and application by the courts.

Part III will explore the leverage, in litigation and negotiation, that NEPA and section 4(f) have given highway opponents confronting the FHWA. The Note will then examine actions brought under NEPA, specifically the Westway controversy in New York, and under section 4(f), illustrated by the H-3 litigation in Hawaii (highway opponents were victorious in both these cases).

Part IV will describe the administrative changes with which the FHWA has responded to these legal challenges. With the changes in place, the FHWA has been able to protect its projects from opponents in the majority of cases. It has done this in two ways: first, by litigating disputes successfully, as in the case of the Blue Route; and second, by negotiating settlements with opponents, as in the case of the Century Freeway.

Finally, Part V will propose amendments to the environmental statutes that would encourage all parties involved in highway disputes to use negotiation and/or mediation before litigation.

This change in the decision-making process would force the parties to spend more time debating the ultimate merits of highway projects—for example, the economic, social, and environmental costs and benefits of the projects—and less time arguing over the less important procedural and administrative issues, such as whether the FHWA adequately undertook the environmental review process. A decision-making process that includes the proposed amendments would better consider and reflect the needs of the public.

I. Developing a Theory of Public Choice

A. *Development of Current Decision-making Process*

Arriving at a decision about building a new highway entails making a societal choice from among alternative conceptions of a public good.[25] The plan to construct a large highway project forces a region to weigh the anticipated benefits to the area against the burdens placed on those in the right of way. Highway construction also produces long-term changes in the environmental and socioeconomic character of the region.[26] Since decisions of this nature involve apportioning benefits and burdens, allocating scarce resources, and expending tax revenues, public choices about transportation projects are essentially political choices and products of citizen interaction with the political system and its administrative appendages.[27]

Our society places the authority to make this choice with the elected political leaders acting through the appointed transportation-agency officials at the local, state, and national level.[28] Decisions to create public goods require such centralized au-

[25] Policy analysts attempt to reduce the public choice phenomena to identifiable and scientific steps. For a discussion of the ultimately political nature of public decisions, see E. Stokey & R. Zeckhauser, A Primer for Policy Analysis 257–66 (1978); *see also* E. Haefele, Representative Government and Environmental Management 15–19 (1973) (using the problem of environmental management to examine how American society in the late twentieth century makes social choices).

[26] For a discussion of the changes to the environment wrought by growth, see The Use of Land (W. Reilly ed. 1973).

[27] *See* E. Stokey & R. Zeckhauser, *supra* note 25.

[28] While the highest secretary and assistant secretary positions in the federal Department of Transportation and the state transportation departments are political appointments, civil service level administrators carry out most of the program implementation. *See* Federal Staff Directory 568 (8th ed. 1989).

thority to coordinate disparate interests.[29] The need for correlative decision-making on public goods, especially those that distribute burdens, stems from the inability and inefficiency of individuals and localities to affect such actions on their own.[30]

The multiple locality character of highways—even urban spurs or bypasses usually traverse many different towns, cities, and even counties—makes it inefficient for an individual political unit of a town or a city to construct a major highway project. Much as a sidewalk or park may be a public good in relation to the individuals who live on a city block, an interdistrict highway may be a public good for individual political subdivisions of a region. A higher political entity, either a state or the federal government, must coordinate local cities and towns and provide large roads.[31]

The original model of highway decision-making saw technically oriented public agencies making unilateral and unreviewable decisions.[32] Before the passage of section 4(f) and NEPA, powerful planners made decisions about highway placement, location, scope, and type.[33] Engineering concerns and technical feasibility, rather than human impact, often dominated these decisions.[34] Traffic demand projections and growth forecasts would be incorporated into a scientific highway plan designed

---

[29] The essential characteristics of a public good are (1) nonprovision, that is, no individual will receive enough benefit from the good to provide it on his or her own, (2) nonrivalry, meaning one person's use of the good does not diminish another's ability to consume the good, and (3) nonexcludability, or the impossibility of prohibiting noncontributors from using the good if only a few individuals had originally joined to create the good. Because no individual is willing to provide a public good, a central authority must step in and provide this service. *See* E. STOCKEY & R. ZECKHAUSER, *supra* note 25, at 305–08. Highways have these essential characteristics and are thus classic examples of public goods. Some argue, however, that a higher percentage of the road maintenance and construction burden should be left to the private sector through an alternative pricing system. *See, e.g.*, K. SMALL, C. WINSTON & C. EVANS, ROAD WORK (1989).

[30] *See* E. STOKEY & R. ZECKHAUSER, *supra* note 25, at 315–16.

[31] *See id.* While state or federal governments must build large-scale highways, since their construction involves choosing which communities and neighborhoods must bear the burden of the right-of-way, towns, cities, or counties may undertake smaller road projects. Road projects may range from installing a traffic light to building I-90 from Seattle to Boston, and state or federal governments primarily provide the larger end of the spectrum.

[32] *See* THE USE OF LAND, *supra* note 26, at 179.

[33] Perhaps the most glaring example of power vested in technocratic hands was Robert Moses' grip on the highway planning process in New York City. Moses, as chairman of the Triborough Bridge and Tunnel Authority, and Commissioner of Parks, was able to exercise almost unfettered discretion in highway construction in New York City. *See* R. CARO, *supra* note 15.

[34] *See* Brooks, *Environmental Decision Making: When Values Conflict*, in WHEN VALUES CONFLICT 131 (L. Tribe, C. Schelling & J. Voss, eds. 1976).

to meet the transportation needs of the community.[35] This model reflected the public belief of the 1950's and early 1960's in solving problems using apolitical technical analyses proffered by administrative experts.[36]

After the passage of section 4(f) and NEPA, the highway planning process evolved into a process stressing participatory decision-making.[37] Participatory decision-making involved committed segments of the public using public hearings and legal challenges to inject their particular views into the planning process.[38] The public hearing requirements and the threat of citizen suits forced highway planners to consider the road's human impacts as well as the traditional technical concerns. Public perception changed to recognize that highways affect communities not only by changing traffic patterns, but also by prompting other socioeconomic changes.[39]

A more satisfactory transportation planning process would balance the needs for centralized and comprehensive national transportation planning with a concern for the specific local impacts that each highway project causes.[40] Decisions on national transportation priorities, such as determining the allocation of scarce transportation resources both intra-modally and inter-modally, must be made at the national or state level through the electoral political process. In this view, administrative agencies should make many decisions between competing intra-modal concerns, but overall policy should be resolved through political mechanisms.

The most difficult aspect of this conception of transportation policy involves the apportionment of power to determine the final form of roads between transportation agencies empowered to build roads, and affected citizens who want to alleviate their

---

[35] *See* ST. CLAIR, *supra* note 2, at 110.

[36] *See* Brooks, *supra* note 34, at 130.

[37] *See id.* at 131.

[38] For a discussion of the effect of the threat of litigation on the highway planning process, see *infra* Parts II, III, & IV.

[39] Highway construction creates two measurable categories of effects: outputs and impacts. Outputs emanate from the functional aspects of the project; they are reflected by changes in traffic patterns and flow, as measured, for example, by corridor speed, traffic volume, travel time, trip distance, and accident rate. Impacts are the non-traffic consequences of transportation improvements; they include the economic, social, psychological, and environmental changes experienced by those living in the vicinity of the project. *See* Burkhardt & Shaffer, *Social and Psychological Impacts of Transportation Improvements* 1 TRANSP. 207, 208.

[40] *See* E. HAEFELE, *supra* note 25, at 16–17.

burden at all costs.[41] Inherent in the conflict over the public choice involving a public good is its insulation from the political process. Where the appropriation of funds for transportation projects requires legislatures to debate issues publicly and determine priorities,[42] the interaction between agency and citizens becomes a struggle waged with litigation and skirmishes for publicity.[43]

B. *A Proposed Model*

A model system of public choice for highway projects would maintain the requirement that agencies give weight to citizens' concerns, but it would do so through an interactive process that ultimately might balance individual concerns with community and regional needs. After the political machinations that result in the approval of a route, the next step in the process would be engineering feasibility and corridor analysis, to determine possible routes that would serve the agreed-upon needs. The determination of the highway's specific traits, such as size, width, right-of-way, entrance and exit ramp configuration, and noise and sight mitigation measures would be influenced by and tailored to meet the specific needs of the community.

Instead of ad hoc scrambling for projects designed to attract federal funds, regions and localities should formulate comprehensive planning documents that evaluate the potential effects of highways on regions and communities.[44] These plans should discuss the social and economic characteristics of the region and then determine the role of transportation in the region's future. Such regular planning would provide some indication of

---

[41] *See* THE USE OF LAND, *supra* note 26, at 179.

[42] Legislatures may fund specific projects as part of transportation bills, so the determination of priorities also may be done legislatively. *See, e.g.*, Federal Aid Highway Act of 1987, 23 U.S.C. § 103 (1989) (setting forth the procedures to be followed by agencies and Congress in appropriating highway funds).

[43] *See, e.g.*, THE END OF THE ROAD, *supra* note 14. *See also* INSTITUTE FOR PUBLIC TRANSPORTATION, IT'S UP TO YOU: A CITIZEN'S GUIDE TO TRANSPORTATION PLANNING (1975) (criticizing the economic, social, and environmental costs of a transportation system built around the private automobile, and providing examples of how concerned citizens have altered highway design in their communities).

[44] For an example of a comprehensive planning document, see D. HAMMERSCHLAG, B. BARBER & J. EVERETT, THE INTERSTATE HIGHWAY SYSTEM AND URBAN STRUCTURE: A FORCE FOR CHANGE IN RHODE ISLAND 143, 145–46 (1976) (summarizing the effects of highway projects in Rhode Island and making specific recommendations of how planners can better respond to these effects).

community values and needs that could guide decision-makers formulating the specific engineering plans of the highway.

Local political leaders and citizens' groups should use the public hearing and comment process to evaluate the project's ability to meet transportation needs in light of the project's negative regional impacts. Political leaders should articulate a set of criteria and goals for transportation planning in their community. This might include an estimation of business and industry location, the needs for public transportation, the specific character of the community, and the effect the highway will have on local citizens, especially if circumstances have changed since the last comprehensive plan.

Choosing the exact location of a highway means deciding which specific individuals must suffer the burdens of the project. A highway's location must respect the nationally-mandated values of NEPA and section 4(f) and, consequently, must defer to environmental considerations when economically feasible.[45] A highway project should also defer to community values and be sensitive to local concerns, and it should include measures designed to mitigate the negative social impacts.[46] Decision-makers should opt for the location that minimizes social dislocation, provides maximum access, draws local traffic, and is consistent with local demography and topography.

Local citizens and politicians should negotiate with planners to include design features which lessen the negative impact of the road on the community. Noise barriers, design covers, reduced exit scope, restricted entry, viaducts instead of open cuts, extra cross-bridges, and landscaping may all help reduce the impact of a highway on a neighborhood and thus satisfy local concerns.

---

[45] Both NEPA and section 4(f) provide broad value statements for highway planners. NEPA requires a formal assessment of the environmental impact of a project and mitigation measures where possible. *See* 23 U.S.C. § 103 (1982 & Supp. 1987). *See also* NEPA regulations, 40 C.F.R. § 1502.14(f), § 1502.16(h), § 1503.3(d), § 1505.2(c), § 1505.3 (1988). Section 4(f) mandates that parkland may only be taken for highways if "there is no prudent and feasible alternative to using that land; and . . . the . . . project includes all possible planning to minimize harm . . . ." 49 U.S.C. § 303 (1982 & Supp. 1987). *See also* section 4(f) regulations, 23 C.F.R. § 771.135 (1988).

[46] Two examples of highway projects that were tailored to meet local concerns are: I-66 through Arlington County in Virginia near Washington D.C., contested in Agnew v. Adams (E.D. Va. [filed March 20, 1979]); and I-95 through Philadelphia along the historic waterfront, where planners severely restricted exit ramps and installed noise and visual mitigation features, contested in Neighborhood Preservation Coalition v. Claytor, No. 73-1506 (E.D. Pa. May 22, 1987) (amended consent decree approved).

Once committed to building a highway, the process must arrive at a plan that minimizes environmental harm and negative social impacts to the extent that it can do so and still maintain some overall cost efficiency in highway construction. Minimizing environmental harm may make some stretches of urban highway enormously expensive, thereby reducing the overall amount of highway construction in the United States.[47] Moreover, the nationwide requirements for highway repair and modification are burgeoning as the system ages; they too are competing for limited highway budget resources. Thus, a socially efficient level of mitigation must be found.[48]

Mitigation measures should be pursued to the level where the marginal cost of each dollar of mitigation equals the marginal benefit of each dollar of highway funds spent elsewhere. The FHWA should spend only enough dollars on mitigation to make the project feasible and acceptable, a sum likely to be less than the level of mitigation desired by the community. While political realities may force the FHWA to complete expensive sections of highway, some attention to the overall socially efficient level of highway construction would allow for reallocation of scarce resources to more efficient areas.

Communication and discussion among administrative agencies, political leaders, and citizens are the most important features in this model of highway decision-making. The model would ideally see a non-confrontational forum for negotiation of specific design and location issues. But the realities of agency entrenchment, strident citizen opposition, and lack of consensus on overall transportation goals have produced a different and stagnated system.

II. Historical Background to Highway Disputes

A. *Interstate Highways*

The federal government has long been involved in public highway construction, taking an active role on this issue in coordination with the states. Federal involvement in roads pre-

[47] *Cf.* Kussy, *supra* note 11, at 253–57 (discussing the use of federal highway funds for mitigation measures).

[48] *See* Congressional Budget Office, The Interstate Highway System: Issues and Options app. at 60 (1982) [hereinafter cited as CBO Study].

dates the automotive age,[49] and the first federal highway legislation was passed in 1912.[50] The notion of federal funding for highway construction was first realized in 1916 with the passage of the Good Roads Act,[51] which authorized fifty percent federal funding for some projects. Five years later, Congress enacted the Federal Highway Act of 1921.[52] The Federal Aid Primary System, which the 1921 Act established, used a fifty-fifty federal-state cost sharing formula.[53] Before the Depression, federal highway programs focused on connecting distant cities and isolated rural towns and, not surprisingly, rural roads were the bulk of those constructed.

The Depression and World War II eras saw increased pressure for federal involvement in interurban, urban, and suburban highway construction. Industrial and social forces sought upgraded road facilities and a federal commitment that transcended the largely successful effort to link previously isolated rural communities.[54]

Toward the end of World War II, plans were drawn for an extensive national network of limited-access highways, essentially laying out the routes for the Interstate System.[55] This Interstate System was seen as a necessary outgrowth of consumers' pent up demand for new housing, mobility, and efficient transport of goods, needs that had been intensifying during the fallow periods of the Depression and World War II. The new highways, especially the urban and suburban portions, opened up previously inaccessible or inconvenient areas to settlement. This in turn stimulated the suburbanization of post-war baby boom families.[56]

---

[49] In 1893, the Office of Road Inquiry was created under the auspices of the Department of Agriculture to coordinate road research by, for example, publishing maps. In 1905 this office was merged into the Office for Public Roads and Rural Engineering. *See* ST. CLAIR, *supra* note 2, at 19.

[50] $500,000 was appropriated to improve rural postal roads. *See id.*

[51] Federal-Aid Road Act of 1916, Pub. L. No. 64-156, 39 Stat. 355.

[52] Federal Highway Act, Pub. L. No. 67-87, 42 Stat. 212, 213–19 (1921).

[53] *See* ST. CLAIR, *supra* note 2, at 21. The remnants of these federal highways still exist and are often important arteries, such as the U.S. Routes.

[54] For a discussion of the factors that led to federal development of urban roads, see *id.* at 121–24.

[55] *See generally id.* at 151–57.

[56] For the standard exposition of this theory, see generally J. RAE, THE ROAD AND CAR IN AMERICAN LIFE 225–27 (1971). Two other explanations have been offered for the funding of the interstate system. One holds that the government was impressed with the military efficiency of the German Autobahns during World War II. In fact, the Interstate Highways are officially called the National System of Interstate and Defense

Once plans for the Interstate System had been drafted, the extent of federal funding became the major issue. In the Federal-Aid Highway Act of 1956, Congress created the National System of Interstate and Defense Highways.[57] The federal government would fund 90% of the cost of 42,944 miles of Interstate highway, with the states assuming the other 10%.[58] By 1980, all but 1575 miles had been completed or were under construction,[59] and while only accounting for 1.07% of the country's total road mileage, interstate highways carried 19% of the domestic auto and 32% of the truck traffic.[60] By 1986, the size of the system had expanded to 43,844 miles.[61] The 1987 Federal-Aid to Highways Act funded previously approved projects, such as the Depressed Central Artery (I-93) in Boston,[62] and declared the Interstate System complete with these final appropriations.

Transportation planners have recognized that the highway system in general, and the Interstate System in particular, is nearing the end of its useful life and needs substantial repair.[63] Many urban areas also plan to upgrade their current systems by widening existing arteries.[64] In addition, a disproportionate per-

---

Highways. *See* THE END OF THE ROAD, *supra* note 14, at 18–19. A different and more conspiratorial theory posits that large industrial enterprises, such as large automobile manufacturers, suburban developers, construction interests, and urban and rural right-of-way owners pressured the government to fund a massive system in the absence of public galvanization on the issue. For an examination of conspiracy theories, see ST. CLAIR, *supra* note 2, at 56–77. For a more complete discussion, see generally H. LEAVITT, *supra* note 16, at 111–55, and K.R. SCHNEIDER, AUTOKIND VS. MANKIND 16–17 (1972); *see also* Checkoway, *Large Builders, Federal Housing Programmes, Postwar Suburbanization* in READINGS IN URBAN ANALYSIS 173–91 (1983) (discussing the context in which suburban areas were established and developed).

[57] Pub. L. No. 84-627, 70 Stat. 374 (1956) (current version at 23 U.S.C. §§ 103–57 (1982)).

[58] For a discussion of the completion of the system and the scope of federal funding, see generally CBO STUDY, *supra* note 48, at xv–xxi, 1–6.

[59] *Id.* at xv.

[60] *Id.*

[61] FHWA HIGHWAY STATISTICS 1986, *supra* note 11, at 117 (Table HM-18).

[62] *See* Federal Aid Highway Act of 1987, Pub. L. No. 100-17, 101 Stat. 134, 23 U.S.C. § 103 (1982).

[63] *See* CBO STUDY, *supra* note 48, 1 app. at 60 (discussing current and projected problems with the Interstate System).

[64] Since building new urban highways is virtually impossible in some areas, the only way to confront mounting traffic problems in many cities is to increase the capacity of existing roads. For example, California has enjoyed great success through narrowing lane widths to accommodate more vehicles. Other cities, such as Atlanta and Boston, are seeking to widen or rebuild existing roads within their current placements. Since highways are now frequently surrounded by development, widening may incite opposition from those abutting the right-of-way. *See* Koepp, *Gridlock!*, TIME, Sept. 12, 1988, at 52–60.

centage of the uncompleted or unbuilt segments are in urban areas, where the economic and social costs of highways are greatest, and where the opposition to such projects may be concentrated.[65]

Against such a background emerge legal challenges to highway improvements.[66] The specter of a completed highway may galvanize local opponents, compelling them to organize into a group determined to fight the highway's construction. The membership of such groups may be quite diverse, including those faced with the loss of their homes, proponents of mass transit, and would-be protectors of the environment. Since the final design planning for such projects is done at the state level, local opposition may be heard loud and clear. Whether the road is a spur bisecting a neighborhood or a national connector traversing an area of great natural beauty, opponents may use a variety of statutes to embroil the project in lengthy litigation.[67]

B. *Statutory Authority Guiding Highway Litigation*

1. National Environmental Policy Act of 1969

A climate of increasing environmental concern led to the passage in 1969 of the National Environmental Policy Act (NEPA),[68] which requires that environmental concerns be inte-

---

[65] 32,784 of the total of 43,844 miles, or 74.77%, of Interstate Highways are rural. In addition, 226,416 of the 259,389 miles, 87.29%, of Primary Federal Roads, are categorized as rural. FEDERAL HIGHWAY STATISTICS 1986, *supra* note 11, at 117 (Table HM-18). By contrast, 477 of 1,742, 37.7%, of the uncompleted miles are in urban areas. Further, 100 of the 159 discrete unfinished segments, 62.89%, are in urban areas. CBO STUDY, *supra* note 48, at 77–95. In other words, while approximately 25% of Interstate Highways are urban, 38% of the uncompleted projects are urban.

[66] Opposition to transportation projects is by no means limited to highway projects. Extending airport facilities, for example, often raises the antagonism of local residents who feel increased noise and traffic are antithetical to maintaining their neighborhood's character. Mass transit projects, on the other hand, are often contested in a different way. Local residents often petition and organize to be included on the path of subway or light rail improvements, as subway access often greatly increases property values.

[67] For a discussion of legal challenges, see *infra* Part III.

[68] NEPA, *supra* note 6.

grated into the planning of all major federal projects.[69] NEPA provides that for all major federal construction projects significantly affecting the environment, a statement of environmental impact must be prepared outlining the substantive environmental consequences of the action and its alternatives.[70] The process of preparing the environmental impact statement (EIS) involves an interdisciplinary study of the environmental, social, and economic consequences of the proposed project.[71]

NEPA did not create any new substantive environmental rights, but rather established a baseline level of procedural review for each major project.[72] Acknowledging the growing political pressure to respect the environment,[73] the Act declared a national policy of incorporating environmental concerns into federal projects.[74] The legislative history suggests Congress intended NEPA to apply broadly to all aspects of federal action.[75]

---

[69] NEPA's most important provision requires that officials:

> (C) include in every recommendation or report on proposals for legislation and other major Federal actions significantly affecting the quality of the human environment, a detailed statement by the responsible official on—
> (i) the environmental impact of the proposed action,
> (ii) any adverse environmental effects which cannot be avoided should the proposal be implemented,
> (iii) alternatives to the proposed action,
> (iv) the relationship between local short-term uses of man's environment and the maintenance and enhancement of long-term productivity, and
> (v) any irreversible and irretrievable commitments of resources which would be involved in the proposed action should it be implemented.

42 U.S.C. § 4332(C) (1982).

[70] For the regulations concerning the preparation of environmental impact statements, see 40 C.F.R. § 1502 (1988).

[71] Regulations provide that the EIS shall incorporate the "use of the natural and social sciences and the environmental design arts . . . ." 40 C.F.R. § 1502.6 (1988).

[72] 42 U.S.C. § 4332 (1982). *See also* 2 F. GRAD, TREATISE ON ENVIRONMENTAL LAW § 9.01, at 9–19 (1988).

[73] *See* 42 U.S.C. § 4331(a) (1982).

[74] NEPA provides for two new affirmative governmental actions. First, it requires the government to prepare environmental impact statements. *See* 42 U.S.C. § 4332 (1982). Second, it creates the Council on Environmental Quality (CEQ), an advisory panel to the President intended to coordinate federal environmental compliance. *See* 42 U.S.C. § 4342 (1982). Subsequent statutes gave the CEQ even broader powers. *See* 2 F. GRAD, *supra* note 72, § 9.01, at 9–19. But its original function, as created by NEPA, was to advise the President on all environmental matters and to coordinate governmental action. *See* 42 U.S.C. § 4344 (1982). The legislative history indicates that the CEQ was intended to be the primary thrust of the NEPA bill. Indeed, the earlier sections declaring the national environmental policy, *see* 42 U.S.C. § 4331 (1982), and providing for the EIS, *see* 42 U.S.C. § 4332 (1982), were added later. *See* 2 F. GRAD, *supra* note 72, at § 9.01, at 9–31.

[75] The debate about one clause in particular, with special relevance to highway litigation, was the result of compromises during debate about NEPA. That federal agencies shall comply "to the fullest extent possible" was a concession to Senator Muskie's (D-Me.) concerns about the applicability of NEPA to different federal agencies, and their

NEPA compels agencies to administer its process with an expansive notion of the types of federal actions that trigger its implementation.[76]

In addition to the EIS requirement and creation of the Council on Environmental Quality (CEQ), NEPA's mandate includes a charge to federal agencies to incorporate all existing environmental laws into program operations "to the fullest extent possible."[77] Coupled with the legislative history indicating a general intent of broad application, this charge sets sweeping requirements for federal agencies.[78] NEPA requires rigorous study and assessment of alternatives to a planned action that has harmful environmental consequences.[79] In so doing, it guarantees environmental review through the environmental statement process. By requiring federal agencies to file an EIS, NEPA compels consideration of relevant alternatives, with judicial review available for an improperly considered decision. Since NEPA creates no substantive minimum standard for environmental consideration, the EIS process itself becomes NEPA's method for assuring compliance with its standards.

NEPA regulations guide an agency through its preparation and deliberation over an EIS.[80] An agency that determines a project is a "major Federal action[] significantly affecting the quality of the human environment"[81] must then begin the EIS

---

degree of compliance. 115 CONG. REC. 40,423 (1969). For a discussion of Senator Muskie's concerns about NEPA, see 2 F. GRAD, *supra* note 72, § 9.01, at 9–39. For a general discussion of NEPA's legislative history, see Ackman, *Highway to Nowhere: NEPA, Environmental Review and the Westway Case* 21 COLUM. J.L. & SOC. PROBS. 325, 342–46 (1988).

[76] Subsequent NEPA regulations codify this broad mandate. Regulations interpreting NEPA policy use the broadest possible language in discussing its intended effects. *See* 40 C.F.R. § 1500.2 (1988).

[77] 42 U.S.C. § 4332 (1982).

[78] *See, e.g.*, 42 U.S.C. § 4331(b) (1982), on the scope of NEPA's application:

> (b) In order to carry out the policy set forth in this Act, it is the continuing responsibility of the Federal Government to use all practicable means, consistent with other essential considerations of national policy, to improve and coordinate Federal plans, functions, programs and resources . . . .

See also 42 U.S.C. § 4332 (1982), which begins:

> The Congress authorizes and directs that, to the fullest extent possible: (1) the policies, regulations, and public laws of the United States shall be interpreted and administered in accordance with the policies set forth in this Act . . . .

[79] *See* 42 U.S.C. § 4332(C), (E) (1982); 40 C.F.R. § 1500.2(e) (1988); Herson, *Project Mitigation Revisited: Most Courts Approve Findings of No Significant Impact Justified by Mitigation*, 13 ECOLOGY L.Q. 51–52 (1986).

[80] *See* 40 C.F.R. §§ 1500–08 (1988).

[81] 42 U.S.C. § 4332(C) (1982). No matter what the size or scope of the project, FHWA-funded actions require the preparation of an Environmental Assessment (EA) as a threshold matter. 40 C.F.R. § 1501.3 (1988) delineates when an agency should prepare

process.[82] Drafts of the EIS documents must be made available for public review and commentary.[83] The final EIS presents a list of alternatives and the environmental consequences of each, including a discussion of ways to mitigate the environmental damage.[84]

NEPA applies to a large percentage of highway construction,[85] for federally financed highway projects will generally have significant effects on the environment.[86] As a result, the FHWA will almost always comply with NEPA's strictures as a matter of policy.[87] The FHWA classifies highway projects into three categories, each requiring a different level of NEPA compliance.[88] Federally funded highway projects usually have an EIS prepared before construction begins.[89]

2. Section 4(f) of the Department of Transportation Act

Congress enacted section 4(f) of the Department of Transportation Act[90] as a substantive law intended to curtail the routing of highways through parkland. The provision permits highway construction to use parkland only if the Department determines that "no prudent and feasible alternative"[91] route

---

an EA. An EA does three things: first, it justifies the decision to prepare an EIS or a Finding of No Significant Impact (FONSI) (provided for in § 1508.13); second, it shows compliance with NEPA when no EIS is necessary; and third, it "[f]acilitate[s] preparation of a statement [EIS] when one is necessary." 40 C.F.R. § 1508.9(a) (1988). Section 1508.9(b) allows for a preliminary EA, when an EIS is necessary, that provides an initial summary of relevant issues and the EIS process. *See* Herson, *supra* note 79.

[82] NEPA should be applied early in the decisionmaking process. *See* 40 C.F.R. § 1501.2 (1988).

[83] 40 C.F.R. § 1503 (1988).

[84] 40 C.F.R. § 1502.14 (1988).

[85] *Cf.* 42 U.S.C. §§ 4321, 4331 (1982) (declaring the purposes of and national environmental policy set forth in NEPA).

[86] As an extreme example, one mile of interstate highway may require up to 48 acres of land. *See* 3 F. GRAD, *supra* note 72, § 11.02, at 7 (citing COUNCIL ON ENVIRONMENTAL QUALITY, ENVIRONMENTAL QUALITY—FIFTH ANNUAL REPORT 39 (1974)).

[87] *See* F. ANDERSON, *supra* note 10, at 84.

[88] 23 C.F.R. § 771.115(a)-(c) (1989).

[89] The precise meaning of the terms and phrases in NEPA, such as "major federal actions" and "to the fullest extent possible," and what consequences of highway building are significant enough to trigger NEPA's EIS process, have been the subject of a great deal of litigation and controversy. *See supra* note 75.

[90] 49 U.S.C. § 303 (1982).

[91] 49 U.S.C. § 303(c)(1) (1982).

exists, and that the routing will "minimize harm to the park"[92] if it must go through the park.[93]

Section 4(f) curtails the tendency of highway planners to route urban highways through scarce public parkland. Planners seeking to expedite a highway's construction have an incentive to build on parkland for two reasons. First, using parkland lessens the total cost of acquiring the right-of-way. The parkland is already publicly owned, so the government does not need to purchase many acres of expensive, developed urban land. In addition, no one lives on the parkland, so the government also avoids the costs of condemning existing property and relocating those residents who live in the path of the highway. Second, using parkland and thereby minimizing the displacement of residents helps avoid community protest,[94] which can cause costly delay and generally reduce the political support critical to completing a highway.[95]

Legal challenges to highway projects based on section 4(f) involve litigants seeking a review of the Department of Transportation's[96] consideration of the approved route to assure that the route chosen minimizes the use of parkland. Section 4(f) is a substantive environmental statute, and agencies often implement it while complying with NEPA's EIS provisions.[97]

---

[92] 49 U.S.C. § 303(c)(2) (1982).

[93] Relevant portions of section 4(f) provide as follows:

> (c) The Secretary may approve a transportaion program or project . . . requiring the use of publicly owned land of a public park, recreation area, or wildlife and waterfowl refuge of national, State, or local significance, or land of an historic site of national, State, or local significance . . . only if:
>
> (1) there is no prudent and feasible alternative to using that land; and
>
> (2) the program or project includes all possible planning to minimize harm to the park, recreation area, wildlife and waterfowl refuge, or historic site resulting from such use.

49 U.S.C. § 303 (1982 & Supp. 1987).

[94] *See generally* R. CARO, *supra* note 15. Moses recognized that using parkland irritated no residents. Hence, major New York roads went through Riverside Park, Flushing Bay Park, and other parks.

[95] While legal remedies might not be available, sometimes community protest storms so loudly that political support for a highway project fades, killing construction. *See* R. BAUMBACH & W. BORAH, *supra* note 3; A. LUPO, *supra* note 15. *But see* R. CARO, *supra* note 15.

[96] In the planning process for federally-funded highways, the federal Department of Transportation usually delegates routing decisions to its state counterpart. The state department of transportation will then implement the necessary process, including such steps as commissioning engineering studies, verifying that the plan conforms to safety specifications, or holding public hearings on the proposed plans. FHWA approval comes only when a more final plan takes shape. For the list of procedures guiding right-of-way acquisition and the highway planning process, see 23 C.F.R. § 712.204 (1989).

[97] *See* 23 C.F.R. § 771.109(a)(3) (1989).

Very often litigants assert both section 4(f) and NEPA claims in the same lawsuit, but litigants may file suits based solely on section 4(f) grounds.[98] Litigants may make a procedural challenge under section 4(f), contesting the decision not to prepare a section 4(f) statement or challenging the quality of work done in preparing the 4(f) statement, particularly the adequacy of its consideration of alternatives. Substantive suits under section 4(f) contest the agency's decision to use parkland based on the statement.[99]

As a substantive statute, section 4(f) provides a broad mandate to the Department of Transportation to minimize the use of parkland for highway construction. The courts have set the substantive bounds of section 4(f), determining what constitutes "use" of parkland, whether harms are mitigated to the fullest extent possible, and what constitutes adequate consideration of alternatives to determine that they are not "feasible" or "prudent."[100]

The first major judicial interpretation of section 4(f) came in the case of *Citizens to Preserve Overton Park v. Volpe*,[101] where the Supreme Court interpreted the statute as a "plain and explicit bar" to constructing a six-lane expressway through a downtown Memphis, Tennessee park.[102] The Supreme Court ruled that section 4(f) should be interpreted broadly, reading the statute as a clear command by Congress to preserve parkland.[103]

In the wake of *Overton Park*, courts have interpreted section 4(f) to the broadest possible extent. Courts divide the use re-

---

[98] The FHWA environmental lawsuit study found that of the 40 cases decided during the study period, 15 (40%) contained both NEPA and § 4(f) claims. Fourteen of the 40 (35%) contained only NEPA claims, 10 of the 40 (25%) contained only section 4(f) claims, and 1 contained neither. A suit involving both section 4(f) and NEPA claims might even argue that the NEPA violation was the failure to prepare an EIS that contained a section 4(f) statement. *See* FHWA ENVIRONMENTAL LAWSUIT STUDY, *supra* note 9, at 9.

[99] *See id.*

[100] *See* Miller, *Department of Transportation's Section 4(f): Paving the Way Towards Preservation* 36 AM. U.L. REV. 633, 644–60 (1987) (analyzing judicial decisions that set boundaries of section 4(f)'s scope). Miller discusses another aspect of section 4(f) open to judicial interpretation: whether there has been an impairment of an historic site sufficient to qualify as "constructive use." *See id.* at 650.

[101] 401 U.S. 402 (1971).

[102] *Id.* at 411.

[103] The court also enunciated the "hard look" doctrine as the standard for reviewing agency decisions. This doctrine indicates that reviewing courts will take a "hard look" at the agency's decision to determine if it was "arbitrary" or "capricious." For a discussion of the hard look doctrine, see S. BREYER & R. STEWART, ADMINISTRATIVE LAW AND REGULATORY POLICY 341 (1985).

quirement into two subdimensions: actual use and constructive use. Courts look first to see if an actual use of parkland is planned. Any actual physical use, no matter how trivial, merits preservation under section 4(f).[104] If there is no actual use, the courts will look for constructive use; they will acknowledge that air pollution, noise, and the visual byproducts of a highway may amount to constructive use of a park and thus warrant application of section 4(f).[105]

Courts have recently expanded their conception of constructive use.[106] In *Citizen Advocates for Responsible Expansion, Inc. (I-CARE) v. Dole*,[107] the court held that the mere proximity of the highway to a protected site constitutes a use. A recent court decision has followed the *I-CARE* court's philosophy and expanded the constructive use doctrine.[108] Constructive use may also occur when the highway passes near an historic place. In *Stop H-3 Ass'n v. Dole*,[109] the Ninth Circuit recognized a Hawaiian expressway's proximity to highway-created parkland as constructive use.[110]

III. Highway Opposition Litigation Using NEPA and Section 4(f)

Highway opponents recognize that NEPA and section 4(f) provide them with unequal degrees of leverage over highway planners, because the statutes involve different sorts of chal-

---

[104] *See, e.g.*, Louisiana Environmental Society, Inc. v. Coleman, 537 F.2d 79 (5th Cir. 1976).

[105] The first case to recognize constructive use was Brooks v. Volpe, 460 F.2d 1193 (9th Cir. 1972) (preliminary injunction granted halting construction of highway proposed to span Cross Lake, in order that the Secretary could consider alternative routes which might minimize use of lake).

[106] *See* FHWA Environmental Lawsuit Study, *supra* note 9, at 25.

[107] 586 F. Supp. 1094 (N.D. Tex. 1984), *rev'd*, 770 F.2d 423 (5th Cir. 1985).

[108] *See* Town of Belmont v. Dole, 766 F.2d 28 (1st Cir. 1985), *cert. denied* 474 U.S. 1055 (1986) (upholding regulation which included archaeological sites within the protective scope of section 4(f), where close placement of highways may decrease the sites' historic value).

[109] 538 F. Supp. 149 (D. Haw. 1982), *aff'd in part and rev'd in part and remanded* 740 F.2d 1442 (9th Cir. 1984), *cert. denied* 471 U.S. 1108 (1985).

[110] For a discussion of the *H-3* situation and the FHWA's reaction, see *infra* notes 145–161 and accompanying text.

lenges to highway projects.[111] Litigation based on NEPA focuses on the procedure used by the agency to consider the environmental impacts of a highway project. Successful opponents in NEPA cases force the FHWA to conduct lengthy, costly environmental impact studies (or to supplement previous studies) that significantly delay highway projects.[112] Once the FHWA produces an adequate study, though, the project continues. By contrast, section 4(f) attacks the substance of a highway project, barring any construction which uses parkland (unless no alternative exists). Litigants using it may compel the FHWA to change the actual plans for the highway project and use an alternate route. Thus a successful challenge under section 4(f) will also result in delay. But while a NEPA delay may end with the highway project continuing as planned, a section 4(f) delay may force the FHWA to revamp or basically alter the routing of the highway so as not to use parkland.[113]

A. *NEPA Suits*

1. Highway Opponents' Leverage in NEPA Litigation

Challenges to highway projects based on NEPA violations seek to prove that the FHWA failed to consider fully the environmental impacts of the proposed construction.[114] A successful NEPA challenge will result in the court issuing an injunction halting construction until the FHWA issues an EIS analyzing

---

[111] The National Wildlife Foundation and the Environmental Action Foundation became so adept at fighting highways that they had the swagger to publish a manual, THE END OF THE ROAD, *supra* note 14, a citizens' primer on how to stop a highway project. While their helpful hints primarily focus on public organization and creating large-scale opposition to projects, they dedicate one full chapter of the manual to legal challenges. *Id.* at 115–35.

[112] THE END OF THE ROAD, *supra* note 14, recognizes delay as the indisputable friend of the highway opponent. The authors concede, however, that delay must be used in conjunction with other tactics to diminish political support for the project. *Id.* at 115.

[113] *See, e.g.*, Citizens to Preserve Overton Park v. Volpe, 401 U.S. 402 (1971), in which the Supreme Court sent the case back for a redeterminination of alternative routes for I-40. The Secretary then decided that any redesigned alignment of I-40 would either cause too much local dislocation or constructively use land in Overton Park. As a result, that segment of I-40 was never completed. *See* Citizens to Preserve Overton Park v. Brinegar, 494 F.2d 1212 (6th Cir. 1974) (secretary is not required to propose a feasible and prudent alternative route after disapproving the Overton Park route; the State must make an affirmative proposal for the Secretary's evaluation).

[114] *See supra* notes 68–89 and accompanying text for a discussion of requirements and provisions of NEPA.

the environmental impacts that had been ignored in the original planning.[115] In the best case scenario for opponents, the injunction lasts until the FHWA completes the new or enhanced EIS to the court's satisfaction.[116]

No matter how meritorious the claims of the highway opponents, suits under NEPA may at best achieve a delay in the project until the FHWA corrects the faulty procedure.[117] Nevertheless, opponents understand the value of delay in stopping highway projects.[118] Successful litigation brings publicity and focuses attention on a potential fracture within a community. NEPA litigation may actually kill projects if the delay is able to diminish political support.

In the case of highway projects with only marginal support from the start, the goal of a NEPA suit may simply be to cause excessive delay and cost by forcing the FHWA to develop or supplement an EIS.[119] While a court order requiring the agency to do this may only delay a solid project, it can be fatal to a marginal proposal.[120] In such instances, the FHWA might at-

---

[115] THE END OF THE ROAD, *supra* note 14, at 117, lists three possible goals of NEPA suits. First, the suit could force the FHWA, through the state DOT, to disclose all the foreseeable environmental impacts. This usually involves preparing a new EIS, or issuing a supplemental EIS (SEIS). Second, the judge could order the government to prepare a new EIS (or SEIS) to discuss and consider alternatives inadequately addressed or omitted in the original EIS. Third, the suit could result in the preparation of an EIS in which the FHWA and state DOT determine that a full EIS is unnecessary.

[116] *See, e.g.*, Sierra Club v. United States Army Corps of Eng'rs, 772 F.2d 1043 (2d Cir. 1985) (affirming district court decision to vacate issuance of landfill permit).

[117] *See* THE END OF THE ROAD, *supra* note 14, at 115–17.

[118] THE END OF THE ROAD, *supra* note 14, at 97–114, implores litigants interested in successfully pursuing a NEPA claim to use the delay to galvanize public opposition and ultimately erode political support for the project. Delay drives up the cost of the project and gives the media time to focus on the affected citizens. The ultimate goal is to persuade local, state, and national politicians to withdraw support from the project. For example, erosion of political support killed urban transportation networks in Boston, *see* A. LUPO, *supra* note 15; Washington, D.C., *see* THE END OF THE ROAD, *supra* note 14, at 97–114; and Baltimore, *see* BELTWAY STUDY, *supra* note 12, at B-8 to B-11. Contrast this paralysis with the situation in New York, where unaccountability to political pressure in an earlier age allowed Robert Moses to construct the great network of highways in that city. *See generally* R. CARO, *supra* note 15.

[119] Of the 29 suits filed under NEPA during the study period, the FHWA won 15, meaning no additional EIS was necessary, and lost 7, meaning the court required the preparation of additional environmental documents. In addition, 6 suits were settled. In losses, the court either ruled the EIS inadequate or mandated preparation of an EIS. FHWA ENVIRONMENTAL LAWSUIT STUDY, *supra* note 9, at 17.

[120] A suit may take a substantial amount of time to come to trial, but courts will grant preliminary injunctions to halt construction pending environmental review. *See, e.g.*, Action for Rational Transit v. West Side Highway Project, 536 F. Supp. 1225 (S.D.N.Y. 1982), *aff'd in part*, 701 F.2d. 1011 (2d Cir. 1983) (upholding preliminary injunction).

tempt to negotiate a solution rather than face complete loss of the project.

2. Westway

Perhaps the greatest NEPA litigation victory for highway opponents came in *Sierra Club v. United States Army Corps of Eng'rs*,[121] where determined highway opponents succeeded in killing the Westway project by suing under NEPA. In the Westway litigation, opponents successfully challenged a series of EIS proposals. They managed to delay the project from 1973 until 1985, when, facing an imminent loss of federal funding, New York City officials reluctantly traded the federal funds earmarked for Westway for an approximately equal amount of federal money to be used for local mass transit projects and construction of a scaled-down road.[122]

Analysis of opponents' victory in the Westway case reveals how a statute like NEPA, which concerns procedural issues, may ultimately affect the substance of a project. Westway represents the worst possible outcome for highway proponents, namely, the cancellation of the entire project. The case did involve a unique confluence of circumstances, including federal deadlines, which make it an exaggerated example of a highway dispute. Nevertheless, Westway highlights the types of factors that may unite to end a project.

---

[121] 772 F.2d 1043 (2d Cir. 1985).

[122] In 1973, the government amended the Federal Aid Highway regulations to allow cities to "trade-in," or exchange, the amount of the federal share of a highway's cost for mass transit purposes. Since the cost of these urban highways often ran into hundreds of millions of dollars, this infusion of money into mass transit was intended to be a significant boon to the local transportation system. Before Westway, 338.7 miles of highway had been traded in for a value of $8.1 billion. J. Lusk, Memorandum: Westway Trade-In 1 (July 17, 1985) (unpublished memorandum) [hereinafter Lusk Memo].

The first and most striking example of such a trade-in was Boston's exchange of the funds for the inner beltway (I-695) and southwest expressway (I-95) for rapid transit money ultimately used to construct the new southwest corridor Orange line project and augment commuter rail service to the areas southwest of Boston. Opponents argue, however, that despite the success of the new Orange line, it diverts mass transit from a poor neighborhood to a wealthier one and still does not provide automotive access to the Route 128 high-technology corridor, where a large percentage of the area's job growth has occurred. *See* Boston Globe, May 5, 1987, at 21, col. 4.

The Westway situation originated in 1973, when a truck travelling on the old elevated West Side Highway[123] fell through the road, closing the highway from 59th Street south.[124] When inspections showed the road to be in an advanced state of disrepair, the favored replacement proposal involved an "outboard" below-ground expressway built on landfill in the Hudson River.[125] Despite federal, state, and local political support, the multi-billion dollar project ignited storms of protest from community groups and mass transit proponents.[126] Nevertheless, environmental and design study continued. During the period from 1975 to 1982, the FHWA, the New York Department of Transportation, and the Corps of Engineers were engaged in preparing and circulating the environmental impact studies for the enormous project. In 1982, opponents obtained an injunction blocking the Corps of Engineers' landfill permit until completion of a study on the effect of the landfill on the striped bass.[127]

In 1985, after finishing a new EIS on the striped bass question, the Corps of Engineers issued a new dredge permit, which subsequently was judicially invalidated under NEPA.[128] With the deadline for trading in interstate funds for mass transit approaching,[129] New York City and state officials decided to terminate the Westway project and use the federal portion of the

---

[123] The old West Side Highway, completed in the 1930's, was considered technologically advanced for its time. Heavy traffic and the passage of time ultimately took their toll on the road, though, rendering it obsolete. *See* L. BACOW & M. WHEELER, ENVIRONMENTAL DISPUTE RESOLUTION 105 (1984).

[124] Chronology of Westway History (1985) (unpublished internal memorandum prepared by Mayor's Office, City of New York) (on file with author).

[125] The proposal involved construction of a submerged six-lane highway created by 242 acres of landfill in the Hudson River. In addition to the road, the project, estimated in 1977 to cost $1.1 billion, would have created a 93-acre waterfront park and allowed other westside development. *See* Ackman, *supra* note 75, at 333–34. *See also* L. BACOW AND M. WHEELER, *supra* note 123, at 105–06.

[126] For a full discussion of political support for Westway, see Ackman, *supra* note 75, at 333.

[127] For a more complete review of the action of the government agencies and an analysis of the factors that led to the first injunction, see *id.* at 334–38.

[128] *See* Sierra Club v. United States Army Corps of Eng'rs, 772 F.2d 1043 (2d Cir. 1985).

[129] The Surface Transportation Assistance Act of 1982 set the deadline for trading in unused interstate funds for mass transit money at September 30, 1983. Pub. L. No. 97-424, 96 Stat. 2097 (1983). For programs in litigation at that time, such as Westway, the deadline was extended to September 30, 1985. Facing the potential loss of $1.725 billion, Mayor Koch and Governor Cuomo decided to trade in the funds on September 26, 1985. M. Cuomo & E. Koch, Westway Trade-In Memorandum of Understanding between Governor Cuomo and Mayor Koch (September 26, 1985) (unpublished memorandum) (on file with author).

Westway funds for local transit and smaller roads.[130] In the end, a combination of delay in the face of federal deadlines led to the death of Westway.

While Governor Cuomo and Mayor Koch did act rationally in stopping the project to avoid losing an infusion of $1.725 billion dollars into the city,[131] the procedural delay imposed by NEPA litigants forced the abandonment of a project supported by all major elected officials between 1974 and 1985. Critics of the project argue that botched environmental review led to Westway's demise, but in the absence of NEPA and the procedural delays it authorized, the huge project might have been completed.[132]

Westway's environmental review process, flawed from the start, allowed the opponents to turn procedural delay into a complete victory. The original EIS, issued in 1977, anticipated few effects of the landfill on the Hudson River ecology.[133] But the Corps of Engineers based this conclusion on an outdated study,[134] and the court applied NEPA to correct serious deficiencies in the original EIS.[135] When the second EIS also un-

[130] By 1985, the estimated total cost of the Westway project had escalated to $1.983 billion. The Surface Transportation Assistance Act (STAA), Pub. L. No. 97-424, 96 Stat. 2097 (1983) provided that the value of the trade-in would be 85% of the construction cost of the project. This would amount to $1.725 billion for New York. Of this money, at least $690 million up to $972 million would finance a replacement road. The remainder, originally targeted for the Second Avenue subway, would instead be spent on mass transit and local streets and bridges. Cuomo & Koch, *supra* note 129, at 3.

[131] The City of New York, in weighing its decision whether to trade in funds, recognized that a loss in the pending legal action would leave the City with no trade-in money and no Westway. Faced with such a choice, the only rational decision was to elect for the trade-in. *See* Lusk Memo, *supra* note 122, at 3.

[132] Ackman argues that NEPA gave environmental opponents a "procedural veto" over Westway. This veto allowed opponents to halt work solely because of procedural defects, while the merits of Westway were never passed on by any court. He further argues that the political support of so many elected officials over such a long period of time indicates that, via the ballot box, the public approved of Westway. *See* Ackman, *supra* note 75, at 327–30.

[133] The first Final EIS (FEIS), issued in 1977, called the interpier area that would be landfilled a "biological wasteland." *Id.* at 336. But commentary by the National Marine Fisheries Service (NMFS) on a draft EIS indicated that the NMFS did not believe the FEIS accurately assessed the impact of the landfill on Hudson River ecology. *Id.*

[134] The Corps based its 1977 EIS on a 1973 study, rather than conducting its own study. *Id.*

[135] The Second Circuit recounted the whole, long, legal history of Westway in its second review of the adequacy of the EIS. *See* Sierra Club v. United States Army Corps of Eng'rs, 772 F.2d 1043, 1047 (2d Cir. 1985). In addition to ruling on NEPA procedural grounds, the court also found that the EIS ignored public commentary from the Fisheries and Wildlife Service and EPA about the impact of the dredge permit on wildlife. *See* The Clean Water Act, 33 U.S.C. § 1344 (1972), *amended by* Act of Dec. 27, 1977, Pub. L. No. 95-217, § 67(a)(b), 91 Stat. 1600, and Act of Feb. 4, 1987, Pub. L. No. 100-4, Title III, § 313(d), 101 Stat. 45 (requiring consideration of public commentary).

derstated the harm to the spawning habitat of the striped bass, the Southern District vacated the landfill permit and granted a permanent injunction until the Corps conducted the proper study; the Second Circuit affirmed the decision to vacate the permit, although it did reverse the injunction.[136]

By attempting to minimize the ecological impacts of Westway, the Corps of Engineers allowed an attack based on procedure, rather than the merits of the highway, to kill the project. This result might have been avoided if the Corps had simply acknowledged that the landfill would have an ecological impact on the striped bass. The FHWA could then have decided that the economic, social, and political factors that originally led to the selection of the outboard option outweighed these harms.[137] Under these circumstances, a court following NEPA could have come to only one conclusion, that the agencies followed proper procedure; the court could not have stopped the project under NEPA. If still unpopular with the public, the decision to proceed with Westway at that point would have been essentially political, rather than judicial. In short, the decision-making process actually used by planners in the Westway case was far from ideal.

The decision on whether to proceed with a project of the scale and magnitude of Westway should be political, rather than the product of legal action. With the environmental risks properly quantified, the process of weighing those risks and the socio-economic changes accompanying the project against its desired effects should rest with those accountable at the ballot box.

Westway illustrates how highway opponents may utilize the procedural constraints of NEPA to stop a popular project. Westway was a unique project; an atypical set of factors, notably poor administrative performance and the resultant delay, killed it. The leverage that a NEPA suit gave the Westway opponents may not be attainable in other cases; litigants cannot look to Westway as a current model of successful highway opposition.[138] But despite the circumstances of Westway, the FHWA had to

---

[136] *See* Sierra Club v. United States Army Corps of Eng'rs, 614 F. Supp. 1475 (S.D.N.Y. 1985), *aff'd in part and rev'd in part*, 772 F.2d. 1043 (2d Cir. 1985).

[137] For a discussion of the five alternatives discussed in the first EIS, see L. BACOW AND M. WHEELER, *supra* note 123, at 105.

[138] An interesting project to watch is Boston's third harbor tunnel/Depressed Central Artery. As politically controversial as Westway, this immense urban project has a large number of environmental problems as potential impediments to its progress. Indeed, myriad issues face the project, ranging from disposing of the unearthed dirt to controlling rats irked at the loss of their habitat. *See* Boston Herald, March 26, 1989, at A1, col. 3–4.

respond to its defeat and amend its procedures to prevent a recurrence of such a situation.

B. *Litigation Under Section 4(f)*

1. Highway Litigation in Section 4(f) Suits

Section 4(f) allows litigants to challenge a highway construction project's proposed use of a protected area, defined to include parkland and historic sites.[139] The FHWA must then defend its decision by demonstrating (1) that the agency followed all necessary procedures in assessing the environmental consequences of a project and alternatives to that project, and (2) that it made a substantive and affirmative determination that no feasible and prudent alternative exists.[140] Unlike their NEPA counterparts, successful section 4(f) suits can force the FHWA to reroute the highway, or, if no "feasible or prudent" alternative design exists, to cancel the highway's construction.[141]

These substantive remedies contrast sharply with NEPA's procedural relief, and their availability makes section 4(f) potentially a more lethal weapon than NEPA.[142] The judiciary's broad construction of section 4(f) requirements, partly evi-

---

[139] 49 U.S.C. § 1653(f) (protecting "any publicly owned land from a public park, recreation area, or wildlife and waterfowl refuge of national, state, or local significance as so determined by such officials, or historic site . . .").

[140] *See, e.g.*, Stop H-3 Ass'n v. Dole, 740 F.2d 1442 (9th Cir. 1984) (stringent requirements of section 4(f) were not met, in part, because the agency record did not establish that the Secretary could reasonably conclude that a no-build alternative must be rejected as imprudent); Ashwood Manor Civic Ass'n v. Dole, 619 F. Supp. 52, 74 (E.D. Pa.), *aff'd*, 779 F.2d 41 (3d Cir. 1985) (section 4(f) requires the decision-maker to determine that no alternative is feasible and prudent).

[141] *See* Citizens to Preserve Overton Park v. Brinegar, 494 F.2d 1212 (6th Cir. 1974), *cert. denied*, 421 U.S. 991 (1975) (challenge to the routing of I-40 resulted in Secretary of Transportation's agreement not to complete I-40 through Memphis). *See also* Coalition Against a Raised Expressway (CARE) v. Dole, 835 F.2d 803 (11th Cir. 1988) (challenging the completion of I-210 past the oldest American city hall in use, in Mobile, Alabama). Listed on the national register of historic places, the building received section 4(f) protection. *Id.* at 811. The FHWA reports they recently negotiated a settlement abandoning the plans for the raised expressway.

[142] *See* Stop H-3 Ass'n v. Dole, 740 F.2d at 1461 (9th Cir. 1984) ("While the mandate of section 4(f) essentially is prohibitory, the mandate of NEPA essentially is procedural"). Section 4(f) suits also may contain a procedural claim, *i.e.*, one that alleges improper or incomplete compliance with the statute's format for considering alternatives which do not use protected areas. 49 U.S.C. § 1653(a) (1982 and Supp. 1987). Such claims were bolstered in *Overton Park*, 401 U.S. 402 (1970), where the Supreme Court ordered lower courts to take a "hard look" at an agency's decision-making under section 4(f). For a discussion of the hard look doctrine, see S. BREYER & R. STEWART, *supra* note 103, at 331.

denced by its development of the constructive use doctrine,[143] also makes section 4(f) more potent. Faced with judicial sympathy toward strict compliance with section 4(f), the FHWA generally has chosen to accomodate, rather than to oppose, challenges under the statute. However, the pursuit of alternative construction plans in some urban areas has threatened so much community disruption that the FHWA has endured litigation despite its attempts at accommodation and compromise.[144]

2. *Stop H-3*

Opponents of an interstate highway in Hawaii recently won a considerable victory after lengthy section 4(f) litigation which targeted that expressway's constructive use of Hawaiian parkland.[145] The plan to build interstate highway H-3 across the island of Oahu, from Pearl Harbor Naval Base near Honolulu to Kaneohe Marine Air Corps Station on windward Oahu, raised storms of local protest.[146] The years of injunctions obtained by opponents were effective in halting the project, but, as discussed below, were issued for reasons peripheral to the real question of whether changing growth patterns on Oahu rendered the highway unnecessary.[147]

The plan for interstate highway H-3 originated from a 1973 study predicting a large population growth in windward Oahu,

---

[143] The FHWA has never won a constructive use case. FHWA ENVIRONMENTAL LAWSUIT STUDY, *supra* note 9, at 26.

[144] *See, e.g.*, Ashwood Manor Civic Ass'n v. Dole, 619 F. Supp. 52 (E.D. Pa.), *aff'd*, 779 F.2d 41 (3d Cir. 1985), where residents of Philadelphia's suburbs charged that completion of beltway I-476 around Philadelphia "used" parkland in the Crum and Ridley Creek basins. Following years of section 4(f) litigation, the highway—dubbed the "Blue Route" because of the choice from three possible routings, the blue, red, and yellow/green—was deemed a permissible use of parkland. The court's ruling rested on a finding that the other eleven possible corridor locations, while not "using" parkland, would cause "community disruption of extraordinary magnitude" by traversing heavily populated residential areas. Thus these alternatives were not "feasible and prudent" within the meaning of section 4(f). *Id.* at 76–78.

[145] Stop H-3 Ass'n v. Dole, 538 F. Supp. 149 (D. Haw. 1982), *aff'd in part and rev'd in part*, 740 F.2d 1442 (9th Cir. 1984).

[146] *See* 538 F. Supp. at 154.

[147] The two opinions cited here were merely the final installments of a 20-year struggle over H-3, issued in response to EIS and section 4(f) statements newly prepared in compliance with an earlier order from the same district judge. The origins and development of the dispute are discussed in Stop H-3 Ass'n v. Coleman, 389 F. Supp. 1102 (D. Haw. 1974); *see also* Stop H-3 Ass'n v. Volpe, 349 F. Supp. 1047 (D. Haw. 1972).

based on that island's 1969 General Plan.[148] H-3's corridor was tailored to facilitate commuting and to channel growth to a designated area.[149] From its inception, H-3 aroused disfavor from those opposed to the development of windward Oahu in general, and to H-3 in particular.[150]

While *Stop H-3* addressed numerous related claims, the district court ultimately decided to halt construction based on the highway's constructive use of Ho'omaluhia Park and Pali golf course.[151] Ho'omaluhia Park was originally conceived by the Army Corps of Engineers as a flood control project. The project was expanded between 1966 and 1970, from thirty-five acres surrounding a flood-control dam and reservoir, to seventy-five acres. In 1973, the proposed park was expanded to 115 acres (the "core"); then an additional 115-acre area was designated as a buffer zone between the highway right-of-way and the park's 115-acre core. The 1974 park master plan restricted all of the park's intensive uses, such as hiking, picnicking, and camping, to the core. The buffer zone would be used for park access, parking, trails, and open space.

While acknowledging the difficulty of quantifying H-3's constructive use of the core park, the district court reasoned that finding for opponents in close cases would prompt the FHWA to go further in mitigating the ecological impact of future projects.[152] The court relied on this rationale to supplement its purely hypothetical discussion of constructive use in this case.[153]

---

[148] A common problem with lengthy transportation litigation is that the projections and assumptions on which a highway's need is assessed change during the course of the lawsuit. While *H-3* was in court, Oahu's general layout and population figures had changed, with the relevant portion of Oahu no longer slated for rapid growth. *See* Stop H-3 Ass'n v. Dole, 740 F.2d at 1456 (citing Stop H-3 Ass'n v. Lewis, 538 F. Supp. at 166–67 (D. Haw. 1982)) (9th Cir. 1984).

[149] A much-debated question regarding highway construction is whether highways follow or spur development. If highways spur development, then corridor designations may unwittingly and improperly substitute for much larger economic and political policy choices. Whether a beltway through unimproved land proximate to a large city is constructed or not, may determine, for example, whether development is radial or symmetrical—an effect usually arrived at through political decision-making. *See* BELTWAY STUDY, *supra* note 12.

[150] *See* Stop H-3 Ass'n v. Lewis, 538 F. Supp. at 154–56 (D. Haw. 1982); Stop H-3 Ass'n v. Dole, 740 F.2d at 1461–63 (9th Cir. 1984).

[151] Stop H-3 Ass'n v. Lewis, 538 F. Supp. at 176 (D. Haw. 1982) (noting that issue of constructive use was not raised on appeal).

[152] *Id.* at 177.

[153] The court did not adduce evidence of constructive use in the instant case, but stated that

> [t]he degree or existence of constructive use depends upon both the nature of the park and the degree of impacts from the highway. For example, a park circulation road or parking lot would be much less affected by highway noise

In reviewing the district court decision, the Ninth Circuit tackled the larger issue of whether H-3 should be built at all. The court focused attention upon the substantive merits of H-3 by considering whether changed socioeconomic circumstances on Oahu made a no-build alternative feasible and prudent.[154] Section 4(f) review, however, is not supposed to address the merits of the highway, but only consider whether the FHWA could possibly have found a feasible and prudent alternative to the contested (constructive or actual) use of parkland.[155] The circuit court found that the Secretary failed adequately to consider downward revision of recent population projections and the corresponding shift in Oahu's growth focus from windward Oahu to other areas.[156] This failure contributed to the court's finding that the administrative record did not support the Secretary's decision that a no-build alternative was infeasible and imprudent.[157]

The court's analysis highlights the relationship between NEPA and section 4(f) claims. Section 4(f) prohibits park use unless no "feasible and prudent" alternative exists. If the NEPA EIS determines the no-build alternative to be a possible solution, then not building the highway is a feasible and prudent alternative to taking parkland.

H-3 opponents used section 4(f) to delay for at least a decade—and perhaps forever—the highway's completion. The Ninth Circuit reinstated the injunctions dissolved by the district

---

than would a picnic area or campsite. Mitigation measures such as landscaping and noise barriers may reduce the impact of a highway sufficiently to eliminate park "use" . . . . The fact that a park and a highway are jointly developed may be sufficient to establish that there are not feasible and prudent alternatives to the use of such park.

Id. at 176–77.

[154] Stop H-3 Ass'n v. Dole, 740 F.2d at 1461–62 (9th Cir. 1984).

[155] *See id.* at 1463 ("[O]ur role is not that of a 'superplanner,' (citation omitted) and, under NEPA, we are not allowed to substitute our judgment for that of the agency concerning the wisdom of a proposed action (citation omitted). Our role is limited to insuring that the appellees have taken a 'hard look' at H-3's environmental consequences"). *See also supra* notes 90–110 (discussing the standards for section 4(f) review).

However, the court noted that "[t]he mere fact that a 'need' for a highway has been 'established' does not prove that not to build the highway would be 'imprudent' under *Overton Park*. To the contrary, it must be shown that the implications of not building the highway pose an 'unusual situation,' are 'truly unusual factors,' or represent cost or community disruption reaching 'extraordinary magnitudes.'" *Id.*

[156] *Id.* at 1456.

[157] Stop H-3 Ass'n v. Dole, 740 F.2d at 1462–66 (9th Cir. 1984). The court discussed socioeconomic impact in relation to a NEPA claim included in the appeal. The FHWA did not appeal the constructive use ruling. *See* FHWA ENVIRONMENTAL LAWSUIT STUDY, *supra* note 9.

court until the FHWA fully complied with section 4(f)'s analytic regulations.[158] Thus the court allowed a relatively small opposition group to secure, through a somewhat insular legal forum, a veto over a major project affecting much of Oahu, with none of the broad public discourse desirable in public decisionmaking.

The outcome in *H-3* would have been far more satisfying had that debate centered around the island's revised estimates of growth and development rather than the constructive use doctrine. The central issue should have been whether ten years after the highway's proposal, the plans adequately addressed the region's current and future needs. But because litigation could not supplant public and political choices, the legal analysis did not satisfactorily evaluate Oahu's current needs.

The FHWA's decision not to appeal the constructive use ruling underscores the potency of the doctrine.[159] Ironically, had the buffer area been designated for another use that did not create additional parkland, section 4(f) might not have been able to stop the project.[160] The effects of the highway would have been distanced from the park, and would not have implicated 4(f)'s prohibitions. Moreover, had the section 4(f) statement focused on the safety concerns of the existing versus the proposed roads, then the court might have found that no feasible or prudent alternative existed to H-3.[161]

Both NEPA and section 4(f) provide highway opponents with potent weapons to delay or halt highway projects. Interpreting the statutes broadly, reviewing courts may cripple a project; and, as Westway demonstrated, the functional effect of delay may be to extinguish a project. Strict application of section 4(f) in H-3 halted construction of a highway which, at most, only marginally used parkland—itself created as a concomitant to the highway—in order to preclude constructive use of existing parkland. In neither case, under neither statute, were the substantive merits of the projects the focus of the court decisions. The relief

---

[158] Stop H-3 Ass'n v. Dole, 740 F.2d at 1447–55 (9th Cir. 1984). The circuit court also found that the projected community disruption attaching to the Makai Realignment, an alternative route which would have completely avoided the use of parkland, was insufficient to render it infeasible under *Overton Park*. *Id.* at 1452–54.

[159] *See* Stop H-3 Ass'n v. Lewis, 538 F. Supp. at 176–77 (D. Haw. 1982).

[160] Alternatively, H-3 could have created and constructively used the buffer in order to mitigate damage to the core park, had Congress granted H-3 a statutory exemption for section 4(f). FHWA ENVIRONMENTAL LAWSUIT STUDY, *supra* note 9, at 40.

[161] *See* Stop H-3 Ass'n v. Dole, 740 F.2d at 1456 (9th Cir. 1985).

sought and won under both statutes—remand of the issues to the agencies for further study—resulted in delay rather than public debate to promote a consensus regarding all of the benefits and detractions of the proposed construction.

IV. FHWA Responses to Litigation Under Environmental Statutes

A. *Administrative Change*

The FHWA has reacted to Westway, H-3, and similar court losses by increasing both its litigation effectiveness and its willingness to negotiate design concessions. This two-pronged response suggests a model for administrative change available to many other agencies.

Twenty years after the passage of NEPA, the FHWA achieves its desired outcomes in the large majority of highway challenges, filed under both this statute and section 4(f).[162] Since section 4(f) and NEPA suits challenge a highway's specific corridor placement and design features, the FHWA has focused its efforts on modifying its decision-making process to include the weighing of various substantive factors, and on presenting an adequate administrative record during litigation. This pragmatic approach encompasses a number of possible strategies shaped for the particular cases. Naturally, the agency follows the course most likely to result in the construction of the project.[163] At one extreme, the agency may confront the challenge and attempt to prevail in court. At the other, the agency may bow to opponents' pressure and cancel the project immediately. Since each challenge presents a unique set of circumstances, the FHWA tailors its behavior accordingly.[164]

Other administrative agencies similarly commissioned to construct projects which may implicate environmental concerns have likewise needed to adjust to the new climate of protection memorialized by Congress in NEPA and section 4(f); and many

---

[162] *See* FHWA Environmental Lawsuit Study, *supra* note 9, at v. For a complete discussion and statistical data on FHWA litigation, see *infra* notes 172–188 and accompanying text.

[163] *See* FHWA Environmental Lawsuit Study, *supra* note 9, at 37.

[164] *See generally id.* at 37–42.

have done so.[165] However, agencies might comply with NEPA's EIS requirement by producing documents which are facially acceptable but devoid of substance.[166] Of course, strident litigation might police the production of pro forma EIS's[167] if courts find that such studies comply with neither the spirit nor the letter of broadly construed environmental law. Some agencies have supplemented EIS drafting with changes in personnel (*i.e.*, political appointments),[168] institutional structure, and decision-making in order to incorporate environmental concerns, notwithstanding the lack of internal popularity surrounding these reforms.[169]

Institutional reforms which incorporate greater public participation in agency decision-making encourage administrative agencies to adhere strictly to environmental statutes by forcing them to interact with concerned citizens and gauge public reaction to a proposed project.[170] NEPA and DOT regulations require that the environmental impact assessment phase of project planning include public hearings, participation, and commentary.[171] This sensitivity to public reaction, insofar as it reflects a willingness to address citizen concerns substantively, helps to ensure a better project and reduce the risk of lengthy litigation.

B. *Development of Litigation Strategies*

The physically intrusive nature of its operations—constructing highway projects—means that the FHWA is frequently taken

---

[165] *See* R. LIROFF, A NATIONAL POLICY FOR THE ENVIRONMENT: NEPA AND ITS AFTERMATH 139–42 (1976). *See also* S. TAYLOR, MAKING BUREAUCRACIES THINK: THE ENVIRONMENTAL IMPACT STATEMENT STRATEGY OF ADMINISTRATIVE REFORM 130–33 (1984); R. ANDERSON, ENVIRONMENTAL POLICY AND ADMINISTRATIVE CHANGE 89–91 (1976). In particular, the Army Corps of Engineers provides an interesting example of an agency forced to incorporate environmental values into its operations. *See* S. TAYLOR, *supra* at 41–43.

[166] *See* R. LIROFF, *supra* note 165, at 137–38.

[167] For a brief discussion of administrative reaction to strict enforcement of NEPA and section 4(f) obligations, see 4 H. INGRAM & R. GODWIN, PUBLIC POLICY AND THE NATURAL ENVIRONMENT 206–08 (1985).

[168] *Id.*

[169] Even the reluctant Army Corps of Engineers established environmental review units and made personnel changes in acceptance of its Congressionally delegated environmental responsibilities. D. MAZMANIAN & J. NIENABER, CAN ORGANIZATIONS CHANGE? 58–60 (1979). Experiences such as Westway indicate that the Corps and its fellow agencies will also take a more sober look at the EIS requirement than they have in the past.

[170] *Id.* at 160–66 (discussing Army Corps of Engineers' response to public hearings).

[171] *See* 23 C.F.R. § 771.111 (1988).

to court,[172] primarily under NEPA, section 4(f), or both.[173] Fewer new projects have begun since 1970, yet the total number of suits filed has remained stable, which suggests that these new projects are more contentious than their predecessors.[174] A small percentage of suits has resulted in injunctions halting construction.[175] Clearly while the FHWA has emerged a more formidable adversary from years of litigation experience, highway opponents have also learned a lesson or two, and are now litigating only the most controversial projects—those that hold the greatest promise for a successful challenge.[176]

Agencies have increasingly attempted to chronicle their decision-making in detailed administrative records, to facilitate judicial examination (and approval) of those processes.[177] Yet the FHWA has determined that compiling a judicially sufficient record would exhaust too large a percentage of administrative resources, and has thus attempted to conserve its energies in this regard for those projects over which it strongly anticipates community opposition and controversy.[178] In the absence of the

[172] Since 1974, when a spurt of NEPA cases was filed following the passage of the law, an average of 17 new cases have been filed annually against the FHWA. This number has remained fairly constant over 15 years, despite a steady drop in the number of environmental impact statements prepared during the 1970's. The number of highway-related EIS's prepared levelled throughout the 1980's. *See* FHWA ENVIRONMENTAL LAWSUIT STUDY, *supra* note 9, at 8. *See also* Liroff, *NEPA Litigation in the 1970's: A Deluge or a Dribble?* 21 NAT. RES. J. 315 (1981).

[173] Forty cases were analyzed over a three-year period. Of this group, 39 contained either a section 4(f) or NEPA claim, or both. Fifteen of the 39, or 38.5%, contained both NEPA and section 4(f) claims; 14, or 35.9%, featured only NEPA; and 10 (25.6%) only section 4(f). FHWA ENVIRONMENTAL LAWSUIT STUDY, *supra* note 9, at 9. While all of the cases contained other issues as well, the NEPA or section 4(f) claim generally was dispositive. *Id.* at 14.

[174] *Id.* at 8.

[175] Of 136 total cases filed or active against the FHWA between 1980 and 1985, injunctions were issued in only 11, or 8%. *Id.* at 16. Temporary restraining orders or preliminary injunctions were obtained, however, in one-fourth of the 40 cases examined in the Lawsuit Study. *Id.*

[176] *See* CBO STUDY, *supra* note 48 (describing the urban location and controversial nature of currently unfinished portions of highway).

[177] While not conclusively identifying the advantage such efforts may yield at trial, statistics tend to show that this strategy may work to the agency's benefit. In 15 of 27 cases studied in which the preparation of the administrative record was at issue, the FHWA had anticipated litigation at the project's inception and, presumably, had purposefully compiled a detailed record. The agency prevailed eight times and lost seven. In those 12 suits not anticipated by the FHWA, and for which a detailed record was presumably lacking, the agency prevailed four times and lost eight. FHWA ENVIRONMENTAL LAWSUIT STUDY, *supra* note 9, at 45–46. (It should be noted that these numbers reflect court rulings on the sufficiency of the environmental documents; *i.e.*, a court may order the FHWA to prepare a SEIS or section 4(f) statement but allow construction currently underway to continue. The FHWA considers such an outcome a victory).

[178] *Id.* at 48–49. The FHWA requests that state DOTs, which conduct much of the initial locational analysis, compile the minimum necessary number of administrative documents of a quality suitable for use in litigation. *Id.*

usually clear warnings opponents send when they do not approve of a plan, the agency has preferred to allocate resources to other tasks rather than compile an exhaustive record for every project.[179] Often, the record must be created post hoc, but even records initially deemed inadequate by the courts may be supplemented and emerge victorious on appeal.[180]

In response to judicial decisions, the FHWA has altered a number of its policies and procedures under NEPA, to encompass more frequent consideration of routing or construction alternatives, and has heightened its attention under section 4(f) analysis to environmental impact mitigation and avoidance of parkland usage.[181] The FHWA has tried to conform that analysis to the judicial parameters of constructive use doctrine, but the current expansive standard is an elusive one for the agency to incorporate into its decision-making.[182]

The FHWA cites the anticipation of a lawsuit and support from other governmental agencies during that suit as the two factors most conducive to successful litigation over a project, the latter being somewhat more determinative.[183] A united interagency position seems critical to the success of environmentally controversial projects.[184] Opposition from other govern-

---

[179] The FHWA in effect balances the possible need to create a post hoc record, which would vitiate the initial time and resource savings, against the potential of a negative judgment attributable to a faulty or nonexistent record. *Id.* at 49.

[180] Insufficient administrative records may lead to losses in the first round of litigation. However, projects can drag on for years, during which time the FHWA may review its original decision, including its environmental impact analysis, in order to survive further judicial scrutiny. *See, e.g.*, Ashwood Manor Civic Ass'n v. Dole, 779 F.2d 41 (3d Cir. 1985); *contra* Sierra Club v. U.S. Army Corps of Eng's, 772 F.2d 1043 (2d Cir. 1985).

[181] Most commonly, the agency merely pursues the most cost-effective procedure, whether that be preparing an EIS, supplementing an existing EIS, or conducting further study under NEPA. *See* FHWA ENVIRONMENTAL LAWSUIT STUDY, *supra* note 9, at 37. Strict judicial review of section 4(f) compliance demands more substantive adaptation. *Id.* at 38–40.

[182] Present doctrine recognizes a highway's visual impact and proximity to a protected resource as potential constructive uses. Current FHWA policy, adopted following Adler v. Lewis, 675 F.2d 1085 (9th Cir. 1982), incorporates these extensions of section 4(f) protection. *Id.* Subsequent decisions, such as *Stop H-3*, and Citizen Advocates for Responsible Expansion, Inc. (I-CARE) v. Dole, 586 F. Supp. 1238 (N.D. Tex. 1984), *rev'd* 770 F.2d 423 (5th Cir. 1985), and most recently in Coalition Against a Raised Expressway, Inc. (CARE) v. Dole, 835 F.2d 803 (11th Cir. 1988) highlight the FHWA's inability to respond effectively to constructive use doctrine application.

[183] FHWA ENVIRONMENTAL LAWSUIT STUDY, *supra* note 9, at 47–48. With advance knowledge of litigation and other agencies' neutrality, the FHWA won seven and lost no suits. With advance knowledge and another resource agency's opposition, the FHWA won one and lost seven suits. *Id.*

[184] Governmental agencies involved with the environment or natural resources often have a stake in highway construction. The Environmental Protection Agency, National Park Service, and National Marine Fisheries Service (whose criticism of the Westway

mental resource agencies may vitiate in the court's eyes even a detailed FHWA record of responsible decision-making.[185] A divided governmental front bolsters opponents' arguments that the project is not the most feasible or prudent alternative.

The FHWA has promulgated regulations which adopt *Overton Park*'s standard for determining what constitutes an infeasible alternative.[186] The FHWA now employs a balancing test that weighs the benefits and drawbacks of using a protected resource.[187] However, changing judicial standards for section 4(f) review have increased the difficulty of formulating a coherent policy and the FHWA's reluctance to pursue worthwhile projects.[188] In particular, the significant judicial expansion of constructive use doctrine may dissuade the FHWA from considering or pursuing highway extension and repair programs needed in urban areas.

1. Case Study: The Blue Route

The FHWA's flexibility in modifying compliance procedures and litigation strategy allowed it to persevere and emerge victorious in the struggle to build a partial beltway through the western suburbs of Philadelphia.[189] Nicknamed the Blue Route, this twenty-one-mile road connected I-95 in Chester, Pennsylvania with the confluence of the Pennsylvania Turnpike (I-276) and Pennsylvania Turnpike Northeast Extension in Plymouth Meeting, Pennsylvania. It traversed a corridor through Philadelphia's wealthy western suburbs, a routing which inflamed the area's residents who organized to sue and stop construction of the project. The district court granted an injunction which halted construction pending additional environmental review.[190] Fol-

---

EIS sealed the project's fate), among many others, thus intercede frequently in highway construction disputes. *Id.*

[185] *See id.*

[186] *Overton Park* reiterated section 4(f)'s requirement that parkland only be used if the alternatives were infeasible or imprudent, *i.e.*, if community disruption or costs are of an extraordinary magnitude. *See* 23 C.F.R. § 771.135(a)(2) (1988).

[187] FHWA ENVIRONMENTAL LAWSUIT STUDY, *supra* note 9, at 41 (the test weighs "the importance of the Section 4(f) resource, the harm which the project will cause to the resource, and the magnitude of efforts to minimize harm to the resource").

[188] *Id.*

[189] The Blue Route was initially planned in 1929 and funded in 1959. Construction began in 1967. Halted by injunction in 1971, the unfinished highway awaited judicial sanction for 15 years. *See* BLUE ROUTE EIS, *supra* note 12, at I-6.

[190] For a complete procedural history of the Blue Route litigation, see Marple Township v. Lewis, Nos. 81-4627 and 74-925 (E.D. Pa. 1982).

lowing intensive review, the district court ratified the FHWA decision.[191] That decision and the district court's review were ultimately affirmed on appeal.[192]

At issue had been the highway's routing through creek basins, town parks, historic houses, and the Swarthmore College campus arboretum.[193] Much of the final litigation therefore involved analysis of alternative routes for the highway to determine if a more feasible or prudent alternative was available. Earlier lawsuits had prompted the FHWA to conduct in-depth study of substitute corridors.[194] The agency concluded that alternatives which minimized protected site usage either were too far from Philadelphia to meet the corridor's traffic needs, or would cause inordinate community disruption in a densely populated suburban region.[195] Such circumstances, carefully documented by the agency,[196] led the court to find that the Secretary's determination that no feasible or prudent alternatives existed was not arbitrary or capricious.[197]

During the Blue Route litigation, the FHWA introduced the concept of cumulative impacts, which argues that while an alternative may have many small detractions, any one of which would not make it infeasible or imprudent, cumulatively they render an alternative unacceptable.[198] This creative litigation strategy contributed to the ultimate success of the project.[199]

In addition to presenting a careful and complete administrative record, the FHWA also conducted a thorough review of the project and devised extensive environmental mitigation features to lessen the highway's intrusion on the region.[200] Indeed,

---

[191] Ashwood Manor Civic Ass'n v. Dole, 619 F. Supp. 52 (E.D. Pa. 1985).

[192] 779 F.2d 41 (3rd Cir. 1985), *cert. denied*, 475 U.S. 1082 (1986).

[193] Ashwood Manor Civic Ass'n v. Dole, 619 F. Supp. at 81 (E.D. Pa. 1985).

[194] *Id.* at 63–82.

[195] *Id.* at 76–78.

[196] Most of the text of the opinion is dedicated to evaluating the administrative record which determined that no feasible or prudent alternative existed to using protected resources. *See, e.g.*, 619 F. Supp. at 53. The court found the record presented by the FHWA to be "extensive and complete." *Id.* at 72.

[197] *Id.* at 86.

[198] *Id.* at 79. The court cited two additional reasons for the dearth of feasible alternatives: first, that regional land use planning had been conducted in anticipation of the Blue Route's construction, and thus committed potential areas to other uses; and second, since the Blue Route was chosen before the passage of section 4(f), the court cited a reluctance to "disrupt established land patterns." *Id.*

[199] *See id.* at 87.

[200] The altered Blue Route included two route shifts to viaducts to lessen impacts on streams, reduced the highway from six lanes to four for seven miles, narrowed the median strip, eliminated some interchanges and downscaled those remaining, and added extensive visual and noise abatement measures along the length of the highway. *Id.* at

without agreeing to sweeping and expensive environmental mitigation through negotiation, the project probably would never have been judicially approved.[201] Even with the most strenuous and complete administrative documentation, a persistent and committed litigant may extract significant design concessions from the FHWA before a court will ratify the project.[202]

The administrative and judicial review of the decision to build the Blue Route approximated the model for improving public choice discussed in Part I. The parties and the court engaged in a considerable amount of substantive review of the plan's merits. The FHWA commissioned a new location study which thoroughly examined the need for the project.[203] The court cited its obligation to "conduct a thorough, probing, in-depth review of the substantive decision."[204] Litigation forced all parties to consider the long-term transportation needs of the region, and also to address realistically the dislocation which the corridor would produce.

But the outcome of the Blue Route litigation also suggests the limitations of litigation for highway opponents. The court could not choose between two alternatives, but only could ensure that the FHWA had proper justification to opt for the "build" alternative. Litigation did not substitute for public debate on the merits of the project.[205] In the Blue Route situation, the procedural focus of NEPA and section 4(f) properly constrained the court from tackling the ultimate issue, but without providing for public debate in a proper forum.

C. *Negotiating Design Concessions*

The FHWA's growing success in litigating highway challenges should not obscure the frequency with which the agency settles

---

80–81. Such measures came at a price, however; the completed highway will cost $470 million in 1986 dollars. BLUE ROUTE EIS, *supra* note 12, at IV-113.

[201] The district court emphasized the agency's extensive mitigation efforts and "rigorous scrutiny" of the Blue Route's environmental impact. 619 F. Supp. at 86–87.

[202] Regulations promulgated under NEPA and section 4(f) require the FHWA to engage in mitigation efforts for all projects which negatively affect the environment. *See* 23 C.F.R. § 771.105 (1988); *see also* 23 C.F.R. §§ 752, 770, 777 (1988).

[203] Ashwood Manor Civic Ass'n v. Dole, 619 F. Supp. at 74–84 (E.D. Pa. 1985). In addition to studying alternative corridors, the corridor analysis also exhaustively studied the no-build alternative, in effect reassessing the need for the project.

[204] *Id.* at 73.

[205] This instance further highlights how the FHWA, armed with an effective litigation strategy and a complete administrative record, was able to prevail over a determined opponent.

with opponents in order to avoid suffering the enormous cost increases or loss of political support which litigation entails. Opponents may thus threaten or actually pursue litigation in order to extract design concessions from the FHWA and make the proposed project more environmentally palatable.[206]

The public hearing provisions of the EIS process require that local concerns and considerations be heard, considered, and integrated into the project's design.[207] To ignore public commentary would violate NEPA's strictures and leave the project vulnerable to litigation on the grounds that the process failed to comply fully with regulatory mandates.[208] Entertaining public debate and integrating into the project suggestions raised during that debate may defuse community opposition and preempt litigation. Furthermore, such measures also manifest administrative concern and responsible decision-making. Then, if litigation still cannot yet be averted, these factors are likely to persuade a court to view the agency decision favorably.[209]

Parties to highway litigation have occasionally found significant room for negotiation over environmental mitigation, and have thus been able to accede to a project's continuation.[210] The cost of a lengthy lawsuit provides incentives for both sides to settle as early as possible. Plaintiffs realize that most projects are eventually built,[211] and the FHWA seeks to contain rising construction costs and avoid the prospect of losing the project.[212] While the extent of environmental mitigation measures

---

[206] *See, e.g.*, Ashwood Manor Civic Ass'n v. Dole, 779 F.2d 41 (3d Cir. 1985). This lawsuit resulted in FHWA's voluntary and significant reduction of the scope of the Blue Route.

[207] *See* 23 C.F.R. § 790 (1988) (requiring public hearings and coordination of state, local, and federal planning agency objectives in choosing a highway corridor).

[208] *See, e.g.*, Sierra Club v. U.S. Army Corps of Eng'rs, 772 F.2d 1043 (2d Cir. 1985) (involving the claim that the Corps of Engineers failed "adequately . . . [to] consult and give full consideration to the views of federal fishery agencies" in evaluating the validity of the striped bass spawning study).

[209] *See, e.g.*, Ashwood Manor Civic Ass'n v. Dole, 619 F. Supp. 52 (E.D. Pa. 1985) (existence of mitigation measures supported agency contention that the decision to use parkland was not arbitrary or capricious).

[210] FHWA litigation statistics indicate that of 43 cases closed during the period between 1985–88, five were settled prior to a judicial decision. Federal Highway Administration, Right-of-Way Division, Closed Right-of-Way Case Summaries (unpublished internal Memorandum of FHWA) (1988).

[211] *See, e.g.*, Neighborhood Preservation Coalition v. Claytor, 553 F. Supp. 919 (E.D. Pa. 1982) (FHWA and neighborhood organization negotiated design concessions and mitigation measures to settle a longstanding dispute over the configuration of exit ramps off I-95 in downtown Philadelphia. I-95 was constructed pending final resolution of this dispute, which arose out of a 1973 NEPA suit).

[212] *See, e.g.*, Liquid Air Corp. v. Dole, No. 86-118 (D. Colo. 1986). Plaintiff Liquid Air Corporation sued under NEPA and section 4(f) to enjoin construction of two

incorporated into a judicially enforced settlement varies, settlement at least allows the project to be constructed in some form.[213]

1. Case Study: Century Freeway

Century Freeway (I-105), a 17.2-mile urban freeway, runs from Los Angeles International Airport due east to connect with the San Gabriel River Freeway (I-605) in Norwalk, California, south of downtown Los Angeles. Protracted controversy and a ten-year injunction[214] led the FHWA and the California Department of Transportation (CalTrans) to incorporate significant design modifications into I-105 and settle with the project's opponents.

The settlement agreement allowed construction of the freeway, but committed the FHWA and CalTrans (1) to modify the project to accommodate mass transit by creating separate high occupancy vehicle (HOV) lanes, and bus or light rail transit stations at interchanges;[215] and (2) to provide for the construction or rehabilitation of over 2000 housing units for those in the right-of-way.[216]

Opponents of the Century Freeway scripted the settlement so that the mitigation measures targeted those most likely to be injured by the highway. Those displaced by the corridor, and

---

viaducts in Denver which would pass near plaintiff's cooling towers. Plaintiff argued that water vapor from towers could cause significant traffic hazards of fogging and icing on the viaducts. Parties entered into a settlement providing for a safety study and mitigation, but the court will issue a temporary restraining order on the opening of the viaducts if the plaintiff is not satisfied with safety report and mitigation measures.

[213] *See, e.g.*, Friends of the Park v. Dole (II), No. 87 C 7991 (N.D. Ill. 1987), where the plaintiffs had obtained a temporary restraining order halting work on an exit ramp configuration of Lake Shore Drive in Chicago. The FHWA entered into an agreement with plaintiffs, a nonprofit residents' association, to widen and otherwise alter the project. The FHWA agreed to reduce the length of some entrance and exit ramps, monitor accident rates, provide warning signs, and interact with the City of Chicago and Chicago Park District.

[214] *See* Amended Consent Decree at 5, Keith v. Volpe, (No. 72-355-HP) (C.D. Cal.) (1981).

[215] *See id.* at 5–11.

[216] Under the consent decree's housing provisions, first, California would rehabilitate or construct 1025 units for eligible renters or owners. Second, the state would construct or rehabilitate 1175 units for residents eligible under the Relocation Act. Third, the Federal DOT would allocate $110 million toward these state efforts, subject to a federal efficiency audit. Amended Consent Decree, *supra* note 214, at Amended Exhibit B, 1–4. The consent decree outlines in specific detail federal, state, and local participation in the project, as well as eligibility requirements and the housing production schedule. This plan is judicially enforceable as part of the consent decree. Amended Consent Decree at 4–5.

low-income individuals, were the primary beneficiaries of newly-created housing units.[217] Although Los Angeles had relatively little formal mass transit, the settlement provided for transit stations and reduced the cost of mass transportation to area residents through the HOV lane provision. These provisions addressed needs which in part anteceded the highway's effects and were in part exacerbated by the freeway. Opponents used the settlement to subsidize other regional goals, such as low-income housing and mass transit, by linking them to highway construction.

The negotiated settlement in the Century Freeway dispute addressed needs that each side felt were compelling. The interested parties ultimately engaged in an interactive dialogue, which is a better model for making public choices than the previous litigation-driven process. The settlement of the Century Freeway dispute attempted to strike a balance between the needs of local residents directly harmed by the construction and the transportation demands of the Los Angeles metropolitan area. The parties used the process to serve regional goals, but at a high national price. Traditional frameworks, such as economic analysis, provide unsatisfying results when they are used to balance the transportation requirements of the region against the additional costs of the mitigation.[218]

The negotiated settlement (1) needed the affirmation of numerous federal and state agencies in addition to judicial approval and (2) was reached in a public forum. Both of these aspects allowed debate on the merits of the project. Within the negotiation, political leaders supporting and opposing the project were forced to exchange and reassess their views on the future transportation needs of Los Angeles. Hunger for federal funds may have contributed to their willingness to compromise in order to save the project, but proponents and opponents nonetheless thought long and hard about the role Century Freeway would play in Los Angeles. Those parties in the freeway's right-of-way may have been more fully compensated by the settlement than were the Blue Route litigants, indicating that the former

---

[217] *See* Amended Consent Decree, *supra* note 214, at Amended Exhibit B, 17–24.

[218] *See* H. MOHRING, *supra* note 13. *See generally* AMERICAN ASSOCIATION OF STATE HIGHWAY AND TRANSPORTATION OFFICIALS (AASHTO), MANUAL ON USER BENEFIT ANALYSIS OF HIGHWAY AND BUS-TRANSIT IMPROVEMENTS (1977); UNITED STATES DEPARTMENT OF TRANSPORTATION, EVALUATING URBAN TRANSPORTATION SYSTEMS (1978).

either participated in a more responsive process or had claims more likely to halt the project completely.[219]

V. Environmental Mediation as an Alternative for Public Choice

Alternative dispute resolution (ADR) may provide a framework for resolving highway disputes, as well as a forum for relevant parties to produce a representative and cooperative solution.[220] Environmental disputes have recently been settled using the techniques of mediation and negotiation, as opposed to litigating the disputes to judicial conclusions.[221] Although ADR has had success in resolving highway disputes, it is not commonly attempted.[222]

Alternative dispute resolution techniques work best under certain circumstances which are not always present in complex highway disputes. Mediation and negotiation tend to succeed in disputes which involve small geographic areas, small and well-defined groups of disputants, micromanageable as opposed to broad policy issues, discernible opportunities for compromise, and a genuine desire to reach agreement.[223] In addition, all parties must trust the mediator or negotiator.[224]

Alternative dispute resolution should be used only in the course of considering the substantive merits of the proposal and in a context where its use would produce better public choices.

---

[219] Century Freeway was designed to pass through poorer areas of Los Angeles, where substitute low- and moderate-income housing was unlikely to have been readily available. *See* Amended Consent Decree, *supra* note 214, at 3. In contrast, most residents displaced by the Blue Route were upper-income suburban residents, who could have purchased acceptable substitute housing elsewhere in the region with relative ease. Federal Highway Administration and Pennsylvania Department of Transportation, MidCounty Expressway: Socioeconomic and Land Use Basis Report I10–I15 (1977).

[220] *See* P. Wald, *Negotiation of Environmental Disputes: A New Role for Courts?* 10 Col. J. Env. L. 1 (1985) (discussing the role of ADR in resolving environmental disputes vis-à-vis traditional judicial review of agency actions). *See also* L. Bacow & M. Wheeler, *supra* note 123, at 359; H. Ingram & R. Godwin, *supra* note 167, at 317–22.

[221] For a catalogue of the applications of environmental mediation, see G. Bingham, Resolving Environmental Disputes 169–256 (1986).

[222] Mediation or negotiation was employed to settle transportation disputes concerning the Alewife Brook Subway Station in Cambridge, Massachusetts, *id.* at 175; a segment of the Illinois Tollway, *id.* at 211; and Interstate 90 across Lake Washington near Seattle, Washington, *id.* at 212–13.

[223] *See* D. Amy, The Politics of Environmental Mediation 215–16 (1987).

[224] *Id.* at 215.

Proponents of ADR cite reduced cost, increased speed and efficiency, and party satisfaction as the benefits of non-litigation solutions.[225] But for a negotiated or mediated settlement to become a preferred alternative to environmental litigation, it should also produce a more coherent societal consensus on a project and stimulate public debate.

A. *Case Study: I-90 Across Lake Washington*

The construction of large-scale highway bridges across Lake Washington to connect Seattle, Washington's eastern suburbs with the downtown area ignited heated protest from that region's environmentally conscious populace. Citizen opposition to the project focused on urban sprawl resulting from opening up Seattle's eastern suburbs, and the noise, air pollution, and visual impacts of large bridges across the lake and an underlying fear that Seattle would become a northwestern Los Angeles.[226] Many opponents favored transportation planning which emphasized mass transit, and advocated trading-in the bridge funds for such funds.[227] The original plan, approved in 1960, faced many revisions and court challenges. Yet the dispute raged over whether the road should be built and the specific design of the highway.[228]

With the highway's design in debate and its future in turmoil, the Governor of Washington interceded and appointed a mediator specifically to reach a consensus on highway design, rather than debate the ultimate question of its construction.[229] Because the mediation was designed to resolve the governmental schism

---

[225] *Id.* at 18–23. *Accord* L. BACOW & M. WHEELER, *supra* note 123, at 18–20; *but see* G. BINGHAM, *supra* note 221, at 127–46 (finding that statistics comparing expenses and duration of litigation with ADR are inconclusive).

[226] *See* Cormick & Patton, *Environmental Mediation: Defining the Process Through Experience*, in ENVIRONMENTAL MEDIATION: THE SEARCH FOR CONSENSUS 92 (L. Lake ed. 1956). Cormick and Patton were the mediators appointed by Washington Governor Dan Evans to mediate the dispute. *See* A. TALBOT, SETTLING THINGS 31 (1983).

[227] *See* TALBOT, *supra* note 226, at 28–30.

[228] The plans for I-90 to cross Lake Washington originally began in 1960 as a proposal to build 26 bridge lanes of traffic across the water. By 1976, protest and court challenges under NEPA and section 4(f) had delayed the project, driven costs up to $140,000 a day, and led the Washington State Department of Highways to support a revised project of 10 lanes: four automobile lanes in each direction, and two reversible lanes for mass transit (bus, van, or carpool), known as the "4-2t-4" configuration. The King County City Council also offered, unsuccessfully, a compromise proposal providing for three initial auto lanes expandible to four. *See* Cormick & Patton, *supra* note 226, at 90–93.

[229] *See* TALBOT, *supra* note 226, at 33. This official drive for an acceptable solution stemmed from the desire of local politicians to obtain this infusion of federal money into the region. *See* Cormick & Patton, *supra* note 226, at 93.

over the project's proper design, the parties to the process were limited to elected officials from the affected jurisdictions and representatives from governmental transportation and planning agencies. Environmental and citizen groups, though physically absent from the bargaining table, participated indirectly through their link with the mediators, who raised the concerns of the environmental and citizen groups to the elected and political officials.[230] In addition, through the ballot box voters held the ultimate veto over the project, since they could unseat local officials who supported the road.

Within nine months, the mediation effort, enjoying strong gubernatorial support, produced a hard-fought compromise which incorporated divergent local demands and significant environmental design concessions.[231] While all major political and administrative entities supported the plan, citizen opponents continued to challenge the adequacy of the EIS and 4(f) analyses and conclusions. However, the Ninth Circuit upheld the Transportation Secretary's determinations.[232]

The mediated solution in the I-90 dispute led to a lengthy evaluation of the project by officials at all levels and in all areas of government. The process advanced the assessment of the project's scope, purpose, and role in shaping the future of the Seattle metropolitan area. Much of the debate consisted of officials discussing which version of the road would assure the large infusion of federal funds to the region. But negotiations on the size of the highway, the role of mass transit, and environmental impact of the highway must have forced officials to consider the future parameters of the metropolitan region. Such a stage is a step towards a more sensible model of public choice.

The threat of litigation was important to this mediation process, since the highway agencies would otherwise have had little incentive to negotiate expensive mitigation measures. The purpose of a mediated settlement would be to avoid the delay and expense of litigation, and to arrive instead at a solution which

---

[230] *See id.* at 94–96.

[231] The final project involved a configuration of three lanes for each direction and two reversible lanes ("3-2t-3"), with special access on the transit lanes for local residents of Mercer Island, an eastern suburb. In addition, mass transit linkage between Seattle and Bellevue was extended to other Seattle highways, and a significant portion of the lanes was "lidded-over" to mitigate visual and environmental impact. The agreement also provided for the appointment of joint oversight committees, composed of citizen groups and elected officials. *See* Cormick & Patton, *supra* note 226, at 95–96.

[232] *See* Talbot, supra note 226, at 38 (discussing Adler v. Lewis, 675 F.2d 1085, 1089 (9th Cir. 1982)).

fulfills the criteria of the highway choice model. The FHWA would seek through mediation to avoid the expenses of litigation and delay, and highway opponents might gain some added concessions through publicity and political pressure. While mediation in the shadow of litigation would seem to defeat the purpose and be antithetical to the philosophy of such a settlement, without that incentive, perhaps no meaningful discussion would result.

Mediation, while not a complete answer to settling highway disputes, promotes an active discussion of the substantive needs and merits of a project and its role in the future of the metropolitan region. Certain features of the I-90 dispute which rendered mediation successful may not be present in all highway disputes. Mediation may have succeeded in this instance because the governor limited the participants to government officials who had the common goal of preserving federal money for the region. In other instances, the problems of determining which groups should participate and reconciling the divergent aims of committed opponents may make mediation ineffective.[233] Limiting participation may encourage excluded groups to challenge the project in court and thus result in a process of delay and polarization.

CONCLUSION

The social and economic consequences of highways on urban, suburban, and exurban areas profoundly influence metropolitan regions. The process to determine whether to construct public projects with such extensive repercussions should involve community and political discussion about the project and its impact on the future of the area. However, highway construction under current environmental laws encourages a confrontation between government and the opponents of the project. A lack of community consensus on transportation planning and its role in the growth and development of metropolitan regions leads highway opponents to litigate in federal court under environmental laws, most often NEPA or section 4(f). Litigants seek to exercise a

[233] Indeed, officials attempted to mediate the Westway impasse in 1974, before immutable battle lines were drawn. Citizen groups' diametric opposition to any new construction proved to be one of several too-formidable obstacles to agreement. The city's preference for trading-in funds for use in expanding mass transit was another. *See* L. BACOW & M. WHEELER, *supra* note 123, at 104–08.

neighborhood veto and procedurally delay a project, with an eye toward eroding a project's political support or substantively preventing the project's use of protected resources.

Over the past twenty years, the FHWA has countered the ever-increasing sophistication of highway opponents with altered administrative procedures and litigation strategy or with concessions in highway design. Agency reaction to highway litigation has been to produce carefully scripted and documented administrative records and litigation postures that usually allow a project to proceed in some form. Committed and well-financed highway opponents may extract design concessions, usually tailored to soften the intrusive nature of highways or to avoid using a protected resource.

While years of action and reaction in the end permit most highway construction, the current situation fails to promote a public choice based on community and political debate and resolution. Instead, final decisions are based on administrative maneuvering and procedural accuracy. And while the favor of political figures for projects may be one indication of public support, that favor may also reflect a thirst for an infusion of federal funds into the region—a competing public demand.

Current environmental law and FHWA regulations require public hearings on projects large enough to require environmental impact statements, but the EIS/section 4(f) statement process does not adequately prompt policymakers to incorporate public considerations into planning. The regulations promulgated under NEPA and section 4(f) should be supplemented with regulations requiring mediation to resolve disputes and to promote productive dialogue over the project.

Requiring local, regional, and national political entities to mediate a dispute with input from opposed groups may result in resolution of a difficult regional planning issue and, in rare instances, garner consensus for the project. At the very least, a design formulated by mediation may indicate to a court that the plan received consideration from affected parties and reflects a rough consensus of public support for the plan. In section 4(f) suits, the outcome of mediation may demonstrate to the court a region's willingness to sacrifice some use of parkland for the project.

Mandating a forum that encourages active discussions of the role of transportation planning in the future of a metropolitan region should elevate the public decision-making process.

Should mediation prove unsuccessful, litigation remains an alternative; however, the threat of a lengthy suit will provide incentives to both sides to negotiate. Involving political leaders will require them to formulate postures on the project and face the nature of such projects.

No statutory innovation in relevant rules will quell the vehement opposition of some groups or individuals to highway construction. Highway projects by nature impose significant social burdens on those in the right-of-way, taking their homes and eviscerating neighborhoods in the path. Highways promote automotive travel and its concomitant social costs, and exacerbate sprawl and decentralization. But highways link communities and reduce transit costs, and are inextricably integrated into the social fabric. A choice to build or expand the highway network has significant social consequences, which should be evaluated by public debate rather than administrative maneuvering.

RECENT PUBLICATIONS

CALL TO ORDER: FLOOR POLITICS IN THE HOUSE AND SENATE. By *Steven S. Smith.* Washington, D.C.: The Brookings Institution, 1989. Pp. xvi, 252, appendices, index. $31.95 cloth; $11.95 paper.

For most of the twentieth century, scholarly work on Congress has focused heavily on the importance of committees in determining policy outcomes. This pervasive emphasis on committees can be traced to the prominence of committees in the work of such "giants" of congressional scholarship as Woodrow Wilson, Donald Matthews, and Richard Fenno.[1] By contrast, little attention has been paid to the politics of amending, debating, and voting on bills once they come to the floor of the House or Senate. Instead, the floor has been treated as a "showcase" where much is said, but little of substance is done. To the extent floor activities have been studied at all, inquiry has focused on the coalitional patterns in final passage votes and on events, such as filibusters, which are dramatic enough to capture public attention.

In recent years, however, there has been growing attention among congressional scholars to the intricacies of floor procedures.[2] This increasing attention results from changes within Congress that have altered the context of floor deliberations, and the growing intellectual influence in the field of rational choice models that focus on the influence of institutional procedures on policy outcomes. Steven Smith's *Call to Order* both reflects and extends this new line of inquiry by documenting changes in congressional floor politics from the 1950's to the 1980's.

Smith's central argument in *Call to Order* is that the floors of the House and the Senate have been "transformed into far more important arenas of substantive policymaking" than they were

---

[1] W. WILSON, CONGRESSIONAL GOVERNMENT (1885); D. MATTHEWS, U.S. SENATORS AND THEIR WORLD (1960); R. FENNO, CONGRESSMEN IN COMMITTEES (1973).

[2] *See, e.g.,* Weingast, *Floor Behavior in the U.S. Congress: Committee Power Under the Open Rule*, 83 AM. POL. SCI. REV. 795 (1989) [hereinafter *Floor Behavior*]; Gilligan & Krehbiel, *Complex Rules and Congressional Outcomes: An Event Study of Energy Tax Legislation*, 50 J. POL. 625 (1988); Krehbiel, *Unanimous Consent Agreements: Going Along in the Senate*, 48 J. POL. 541 (1984); Bach, *The Structure of Choice in the House of Representatives: The Impact of Complex Special Rules*, 18 HARV. J. ON LEGIS. 553 (1981); Enelow & Koehler, *The Amendment in Legislative Strategy: Sophisticated Voting in the U.S. Congress*, 42 J. POL. 396 (1980).

in the 1950's and early 1960's (p. 1). The data he presents show greatly increased amendment activity in both the House and the Senate for selected Congresses from 1955 to 1986. As a result of this trend, Smith claims that "decisionmaking in Congress has taken on a more collegial character, one in which rank-and-file and minority party members took advantage of new opportunities on the floor to exercise their formal equality as partners in policymaking" (p. 1).

In presenting his argument, Smith is appropriately sensitive to differences between the House and the Senate. The House, Smith argues, has gone through three distinct eras in the post-World War II period. From the 1950's to 1973, amendment activity was tightly constrained by a hierarchical institutional structure and floor voting procedures that concentrated power in the hands of a few committee chairmen (pp. 20–21). In the early 1970's, however, reform of the committee system and the introduction of electronic voting gave rise to a second era characterized by an explosion of amendment activity. This era provided unprecedented opportunities for rank-and-file members to participate in the policy process, but it posed problems for Democratic party leaders, who found that members of the Republican minority were using amendments to obstruct or eviscerate Democratic legislation and force votes on "divisive" issues (pp. 33–34). As a result, toward the end of the 1970's and into the 1980's, Democratic leaders took steps to control the extent of amendment activity on the floor, especially by using their control over the Rules Committee to fashion increasingly restrictive rules to govern floor consideration of committee bills (pp. 40–45). The growing use of restrictive rules, along with budgetary constraints and the imposition of limits on riders to appropriations bills, leads Smith to characterize the 1980's as a third era in which amendment activity was more constrained, although still more prevalent than the 1950's and 1960's (pp. 2, 49).

Smith's picture of the Senate reinforces the "conventional wisdom" that the Senate is more floor-oriented than the House (p. 88). By comparison with the dramatic shifts in floor behavior in the House, however, Smith's account of changes in the Senate depicts a more "incremental" evolution towards greater floor activity (p. 87). As Smith points out, the rules of the Senate and the chamber's characteristic individualism make it relatively easy for small minorities to block procedural innovation. Con-

sequently, unlike the House, "the procedural arena in which floor debate is conducted in the Senate has remained remarkably stable since the 1950's" (p. 87). Within this more stable context, Smith finds a steady trend toward higher levels of amendment activity from the 1950's to the late 1970's, with some decline in amendment activity in the 1980's. Smith notes that it has been more difficult for Senate leaders than leaders of the House to adapt to and manage this more active floor environment given the absence of germaneness requirements, the increasing use of filibusters, and the requirement of unanimous consent to limit debate.

Call to Order is well-written, carefully researched, and cautious in both its conclusions and recommendations. Smith's data convincingly demonstrate the changing patterns of floor behavior in the House and Senate, and the author must be praised for his herculean efforts in coding amendments, rules, and unanimous consent agreements from the 1950's to the 1980's. Smith's meticulous and systematic data collection efforts represent a significant contribution to both our understanding of Congress's development during the past forty years, and to the empirical study of legislatures in general. The only noticeable "hole" in Smith's data is the absence of data on the 97th and 98th Congresses (1981–1984). Consequently, Smith's generalizations about floor politics in the 1980's rest on only two data points, the 96th Congress (1979–80) and the 99th Congress (1985–86). Smith explains that data were not collected for the 97th and 98th Congresses due to "the difficulty of using the unbound daily edition and monthly indexes of the *Congressional Record* to identify members' amendments" (p. 260).

While *Call to Order* clearly succeeds as an empirical account of changing congressional floor behavior, the book's problems lie in Smith's analysis of the causes and consequences of the patterns he describes. This is partially a product of analytical deficiencies in Smith's presentation, and partially indicative of the need for further research on the nature of change in legislative institutions.

Upon completing *Call to Order*, the reader is left with the impression that the trends in floor behavior described in the book stand as a dependent variable still in search of a focused, cohesive explanation. This is not because Smith has failed to identify the factors that underlie the changes he has described. Indeed, Smith has provided a rich historical account that in-

cludes numerous changes, both internal and external to Congress, which may have promoted change in floor procedures and produced varying levels of floor activity. Nevertheless, Smith does not have a unifying theoretical framework to guide his analysis, assign causal priority to particular factors, and specify causal linkages. As a result, Smith's explanations have an *ad hoc* character and occasionally miss important analytical distinctions.

One example of Smith's failure to recognize significant analytical distinctions is the overly general character of his assertion that "the ebbs and flows in the character of floor decisionmaking were a response to interparty competition" (p. 13). In fact, as Smith's account clearly indicates, *intra*party competition among House Democrats has been at least as important as *inter*party competition in producing changes in House floor decision-making. The opening up of the amendment process during the early 1970's was the result of a reform movement engineered by younger, liberal Democrats and aimed primarily at reducing the power of older, conservative Democratic committee chairmen. The move by the Democratic leadership to exert greater control over floor proceedings in the 1980's was a response to the dual challenge of aggressive minority amendment activity and a Republican president, but it could not have occurred without increasing homogenization of the policy views of Democratic members of Congress (pp. 42–43). By casually lumping all party-related factors together under the conceptual label "interparty competition," Smith sacrifices the ability to develop an explanation of procedural change that would isolate the independent effects of interparty competition, intraparty competition, and the interplay between the two.

In assessing the implications of congressional evolution toward greater floor activity, Smith focuses on two important institutional issues: (1) the participation of individual members in the floor consideration process, and (2) the power of committees. With respect to individual members, Smith's discussion is informative, but it offers little in the way of original insight. Using data on amendment activity and success rates by individual senators and representatives, Smith demonstrates that the norms of apprenticeship and deference to committee recommendations, predominant in the 1950's and early 1960's, have largely disappeared. Junior members of both chambers have greatly expanded their level of participation in floor delibera-

tions, and non-committee members have increasingly been willing to challenge committee recommendations by offering amendments to committee bills (pp. 139–45). By Smith's own account, however, it is already the "undisputed conventional wisdom" (p. 141) that the apprenticeship and committee deference norms have seriously eroded.[3] As a result, Smith's data serve only to help isolate the timing of the changes.

Smith's discussion of how committee power has been affected by greater floor activity shows a significant appreciation for the complexities of assessing committee power and represents a careful attempt to explore empirically a number of propositions that have been advanced in the theoretical political science literature.[4] Smith does demonstrate that committee recommendations are challenged more frequently on the floor, and, not surprisingly, that those committees with the "largest, most salient, and most controversial policy agendas" (*e.g.*, Appropriations, Ways and Means, Energy and Commerce) have experienced the greatest increase in amendment activity (p. 176). At the same time, however, he notes that committee bill managers may be able to counter effectively hostile non-committee amendments by offering amendments of their own (pp. 183–87). He also notes that in the House, committees subject to heavy amendment activity have become the primary beneficiaries of the movement toward heightened use of special rules in the 1980's. As a result, the bills offered by these committees are often protected from hostile amendments. Given these contrasting trends, Smith cautiously concludes that committees have lost some degree of autonomy, but he is unsure about the extent of this decline.

While Smith's discussion of committee power does address a number of important points, it is deficient in its failure to highlight the shift in the balance of power between party leaders and committees that is implicit in the pattern of floor consideration in the House during the 1980's. The vast majority of committee

---

[3] *See, e.g.*, H. SMITH, THE POWER GAME: HOW WASHINGTON WORKS 25, 137 (1988); Rohde, Ornstein & Peabody, *Political Change and Legislative Norms in the U.S. Senate*, in STUDIES OF CONGRESS 147 (G. Parker, ed. 1985); Ornstein, *The Open Congress Meets the President*, in BOTH ENDS OF THE AVENUE: THE PRESIDENCY, THE EXECUTIVE BRANCH AND CONGRESS IN THE 1980'S 185 (A. King, ed. 1983); Asher, *The Learning of Legislative Norms*, 67 AM. POL. SCI. REV. 499 (1973).

[4] *See, e.g.*, *Floor Behavior*, *supra* note 2; Shepsle & Weingast, *Foundations of Committee Power*, 81 AM. POL. SCI. REV. 85 (1987); Krehbiel, Shepsle & Weingast, *Why Are Congressional Committees Powerful?*, 81 AM. POL. SCI. REV. 929 (1987).

bills require a rule from the Rules Committee in order to be considered on the floor. In the 1950's and 1960's, however, the Rules Committee was largely independent of party control. Furthermore, once a bill made it to the floor, committee recommendations were protected, albeit imperfectly, by norms of committee deference and House floor procedures that limited hostile amendment activity. In the more active floor environment of the 1980's, by contrast, the ability of committees to protect their bills is dependent on their success in obtaining special restrictive rules from a Rules Committee which, due to 1970's reforms, is now tightly controlled by the party leadership. As a result, committees now have an incentive to take party leadership preferences more seriously in writing legislation. Moreover, the party leadership can use its control over the Rules Committee to fashion rules that promote floor outcomes not favored by committee members.

The final chapter of *Call to Order* offers Smith's recommendations for reform based on his analysis of differences in House and Senate styles of floor decisionmaking during the 1980's. In this chapter, Smith distinguishes between two forms of political discussion: "debate," in which argument centers on a small set of crystallized options, and "deliberation," in which there is a careful consideration of all alternatives (pp. 238–39). According to Smith, the floor process of the House in the 1980's, with its growing emphasis on special rules and procedures that restrain amendment activity, more closely approximates debate (p. 240). The Senate, by contrast, remains highly individualistic and has permitted few restrictions on floor proceedings. As a result, the Senate floor process more nearly approaches deliberation (pp. 242–46).

In essence, Smith's reform proposals call for each chamber to become more like the other. Smith argues that the use of special rules have constrained House floor proceedings too much, and that the adoption of restrictive rules should be made subject to a three-fifths, rather than a majority vote (p. 248). In the Senate, Smith concludes that the preference for an individualistic, deliberative process has led the Senate to retain rules which tolerate, and indeed encourage, "unchecked procedural obstructionism" (p. 249). He argues that the Senate should adopt procedures to permit limitation of floor consideration with less than unanimous consent, and that it should consider other

reforms such as germaneness requirements to further control amendment activity.

The reforms Smith suggests to impose tighter restrictions on debate in the Senate are timely and appropriate. These reforms would prevent legislation from being "held hostage" by the dilatory tactics of one or two extremist or intransigent senators. Smith's proposal that the House become more "deliberative" is less worthy of implementation. Growing control of the floor process by the majority leadership is one of the few developments in the past twenty years tending toward greater coherence in American public policymaking. Especially given the larger size of the House, Smith's proposal can only exacerbate the problems inherent in the House's fragmented post-reform committee system. This is particularly true in an era when budgetary pressures require Congress to make trade-offs between competing objectives, rather than simply accede to interest group and constituent demands.

Call to Order provides an important and well-researched description of changes in congressional floor politics from the 1950's to the 1980's. While the book has some analytical deficiencies and its recommendations may be challenged, its empirical work on floor politics fills a glaring hole in the existing literature on Congress and raises important issues for future research. It is highly recommended for all those with a serious interest in the American legislative process.

—*Erik H. Corwin*

WELFARE POLICY FOR THE 1990s. Edited by *Phoebe Cottingham & David Ellwood.* Cambridge, Mass.: Harvard University Press, 1989. Pp. vii, 349, references, notes, contributors, index. $30.00 cloth.

Phoebe Cottingham of the Rockefeller Foundation and David Ellwood of the Harvard University Kennedy School of Government have collected in *Welfare Policy for the 1990s* conference papers from leading specialists on welfare policy to assess alternative strategies for the next decade. The well-organized materials on topics of interest to empiricial social scientists and legislative analysts can be categorized according to the welfare

strategies they discuss: (1) reforming the welfare system by emphasizing a stern work component; (2) relying more on non-welfare means of support such as child-care and health benefits; and (3) generating economic prosperity to achieve full employment.

Of most interest to those with a legislative focus is the article by Robert Reischauer of the Brookings Institution on welfare reform legislation. Reischauer correctly assesses the goals and impact of the latest welfare reform legislation, the Family Support Act of 1988, Pub. L. No. 100-485, 102 Stat. 2343 (1989), as modest. He forecasts that welfare reform will be back on the congressional agenda in the 1990s as research accumulates, legislative provisions expire, and the ideological conditions for a "striking liberal rebound" develop (p. 12). In his estimation, a consensus for welfare reform will evolve from increased empathy for the poor as family structures fragment further and real benefit levels wither. Reischauer also summarizes the political constraints on welfare reform: the omnipresent budget deficit, the decline in presidential direction on welfare policy in the Reagan era, and the fragmentation of power over social policy-making in Congress (pp. 35–37).

Most of the other articles assess specific strategies. For example, Douglas Besharov, a resident scholar at the American Enterprise Institute, examines "targeting" long-term dependents on welfare—the poorly educated, unmarried young mothers whose central financial experience is welfare. Ellwood points out that although "targeting" has a surface plausibility, it has been difficult to document empirically the benefits of targeting (pp. 273–74). Ellwood also notes that targeting does not always help because it often stigmatizes the "truly needy" as the "truly irretrievable" and hence reinforces their condition (p. 274).

Denise Polit, president of Humanalysis, Inc., and Joseph O'Hara, president of Institutes for Health and Human Services, Inc., contribute a remarkable summary of evidence to support the reasonable inference that providing day care for children of welfare dependents plays a vital role in allowing those dependents to find and keep employment. Robert Lerman, a senior research associate at the Brandeis University Heller School, also suggests that providing adequate support for children could help reduce welfare dependence. He points to depressing evidence about how little child support payment is collected from absent fathers, which adds to the already formidable burden on

single-parent families (pp. 227–30). Lerman suggests that the tax system offers potential solutions to child support problems. Programs like Wisconsin's widely cited Child Support Assurance System (currently a demonstration project) could help finance uniform child support from absent fathers through the tax system (p. 235).

Unfortunately, some of the policy discussions in the collection are lacking in economic sobriety. William Julius Wilson's paper summarizes and expands upon his work examining how segregation and structural job shifts have lead to the decline in the number of middle-class, black men capable of supporting families. But his exhuming of classical fiscal expansionism as a strategy for coping with the welfare problem seems unrealistic in an era of budget deficits, financial conservatism, and presidential pledges of "no new taxes." David Ellwood admits he "would trade away many of my favorite ideas for welfare reform for a guarantee of the 2-3 percent unemployment and high growth one finds in Massachusetts today" (p. 288).

Ellwood concedes that the minimum wage may not benefit the poor and may cause unemployment (p. 281), but a nostalgic commitment to the familiar remains. Most economic analysts favor a wage subsidy rather than an increased minimum wage. Gary Burtless, a senior fellow at the Brookings Institution, however, provides a skeptical account of the wage subsidy as part of a masterful presentation of employment issues (pp. 134–37). Wage subsidies act to subsidize low-wage employers and are more difficult to administer than the Earned Income Tax Credit ("EITC"). The EITC subsidizes the earnings of workers who have financial responsibility for their children with a modest tax credit (fourteen percent credit for the first $5714 in earnings which declines to zero when earnings reach $17,000) (p. 135). Neither a wage subsidy nor the EITC, however, helps those families without a primary breadwinner (p. 136).

Burtless also has gloomy predictions on training welfare dependents for work. He reports evidence that training does increase wages, but it is not clear how long this effect persists, or whether wage gains come at the expense of those not participating in training programs (pp. 126–27). Furthermore, the most successful training programs cost taxpayers nearly $10,000 a person (p. 118). In any event, the recurring complaint by employers about employees is not specific training, but a lack of basic literacy (p. 143).

The entire collection of papers in *Welfare Policy for the 1990s* could have benefitted greatly from a deeper discussion of the global integration of labor markets for the transition from welfare to employment in the 1990's. The American professional work force is quite competitive in international markets. But low-wage American workers are now exposed to competition from lower-wage workers in the newly industrializing countries. More integrated global markets require more productive American workers to shelter a premium wage. High illiteracy rates present a significant obstacle to training workers to be more productive, especially given the need to work with advanced technolology. Global markets will increasingly siphon off the only viable jobs for the lower-wage population. The problem of international competition will not be solved by full employment.

Welfare Policy for the 1990s reflects the third phase of post-1960 social welfare thinking. The Great Society programs began the first era of welfare policy approaches, which ended with the bitter debate over then-White House aide Daniel Patrick Moynihan's "discovery" of the unraveling of the black family structure and the discrediting of President Nixon's Negative Income Tax proposal. The second phase began, and floundered, on President Carter's abortive welfare reform effort, although the era left the Earned Income Tax Credit as its legacy. The Reagan Administration controlled the third era, influenced by Martin Anderson's *Welfare*[1] and Charles Murray's *Losing Ground*.[2] Anderson argued that welfare should be reduced and focused only on the "truly needy," while Murray blamed the welfare system itself for creating dependency and miring its intended beneficiaries in poverty.

The Family Support Act of 1988 embodied the broad consensus developing about the intractable nature of the problem. Its approach was framed around the themes of mutual obligation, self-sufficiency, family stability, education and training, child support, social science humility, and a reluctance to spend money. The works collected in *Welfare Policy for the 1990s* reflect this consensus about the intractable nature of the problem and the need for modest goals achieved through small steps. The most important characteristic of this collection, however,

---

[1] M. ANDERSON, WELFARE: THE POLITICAL ECONOMY OF WELFARE REFORM IN THE UNITED STATES (1978).

[2] C. MURRAY, LOSING GROUND: AMERICAN SOCIAL POLICY, 1950–1980 (1984).

is that it combines insight, modesty, and compassion, qualities so essential to gradual progress in this demoralizing arena.

—*Vincent Eagan*

INNOCENCE AND EXPERIENCE. By *Stuart Hampshire*. Cambridge, Mass.: Harvard University Press, 1989. Pp. 189, index. $20.00 cloth.

It is a testimony to the creative and interdisciplinary approach often attributed to philosopher Stuart Hampshire's work that as *Innocence and Experience* responds to Niccolo Machiavelli's political philosophy, it evokes a literary association: the contrasting pictures of the young chimney sweeper in William Blake's twin poems, both entitled "The Chimney Sweeper," in his *Songs of Innocence and Experience*.[1] Hampshire's treatment of the question answered in the affirmative by Machiavelli—"Is there a perpetual, even a necessary, conflict in human nature between innocence and experience?" (p. 13)—touches on Blake's notion of innocence and imaginative transcendence of present circumstance, in contrast to experience and the inevitable squalor of reality. However, Hampshire's vision of innocence and experience radiates outward, embracing a myriad of philosophical and humanistic questions, and ultimately provides a reply to Machiavelli's problem of the contrast between the virtues of innocence in private life and the traits of experience seemingly necessary in the hardened world of politics.

This reply seems to find the balance between moral relativism and the assertion of a universal conception of the good for humanity. Hampshire's theory of procedural justice recognizes a slim but absolute moral requirement of fair, rational argument with both sides being heard in situations of controversy. This theory avoids the assertion of a single conception of the universal and necessary morality for all societies at all times, without falling into an entirely relativistic approach to morality. The theory thus limits what goes on in the Machiavellian realm of experience without having to deny the ultimately inevitable conflict between innocence and experience.

---

[1] Blake, *Songs of Innocence* and *Songs of Experience*, in ENGLISH ROMANTIC WRITERS 54, 64 (G. Perkins, ed. 1967).

From the deliberations of the council of war in the *Iliad*, to the inner discussion preceding the actions of a prudent person, to the courts of modern democracies, Hampshire argues, we find a common institution that expresses the core of practical rationality: articulation and review of contrary opinions on policy (p. 52). The procedures of adjudication and the weighing of arguments are understood and applied across religious, national, and moral barriers, even by hostile powers in negotiations. This is not surprising precisely because such procedures are the outward equivalents of the methods of thought that everyone employs, to some extent, in inner debates when reckoning with two competing demands (p. 54). The existence of mechanisms of practical rationality and decision-making used by all humans implies the viability of a universal, ground-level notion of procedural justice and fairness.

In fact, procedure is always a part of justice and fairness, in the sense that an outcome of a biased process of argument and discussion—even an outcome that in substance seems fair—will always be tainted with procedural corruption (p. 53). Similarly, as we know from our own experience in making decisions, we describe a rash choice—even the "right" one—as made on the spur of the moment, without thinking it through. In social policy decisions, the inherent procedural aspect of justice and fairness means that a minimum requirement for a fair and just outcome is adequate reasoning and argument in an impartial tribunal that weighs all relevant considerations. Otherwise, the result will reflect an "incongruity between means and ends" (p. 184).

Aware of the importance of linking a theoretical conception of justice to aspects of our actual experience, Hampshire looks to the Russian Revolution and the Nazi ascendancy as the two most important sources of evidence for moral philosophy of our time (p. 66). We may better understand the meaning and importance of justice upon consideration of the forces of evil, destruction, and tyranny, which the virtue of justice is meant to obstruct. In both the Russian Revolution and the Nazi movement, a particular conception of the right sort of life—it is difficult to use the more common philosophical term "conception of the good" in speaking of Nazis—became a fanatic obsession at the expense of just procedures of negotiation and decision-making.

Hampshire's experience as an intelligence officer in the British military in World War II "altogether changed [his] attitude both to politics and to philosophy" (p. 8). He observed first-

hand "how easy it had been to organise the vast enterprises of torture and of murder . . . once all moral barriers had been removed by the authorities" (p. 8). Such abrogations of procedural justice in the name of one particular theory make reasonably peaceful and coherent life impossible, as the turbulent history of this century reminds us with its spectre of the abolition of justice and morality in public life very nearly realized in Europe (p. 75).

Speaking of friends who were revealed to have been Soviet agents—friends whose dedication to a Marxist theory of history led them to support the Communist Party's mass murders, tyranny, and destruction—Hampshire writes that "[o]ne could therefore ask why perceptions of injustice had to be disguised and deformed by philosophical theory before they were thought to be respectable" (p. 10). There is always the lurking danger of the working up, by a philosophical theory, of a set of ideas into a fanaticism in the face of which injustices are disguised and minimized. Procedural justice can then be understood as a means of providing stability and balance in a world of competing moralities through its ground-level morality of fair dealing (p. 72).

In order to defend his conception of basic procedural jusice, Hampshire argues against the notion, shared by philosophers from Plato and Aristotle to Mill and other utilitarians, that there must be one determinate greatest good, or *telos*, for all humanity. Hampshire writes that "ranked alongside" the intellect that unites humanity across all barriers, and enables the development of logic, mathematics, and the natural sciences, is the human capacity for linguistic, cultural, and moral diversity (p. 30). For us to identify a being as human, it is not enough that such a being can make logical inferences and follow rational arguments; we would expect the being to want to tell stories, to be interested in a history of her own people, and to possess and take pride in a local and unique natural language (p. 44). Because imaginative variation in language, culture, and history constitutes our essential humanity, diversity in human conceptions of the good expresses and celebrates what it means to be human. To assert a universal conception of the good for all humanity, in contrast, denies or subordinates this imaginative variation.

Hampshire also confronts the challenge to his conception of procedural justice posed by "Hume's ghost." Humean skepticism argues that all values are simply projections of our partic-

ular feelings onto things that we then characterize as "good" or "bad" (p. 81). This skepticism challenges Hampshire's insistence on minimum, but absolute, procedures of justice as a means of avoiding what are regarded to be evils across history and circumstance: death, oppression, and tyranny. Hampshire argues against Hume by showing that not just moral judgments but ordinary empirical statements, such as "once this brake is released, this car will run into the wall downhill from it," must assume some constant conditions to be true (p. 84). There is always the theoretical possibility that some strange natural occurrence—a sudden earthquake when there has never before been a tremor in the area—will render false the empirical judgment about the car running down the hill (p. 83). In making statements about the world, whether of an empirical or a moral nature, we necessarily "abstract[] from the infinite totality of things, and . . . presuppose that both the natural world and the social world are proceeding normally" (p. 86). Corresponding to such constant regularities as the effects of gravity, or the alternation of night and day, presupposed in everyday natural explanation, "there is nothing mysterious or 'subjective' or culture-bound" in the "great evils of human experience, re-affirmed in every age and in every written history and in every tragedy and fiction" (p. 90). Thus we can say that the universality of the requirement for procedural justice is rooted in the universality of the predicament—conflicting conceptions of the good alongside the potential for great evils—that gives rise to it.

Having argued against both challenges to his notion of diverse and local conceptions of the good coexisting within a framework of basic and universal procedural justice, Hampshire turns at last to Machiavelli's problem, which he considers a critical, yet little recognized, threat to moral and political philosophy of all varieties. The problem lies in the apparent incompatibility between the virtues of innocence in private life and the virtues of experience in politics (pp. 11–13, 162–68). In reflecting on what we mean by "innocence" and "experience," Hampshire suggests that we think in terms of pictures (p. 172). To grasp the notion of innocence, Hampshire suggests visualizing an early Quaker meeting house with fresh, white walls and no ornamentation, in which the bustle of the world has no sway and the vision is of simpleness, whiteness, straightness, and clearness (pp. 172–73). Experience, then, is the domain of staterooms in great palaces or the Vatican, rooms that witnessed "innumerable wars . . .

unwanted compromises, embarrassing alliances, distressing maneuvers, and secret betrayals" (p. 174). The clarity of vision and pureness of purpose characteristic of innocence cannot be reconciled with the necessities of political leadership—deceit, violence, compromises of ideals—that the corridors of experience have witnessed.

The realities of political life include the expectation of unavoidable squalor and imperfection, of guilty knowledge, of necessary disappointments and mixed results, of half-success and half-failure. "A person of experience," Hampshire writes, "has come to expect that his usual choice will be of the lesser of two or more evils" (p. 170). When the choice is whether to distribute limited funds to the homeless or to Head Start programs, a politician dedicated to an active welfare state cannot maintain clarity and pureness of vision, cannot remain innocent. Other conflicts may involve more traditionally Machiavellian subject matter: the choice between sacrificing the rights of a minority and facing a substantial threat to state security, or the choice between a war that may kill millions of citizens and a dangerous and aggressive nation across the border.

While there is an inherent conflict between innocence—purity of intention and steadfast virtue—and experience—moral compromise and the occasional cruelties and exploitations involved in the political sphere—the notion of minimum procedural justice provides a meaningful constraint on the Machiavellian conception of what a political leader must do to survive in the realm of experience. The politician trying to decide between a risk to national security and an infringement of minority rights will have his own conception of the good that ultimately will tip the balance one way. But procedural justice requires a process of weighing, in which the politician establishes the nature of the claims upon her *in their own terms*. In any negotiation or decision-making process, including one between hostile nations, the drives to dominate are constrained by the basic and universal requirement of fair argument with equal opportunity for all to be heard. Our human faculty of practical reasoning, by which alternative choices are weighed in the chambers of justice just as they are in our own minds in personal decisions, can save us from the worst abuses of Machiavelli's world of experience.

Hampshire's book succeeds in providing a rigorous defense of his conception of procedural justice. His hope is that the book will be of interest to politicians as well as philosophers,

and to this end he gives some background explanation on the philosophies he discusses, although perhaps not enough to make his book accessible to a wide spectrum of American politicians. The book is a delight to read because of Hampshire's expressive use of historical and literary associations to convey his ideas.

Innocence and Experience is profoundly thought-provoking not only because of its central thesis, but also because of the vast array of issues and possibilities raised by Hampshire's work. Particularly interesting are Hampshire's suggestion that any one philosophical theory is simply one possible reading of the themes of moral and political philosophy, a reading determined by the philosopher's own experience of the world (p. 3); his idea that we must borrow from the vocabulary of political and social institutions when we describe the operations of our minds (p. 51); and his explanation of historical injustices, such as slavery, as practices so imbedded in local or cultural tradition that they are considered "natural" and therefore outside the sphere of practical reasoning until gradual questioning brings them within its ambit (p. 56). In proposing such ideas, Hampshire's book not only conveys the openness of his own mind to new possibilities and the creativity of his own thinking, but draws the reader along the same path.

—*Christine Jolls*

TOWARD A FEMINIST THEORY OF THE STATE. By *Catharine A. MacKinnon.* Cambridge, Mass.: Harvard University Press, 1989. Pp. xvii, 330, notes, index. $25.00 cloth.

Many readers of Catharine MacKinnon's new book will be familiar with her work through *Feminism Unmodified*,[1] an earlier collection of speeches, or through her numerous articles in journals such as *Signs: Journal of Women in Culture and Society*.[2] The thirteen essays assembled in *Toward a Feminist Theory of the State* were culled mostly from previously published writings. What makes this volume new, MacKinnon

[1] C. MACKINNON, FEMINISM UNMODIFIED (1987).

[2] *See, e.g.*, MacKinnon, *Feminism, Marxism, Method, and the State: Toward Feminist Jurisprudence*, 8 SIGNS: J. OF WOMEN AND SOC'Y 635 (1983); MacKinnon, *Feminism, Marxism, Method, and the State: An Agenda for Theory*, 7 SIGNS: J. WOMEN IN CULTURE AND SOC'Y 515 (1982). A full list of MacKinnon's publications is provided on page 321 of TOWARD A FEMINIST THEORY OF STATE.

claims, is that it presents her feminist argument "in its original unity, shape, and order" (p. xiv). The result is an exceptionally cogent and passionate exposition of feminist thought.

"Contemporary feminist theory," according to Seyla Benhabib, "is experiencing a *Methodenstreit*, a methodological warfare."[3] One patent motivation for this book was MacKinnon's deep dissatisfaction with current feminist thought. "Feminism has not been perceived as having a method, or even a central argument," MacKinnon complains. "It has been perceived not as a systematic analysis but as a loose collection of complaints and issues that, taken together, describe rather than explain the misfortunes of the female sex" (p. 108). MacKinnon attributes much of the blame for this situation to feminist thinkers whose outlook is insufficiently radical or comprehensive. Her efforts to discredit these brands of feminism are one major theme of *Toward a Feminist Theory of State*.[4]

The most serious criticism MacKinnon levels at her colleagues is that they have failed to bridge the gap between feminist theory and feminist practice. In her eyes, only radical feminism embraces both sides of the "unequal coin" (p. xi). Her book aims to perform a dual role, to be at once a philosophical essay and a political tract, to engage issues on the planes of epistomology and politics. *Toward a Feminist Theory of State* searches for a theory of male dominance which will offer both "an analysis of why and how it happened and why (perhaps even how) it could be ended" (p. x).

Part I of *Toward a Feminist Theory of the State* pursues an extended analogy between two critical theories of society, feminism and Marxism. Why MacKinnon devotes one-third of the book to this comparison may not be immediately obvious. As the preface suggests, however, Marxism played a pivotal role in MacKinnon's own intellectual development. Even within the confines of her feminist theory, MacKinnon's understanding of basic concepts like ideology, consciousness, and society remains strongly Marxist.[5] Marxism, MacKinnon explains, "is the

---

[3] Benhabib, *On Contemporary Feminist Theory*, 36 DISSENT 366, 368 (1989).

[4] MacKinnon occasionally calls her own formulation of feminist theory "feminism unmodified," especially to contrast it with its compromising rivals. More often, in light of her broader reconstructive aims, she simply calls it "feminism." Unless otherwise noted, her practice is followed.

[5] Seyla Benhabib and Drucialla Cornell have criticized MacKinnon's reliance on Marxist categories in her SIGNS articles for missing the "radical challenge posed for Marxist theory by the very presence of women not only as an oppressed group but as

contemporary theoretical tradition that—whatever its limitations—confronts organized social dominance . . . identifies social forms that systematically shape social imperatives, and seeks to explain human freedom both within and against history" (p. ix). Calling feminism "post-marxist" does not mean that feminism rejects Marxism's accomplishments. "It means that feminism worthy of the name absorbs and moves beyond Marxist methodology, leaving theories that do not in the liberal dustbin" (p. xiii). Feminism begins, one might say, where Marxism leaves off.

Part II, entitled "Method," presents the theoretical groundwork of radical feminist thought. It begins by introducing feminism's two analytic categories: sexuality and gender. "Sexuality is to feminism what work is to marxism: that which is most one's own, yet most taken away" (p. 3). Sexuality is an aspect of personal identity given meaning within a set of societal relations. "Not confined to that which is done as pleasure in bed," it is defined by whatever the culture views as sexual (p. xiii). Gender is the name given to the organization of sexuality into two orders, one embodying dominance and the other submission. Mainstream feminism defines these two concepts differently, linking sex with the biological and gender with the social. In MacKinnon's theory, both sexuality and gender are constructions that institutionalize a basic division of power in society. MacKinnon expresses something of the complex relationship within this conceptual triad—sexuality, gender, and power—in the following formula: "As sexual inequality is gendered as man and woman, gender inequality is sexualized as dominance and subordination" (p. 241).

The equation of sex and power is axiomatic to MacKinnon's radical feminism. This thesis, rather than any definitional nicety, separates feminism unmodified from liberal feminism and leftist feminism. Liberal feminism views gender difference as merely an irrational overlay on a social reality composed of otherwise free, rational, and autonomous agents. Once difference is removed, domination will cease. Leftist feminism views gender difference as a function of material conditions in society; it is a consequence of the relations of production and is unalterable

---

collective actors in the historical scene since the middle of the nineteenth century." Benhabib & Cornell, *Introduction: Beyond the Politics of Gender*, in FEMINISM AS CRITIQUE 1, 3 (1987).

on its own. MacKinnon argues that "[t]he failure to face and criticize the reality of women's condition, a failure of idealism and denial, is a failure of feminism in its liberal forms. The failure to move beyond criticism, a failure of determinism and radical paralysis, is a failure of feminism in its left forms" (p. 241). In contrast, radical feminism, MacKinnon asserts, "has begun to uncover the laws of motion of a system that keeps women in a condition of imposed inferiority," because "[i]t has located the dynamic of the social definition of gender in the sexuality of dominance and subordination" (p. 241).

A battery of statistics and studies provides the external evidence for the feminist identification of sexuality with power. That few women are safe from the terror of violence emerges clearly from the data. Nearly half of all women are raped at least once in their lives; more than one-third of all women are sexually molested by male relatives or friends at an early age; one-third of all women are battered at home (pp. 142–43). But the underlying structure of domination expresses itself in other forms beyond physical violence. Ten billion dollars are spent annually on pornography (p. 139); approximately twenty percent of American women are or have been prostitutes (p. 143); eighty-five percent of working women will experience sexual harassment on the job sometime during their career (p. 143). To overlook these practices and focus on deviant, criminal behavior is to dismiss domination as an exceptional, and officially proscribed, activity.

From feminism's standpoint, none of these practices—rape, unequal pay, domestic battery, sexual objectification, and denial of reproductive control—stands independently; together they constitute one seamless reality (p. 244). No one practice is more emblematic of male power than the others. MacKinnon appreciates the difficulty legislators experience writing rape and pornography laws. It is indeed a conceptual challenge to distinguish rape from intercourse when intercourse, "under conditions of gender inequality," often looks so much like rape. It is clearly a conceptual challenge to distinguish pornography from art and advertisement when, in the real world, art and advertisement cannot be told apart from pornography (pp. 112–13, 146). The edges blur. The totality of women's experience emerges as a "cohesive whole within which each [event] resonates," a whole whose defining theme is the male pursuit of control over women's sexuality (p. 112).

Liberal political theory assumes the existence of a private sphere, "personal, intimate, autonomous, particular, individual," a haven for freedom and a refuge from politics (p. 190). The situation of women belies this premise daily. For women, the realm of the private and the intimate is nothing other than the locus of power's operation. The feminist dictum "the personal is political" thus has two meanings. First, it means that "the private is the distinctive sphere of intimate violation and abuse," the realm of "collective subordination" (p. 168). From issues of a woman's right to control her own body to issues of whose interests rape law protects, gender inequality does not respect the public/private barrier that liberalism erects. Second, it means that power not only structures relations between individuals, but also transfigures the self. Female sexuality is an element of women's self-identity that is mediated by society and undergirded by power relations. The results of MacKinnon's investigation of the female self through the method of "consciousness raising"—eliciting women's reflections on their experiences of abuse and powerlessness—provide some of the most moving passages in the book. In one haunting exchange, a pornography model describes what it means to feel split from one's body, what it means to have lost one's sensibility to pain (pp. 147–48).

Part III, entitled "The State," critiques the liberal state based on MacKinnon's theory of gender hierarchy, and advances practical proposals for reforming the law as it affects women. The root problem with the liberal state, from a feminist perspective, is that it seeks legitimation from a male ideology that "sees and treats women the way men see and treat women" (p. 162). Specifically, the "posture and presumption of the negative state" is "the view that government best promotes freedom when it stays out of existing social arrangements" (p. 164). But this outlook can only serve to make women's concerns invisible, since women encounter inequality in civil society itself, the supposed realm of personal freedom. The fact is that "[w]omen are oppressed socially, prior to law, without express state acts, often in intimate contexts" (p. 165).

In these five chapters, MacKinnon addresses, among other issues, rape, abortion, and sexual discrimination. The best example of how male ideology distorts women's reality and then codifies this distortion in law is pornography. Pornography occupies a critical position in MacKinnon's theory since it literally

creates an image of woman that is reflected throughout society; pornography is nothing less than "a means through which sexuality is socially constructed, a site of construction, a domain of exercise" (p. 139).[6] Liberal theory sees pornography as raising questions of first amendment rights, issues of free speech. What it does not see is that pornography humiliates women, objectifies them, subordinates them, debases them, and silences them. "That pornography chills women's expression is difficult to demonstrate empirically because silence is not eloquent," MacKinnon concedes. "Yet on no more of the same kind of evidence, the argument that suppressing pornography might chill legitimate speech has supported its protection" (p. 206). To invoke the free speech argument is to presuppose that speech is free; the reality of gender inequality means that, for women, it is not. This fact alone speaks to the error of liberal thinking. Once it is cleared away, MacKinnon contends, pornography can be seen as "a form of forced sex, a practice of sexual politics, an institution of gender inequality" (p. 197).

MacKinnon's penetrating analysis occasionally lapses into argument that offers more rhetoric than insight. At one point, she expounds:

> [M]ale morality sees that which maintains its power as good, that which undermines or qualifies it or questions its absoluteness as evil. Differences in the law over time—such as the liberalization of obscenity doctrine—reflect either changes in which group of men has power or shifts in perceptions of the best strategy for maintaining male supremacy—probably some of both. But it must be made to work (p. 201).

Here MacKinnon abandons her search for society's "laws of motion" and opts for a garden-variety conspiracy theory. Passages like this are unfortunate if only because MacKinnon is usually so wary of uncritical thinking.

Some readers will feel dismayed at the heavy jargon of the book, which is uniformly dense and at times virtually opaque. MacKinnon is not unaware of the problem. She remarks elsewhere, "Sometimes I think to myself, MacKinnon, you write.

---

[6] MacKinnon was a principal author of a pornography ordinance for Indianapolis, Indiana, which banned pornography defined as "sexually explicit" material which "subordinates" women. The ordinance was held unconstitutional. *See* American Booksellers Ass'n, Inc. v. Hudnut, 771 F.2d 323 (7th Cir. 1985), *aff'd mem.*, 475 U.S. 1001 (1986). Part III of *Feminism Unmodified* is devoted, in its entirety, to a feminist treatment of pornography.

Do you remember that the majority of the world's illiterates are women?"[7] A related complaint is that her terminology is not entirely consistent. The terms sexuality and gender receive new definitions as often as the author sees fit; the original sense of the words is overlaid with successive levels of meaning. Nevertheless, the argument acquires a kind of cumulative momentum, even through repetition, and gains power through introducing a vocabulary adequate to the investigation of women's experience.

A theoretical problem with MacKinnon's argument stems from her attack on the concept of objectivity. Objectivity is the idea that there is one correct knowledge of the world, one privileged access to reality. Absolute, contextless, and certain, objectivism is seen as the epistemological emblem of male hegemony. MacKinnon quotes Simone de Beauvoir approvingly: "Representation of the world, like the world itself, is the work of men; they describe it from their own point of view, which they confuse with absolute truth" (p. 121). If this is only to say that men, like any empowered group—be it scientists, ministers of propaganda, or the Church—package their truths as Truth Itself, then feminism is saying nothing new. The claim can be far more radical: that there is no such thing as objectivity, that objectivity is a fiction. MacKinnon does write: "Disaffected from objectivity, having been its prey . . . women's interest lies in overthrowing *the distinction itself* [between subjectivity and objectivity]" (pp. 120–21).

Feminist theory here follows the lead of Nietzsche, who rejected objectivity, arguing that "there is *only* a perspective seeing, *only* a perspective 'knowing.'"[8] But all theories that attack objectivity undercut themselves, since they are unable to account for their own validity. If there is no such thing as objectivity, no such thing as objective truth, can MacKinnon claim that her theory is objectively true? If not, what reason is there to believe it?[9] A related problem follows from the claim

---

[7] MACKINNON, FEMINISM UNMODIFIED 47 (1987).

[8] Nietzsche, *The Genealogy of Morals* in BASIC WRITINGS OF NIETZSCHE 555 (W. Kaufmann trans. 1968). It is surprising that MacKinnon never directly acknowledges Nietzsche, a thinker who anticipated many feminist ideas—not merely the rejection of objectivity, but also feminism's particular notions of power and domination. The political consequences of Nietzschean philosophy stand in striking consonance with feminist politics. *See* M. WARREN, NIETZSCHE AND POLITICAL THOUGHT (1988).

[9] Hilary Putnam develops this line of criticism against relativism. *See* H. PUTNAM, REASON, TRUTH, AND HISTORY 103–26 (1981).

that the male perspective of objectivity is "systemic and hegemonic," that it literally "defines rationality" (p. 114). Was feminist theory itself authored from within the male perspective? If so, then it seemingly cannot point outside to the possibilty of a female alternative; if not, then the male viewpoint is not all-encompassing and total.

MacKinnon acknowledges and even confirms these objections. "Feminism affirms women's point of view, in large part, by revealing, criticizing, and explaining its impossibility. This is not a dialectical paradox. It is a methodological expression of women's situation" (p. 115). Male hegemony of thought, she might argue, *means* that women's consciousness does not yet exist or exists only inchoately. At this time, "there is no such thing as a woman as such" (p. 119). These maneuvers, though common to any critical theory (including Marxism) that attempts to account for the consciousness that created it, merely evade the difficulty; they do not convincingly explain how feminist theory gets off the ground. There is, then, a lingering problem in the way feminist theory calls itself into question. More serious, however, is the way feminist theory calls feminist practice into question. To put the point another way, feminist epistemology seems incompatible with the possibility of a feminist politics.

Feminist politics advocates a jurisprudence that would generate substantive reforms in the law. Ask proponents of feminist politics what can be done to end domination, and they will give concrete, specific proposals. Rewrite the rape law. Make reproductive control a matter of collective empowerment. To any practice, apply the simple test: "Does [it] participate in the subordination of women to men, or is it no part of it?" (p. 248). The problem is that feminist epistemology cannot accommodate these reforms. In the face of a male viewpoint "metaphysically nearly perfect" (p. 116) and a male power structure "systematic and cumulative" (p. 41), nothing short of a revolution can create meaningful change. Were MacKinnon's feminism liberal, certainly a feminist jurisprudence would have some value. But she herself argues that "so long as male dominance is so effective in society that it is unnecessary to impose sex inequality through law, such that only the most superficial sex inequalities become *de jure*, not even a legal guarantee of sex equality will produce social equality" (p. 164). Again, "[s]o long as men dominate women effectively enough in society without the support of

positive law, nothing constitutional can be done about it" (p. 239).[10]

To change the law is to alter, and hopefully deconstruct, gender inequality in society. The reforms MacKinnon advocates are undoubtedly needed if domination is ever to end. But feminist theory's sobering prognosis for the prospects of feminist practice leaves one troubled about the future of the project itself. Nor is MacKinnon unperturbed. An odd tone of defeatism creeps into her prose at several crucial points in her argument. She writes with a curious mixture of defiance and tragic knowledge: "Women's situation offers no outside to stand on or gaze at, no inside to escape to, too much urgency to wait, no place else to go, and nothing to use but the twisted tools that have been shoved down our throats" (p. 117). Between Andrea Dworkin's nightmarish warnings of the "coming gynocide"[11] and liberal feminism's unthinking optimism, MacKinnon faces the future with considerable ambivalence. The term "toward" in the title of her book now assumes added significance. "A feminist theory of the state has barely been imagined," she concludes; "systematically, it has never been tried" (p. 249).

—Mark Harris

THE BATTLE TO CONTROL BROADCAST NEWS: WHO OWNS THE FIRST AMENDMENT? By *Hugh Carter Donahue*. Cambridge, Mass.: Massachusetts Institute of Technology Press, 1989. Pp. xi, 196, notes, index. $19.95 cloth.

The Federal Communications Commission ("FCC") has spoken, and the Fairness Doctrine is dead.[1] Or is it? Implemented in 1949, the Fairness Doctrine required broadcasters to offer news and public affairs programming and provide reasonable opportunities for presentation of a variety of opposing opinions.[2] Led by FCC Chair Mark Fowler, a strong supporter of the Reagan Administration's deregulation agenda, the FCC elimi-

[10] Feminism thought thus recapitulates one of the great schisms in Marxism, between the reformers who advocated working within the system to improve the situation of the working class, and the orthodox socialists who believed that any reforms short of complete revolution would only buttress the capitalist state.

[11] A. DWORKIN, RIGHT-WING WOMEN 194 (1983).

[1] Syracuse Peace Council, 2 F.C.C. Rcd. 5043 (1987), *aff'd*, 867 F.2d 654 (D.C. Cir. 1989).

[2] In the Matter of Editorializing by Broadcaster Licensees, 13 F.C.C. 1246 (1949).

nated the Fairness Doctrine in 1987. Congressional reaction was swift and furious, and only President Reagan's veto prevented legislation that would have codified the Fairness Doctrine (pp. 173–74). One former FCC chair believes that a sulking Congress has held up legislation dear to broadcasters' hearts and pocketbooks until congressional attempts to resurrect the Doctrine succeed.[3]

In Hugh Carter Donahue's view, the recent furor over the Fairness Doctrine is merely the latest battle in the eighty-year struggle for control of the airwaves, a struggle fueled by mistaken mistrust of the powers of the electronic media. Donahue, an assistant professor of journalism at Ohio State University, argues that the Fairness Doctrine and other FCC regulations such as the equal time and the personal attack rules have failed to supply the public benefits promised by the government. Though designed to foster diversity of expression, the Fairness Doctrine established tepid homogeneity, as broadcasters "ducked controversial issues for fear of triggering fairness objections" (p. x). Though intended to prevent unfair domination of the airwaves, the equal time rule failed to include minor or third-party candidates in national debates or news coverage, as powerful "[p]oliticians evaded or manipulated equal time for temporary pragmatic gain" (p. x). In all, Donahue concludes, the "public lost more than it gained by having equal time and fairness rules" (p. x).

The central problem with *The Battle to Control Broadcast News* is that this thesis appears early in the preface and then disappears until Chapter Ten, where it is not connected sufficiently to the previous nine chapters detailing broadcast history. This diminishes, but does not destroy, the charm and value of Donahue's historical examination. Donahue's research and his ability to establish time, atmosphere, and characters with a few well-worded descriptions and a great many well-chosen quotes deserve praise. His tracing of the evolution of broadcast regulation from the early days of radio to the Bush-Dukakis campaign is remarkable for its vivid color and understandable portrayal of the complex maneuverings and competing interests surrounding broadcast news.

---

[3] Ferris & Leahy, *Red Lions, Tigers and Bears: Broadcast Content Regulation and the First Amendment*, 38 CATH. U.L. REV. 299, 300 (1989).

"All interest groups shared two concerns, control and fear," Donahue says, describing the birth of the Fairness Doctrine (p. 50). In 1926, lawmakers were trying to frame a national policy for the infant radio industry. Particularly vexing was the lack of consensus over allocation of broadcast frequencies. People both craved and feared the massive power of the emerging medium; they sought to control and therefore profit from its scope and reach, but also they wanted protections against what Senator Robert B. Howell (R-Neb.) warned might become "a Frankenstein monster" unless properly checked (p. 14).

In part because oddly allied interests failed to coalesce, the extreme positions of total state control and complete public access lost ground to licensing. At the time, licensing appeared to grant wide private discretion within limited and beneficial governmental requirements. Broadcasters wanted a licensing system and were willing to accept requirements to program in the public interest in exchange for free use of a lucrative spectrum and government regulation to clear the airwaves from chaotic claims to frequencies (p. 4). The justifications for licensing became the cornerstones of federal communications policy: public interest, listener sovereignty, and spectrum scarcity (p. 7).

Starting from this initial compromise, governmental regulation also included such requirements as equal time to all political candidates (p. 15). Lawmakers justified equal time on the basis of protecting the public from an ideological monopoly, but Donahue argues that in reality they manipulated the requirements to further their own interests. President Roosevelt, for example, propagandized his New Deal programs during approximately thirty fireside chats broadcast across the country. "FDR muscled commercial broadcasters and broadcast networks to provide national coverage of his fireside chats by forcing the radio networks to interconnect his broadcasts with all their affiliates," Donahue writes. "That way FDR reached a national audience, something neither Father Coughlin nor Huey Long could do with their popular broadcasts" (p. 20). Skittish broadcasters, who feared that Roosevelt might nationalize radio as part of his fight against the Great Depression, opted for cooperation with the chief executive as a means of protecting their licenses (p. 30). Roosevelt skirted the equal time requirement by confining all but one of his chats to times when he was not a candidate for office, thus preventing his opponents from claiming equal

time and, in effect, creating the ideological monopoly the equal time rule was designed to prevent.

Nor did the Fairness Doctrine serve its purpose of promoting diversity and openness. In its only such action in connection with a fairness violation,[4] the FCC refused to renew the license of racist and neo-fascist radio station WXUR. "WXUR expressed offensive views that appalled many Americans, but it also exemplified the diversity of voices, no matter how repellent, that the Fairness Doctrine was supposedly designed to promote," Donahue says (p. 78). The Doctrine had achieved its aims within the parameters of one radio station; however, silencing the controversial station diminished diversity within the listening area.

Broadcaster timidity was another unexpected consequence of the Doctrine. When the FCC appended the Cullman corollary to the Doctrine, under which broadcasters had to provide free response time to controversial editorials or advertisements if management could not locate an articulate opponent willing to pay for the time, "[b]roadcasters got the jitters about just what product advertisements they could carry without triggering Cullman complaints" (p. 138). Broadcasters were leery of airing controversial programs and advertisements for fear of being required to search for suitable spokespeople or donate time for editorial replies (p. 136). In one case, the FCC ruled that NBC had violated the Doctrine in a prize-winning documentary about workers who had lost their pensions without notice. David Brinkley commented: "To be found guilty of unfairness for not expressing that most people are not corrupt or that most pensioners are not unhappy is to be judged by standards which simply have nothing to do with journalism" (p. 126). Though a court of appeals overruled the finding of violation and the FCC eventually dropped its subsequent litigation, it did not do so until three and a half years after the broadcast (pp. 126–27).

Donahue laces his historical narrative with telling details and distinctive quotes that evoke the texture and personalities of the eras he describes. In recounting how early broadcasters cultivated an impeccable public image to increase respect and stave off calls for nationalization, he reports that "NBC broadcast

---

[4] The FCC actually denied renewal because the radio station made misrepresentations in its application, but the station originally came under scrutiny because of a Fairness Doctrine complaint (p. 78).

only live entertainment, and required any entertainer performing after 6 p.m. to wear evening clothes and black tie. As a rule, NBC, CBS, and Mutual refused advertising for 'beer, wine, liquor, deodorants, depilatories, undertakers, cemeteries or any financial schemes'" (p. 28). Unfortunately, the keen journalistic ear that helps this former television news writer, associate producer, and documentary filmmaker select glittering quotes sometimes leads to superficiality. His reliance on pop star Bruce Springsteen to summarize the political pulse of America in the mid-1980's is simplistic, at best (pp. 104, 135).

Far less forgivable is Donahue's failure to support his basic conclusions. Chapter Ten looses a swarm of soap bubbles, luminous in their originality and promise but doomed because of insufficient substance. Furthermore, offhand assertions about American competitiveness in the international marketplace and the quality of all presidential candidates since Kennedy are beyond the scope of Donahue's study (p. 180). The conclusions fly thick and fast:

"In an environment of plebiscitary politics, political dealignment, and a divided government, the Fairness Doctrine provides an illusory mechanism for interest groups to influence public opinion" (p. 180).

"The longevity of the Fairness Doctrine reflects an inappropriate, groping effort to come to terms with the decline of public intellectuals in contemporary American public life" (p. 181).

Though each assertion is fascinating and rich with controversy and imagination, none lasts more than a paragraph, never long enough to let the author explore its possibilities and connect it to his otherwise admirable historical analysis. Without any grounding in the historical record, the assertions wilt most disappointingly.

Sadder still is the fate of Donahue's ultimate conclusion. Not only is it unconnected to Donahue's historical data, but it actually seems to contradict the bulk of his research. Donahue says:

> Broadcast journalism is neither so dangerous nor so persuasive that a fairness law of dubious constitutionality will accomplish an overriding public good. Although television and radio news may be pervasive, a law that denies broadcast journalists their First Amendment freedoms diminishes politics and assaults critical liberty (p. 183).

Coming after 182 pages that stress repeatedly the power of broadcasting over public opinion, that conclusion rings hollow.

Over and over, Donahue quotes a gamut of politicians who alternately have groveled before and railed against broadcast news because of its massive influence. His denial of the power of broadcast news seems even more disingenuous in the face of recent statistics showing that fifty percent of Americans rely exclusively on television for their news[5] and that the viewing public perceives it as the most credible news source, by a steadily growing margin.[6] While public reliance does not necessarily require protection by the government, it certainly raises questions about the argument that the government should remove its restrictions on the media because broadcast news poses no threat.

Donahue's final conclusion is weak primarily because it is so inconsistent with the history outlined in an ambitious and admirable book. This meticulous and fascinating historical study still is a worthy contribution to understanding continuing struggles over the electronic spectrum.

—*Rosemary Reeve*

[5] *Red Lions*, *supra* note 3, at 315 (citing TIO/ROPER, AMERICA'S WATCHING: PUBLIC ATTITUDES TOWARD TELEVISION 4 (1987)).

[6] *Id.* at 315–16.