HARVARD JOURNAL
on
LEGISLATION

VOLUME 22, NUMBER 2 SUMMER 1985

POLICY ESSAY

UNDERSTANDING CRIMINAL INVESTIGATIONS
Philip B. Heymann 315

ARTICLES

SMALL STEPS ON THE LONG ROAD TO SELF-SUFFICIENCY FOR INDIAN NATIONS: THE INDIAN TRIBAL GOVERNMENTAL TAX STATUS ACT OF 1982
Robert A. Williams, Jr. 335
THE PROPOSED UNIFORM NEW PAYMENTS CODE: ALLOCATION OF LOSSES RESULTING FROM FORGED DRAWERS' SIGNATURES
Steven B. Dow and Nan S. Ellis 399

NOTES

OF THINGS TO COME—THE ACTUAL IMPACT OF *Herbert v. Lando* AND A PROPOSED NATIONAL CORRECTION STATUTE
Dale M. Cendali 441
TELEPHONE PORNOGRAPHY: FIRST AMENDMENT CONSTRAINTS ON SHIELDING CHILDREN FROM DIAL-A-PORN
John C. Cleary 503

COMMENTS

UNION RESTRICTIONS ON THE RIGHT TO RESIGN: A PROPOSAL FOR A NEW REASONABLENESS TEST
Dennis C. Shea 551
STATE DISCRIMINATORY ACTION AGAINST NONRESIDENTS: USING THE ORIGINAL POSITION THEORY AS A FRAMEWORK FOR ANALYSIS
Adam B. Schiff 583

RECENT PUBLICATIONS 609

INDEX

INDEX TO VOLUME 22 629

Copyright © 1985 by the
HARVARD JOURNAL ON LEGISLATION

POLICY ESSAY

UNDERSTANDING CRIMINAL INVESTIGATIONS

PHILIP B. HEYMANN*

Some political issues remain high on the national agenda because they present a conflict between powerful ideals and acute public fears. Policymakers confront a persistent form of this tension when they attempt to strike a balance between civil liberties and the often competing demands of law enforcement. Wherever the commands of the Bill of Rights are less than clear, legislatures as well as courts must decide what burdens on civil liberties are acceptable in order to combat crime more effectively. The focus of these decisions is on specific investigative techniques. Which ones should be prohibited? Under what conditions should others be allowed?

Value judgments alone cannot produce intelligent answers. The focus on investigative techniques demands an extremely complicated marshaling of facts as well. One may begin with some rough judgment about how much crime is tolerable in order to enjoy a certain measure of security against intrusive police activities. But that does not go far towards answering questions about particular investigative techniques, because one must also understand the importance of each technique to the solution of particular types of crime. The police may or may not have other investigative options that are equally effective and less intrusive. The crimes for which a technique is really necessary may or may not be very dangerous.

This Essay attempts to outline the relationship between specific techniques and successful investigation of particular types of crime. The basic ingredients of my argument are simple. "Investigative situations"—defined in terms of the traces of

---

* Professor of Law, Harvard University; B.A., Yale University, 1954; J.D., Harvard University, 1960. Professor Heymann was Assistant Attorney General of the United States Department of Justice, Criminal Division, from 1978 to 1981.

The ideas developed in this article were explored over a number of months with my colleague at the Kennedy School of Government, Professor Mark Moore. Having made no effort to trace the origin of various ideas, I can only express my appreciation for his important help.

criminal behavior that are available to investigators for the solution of a crime—differ for various crimes and may differ for the same crime committed in different ways. A particular technique has a greater or lesser importance to investigators depending on the investigative situation that they confront. Each technique is also associated with certain civil liberties concerns. The focus on investigative situations allows one to assess, at least crudely, the two effects of restricting the use of a particular technique: the types of criminal behavior that will be significantly more difficult or costly to investigate, and the risks to civil liberties that will be alleviated. The key is for policymakers to identify and to understand the different categories of investigative situations when making choices about what restrictions to impose on investigations.

I. The Basic Concepts

Consider a very simple form of investigation. Assume that Jones has decided to compete with several well established haulers of trash, and that those who have tried to enter the business in the past have been threatened. A week later, two of his drivers are forcibly removed from their trucks, which are then set afire.

The crucial investigative steps at this stage are obvious. The police question Jones and the drivers about the events surrounding the crime. They can question other individuals about the activities of each of the competitors or his employees before, during, and after the occasion on which the crime occurred. Their aim is to find a persuasive fit between the offense and the history of one of the competitors or his employees, a fit that could hardly be coincidental.

The police will also question specific suspects; an unwillingness to describe their activities would confirm that the investigation is properly focused on them, a demonstrably false alibi would be strong evidence of guilt, and an admission would be nearly conclusive. The suspects can be presented to the drivers in a line-up. Probable cause to believe that evidence of a crime may be in a competitor's house or office will justify a search for physical evidence. Finally, because the present activities and conversations of the suspects may be highly probative of their past activities, surveillance by officers or informants might be

considered. If there is probable cause, electronic surveillance might also be available.

This comparatively simple example provides a useful summary of most of the major sources of information to which the police can turn at any stage of any investigation: witnesses (including the victim) to the offense or to the offender's prior actions; the suspect; physical evidence linking the suspect to the crime; and surveillance of his present activities or conversations. As additional sources of information, the police can often consult written records made by the suspect, another private party, or the government. They might also examine collected information on similar offenses and their perpetrators.

What is wanted from all these sources of information? The answer goes to the meaning of an investigation. Crimes, like any other human activity, leave traces. Offenders, like anyone else, have individual histories that also leave traces. Solving a crime requires finding a strongly persuasive mesh between the traces of an offense and the history of a particular individual. The two stories must become one. At its strongest, the solution may involve a confession or an eyewitness account, a form of proof so direct that it obscures the fact that both a period of personal history and an account of a criminal event are involved. In other cases, where circumstantial evidence is crucial, the difference between the criminal event and a personal history becomes clearer.

What are the "traces" of a crime? They vary from crime to crime, but another example illustrates the concept. Consider a robbery at gunpoint of a woman returning home from work one evening. If she notifies the police, the victim will have some traces of the crime to report: a rough description of the robber's appearance, when and where the crime occurred, how it was done, and so forth. Police officers will interview other witnesses and may try to determine if there was someone in the neighborhood who met the description. If identifiable property was taken, it may leave a trace in the hands of fences or others. Other traces of the robbery may take time to arise.

Individuals have personal histories that also leave a wide variety of traces. An individual has friends, relatives, and acquaintances who know what he looks like, when he was or was not in a particular place, whether he is employed or unemployed, and with whom he is generally seen. The police may know the names of individuals with criminal records involving

similar crimes. An individual who has committed a crime may have discussed it with others beforehand; he may talk to others in the future about it. If a large sum of money was taken, elevated levels of spending may leave a trace.

There are many ways to match traces of a crime obtained from one source with traces of personal history coming from another source. The police may circulate a description, an artist's drawing, or a picture taken by a surveillance camera. They may invite the victim to look through a "mug book" containing the pictures of suspected robbers. They may talk to fences or informants who circulate in a world where discussions of robbery would likely take place or where traces of the proceeds would easily be recognized. The police may simply happen upon physical evidence in the course of performing other duties.

Ongoing criminal activity, such as provider fraud in the provision of medical services paid for by Medicaid, operation of a large narcotics ring, or political corruption, leaves different traces but requires the same effort to match traces of criminal activity with the personal history of an individual. An investigation may first focus either on a specific crime or on an individual suspected of ongoing criminal activities. The latter focus, which reverses the more familiar pattern, primarily involves identifying the crimes that can be matched with the individual's personal history.

To detect the traces of any crime, the police have available only a limited set of steps that can be combined, in a more or less considered sequence, into a specific investigative plan. (1) They may interview willing witnesses or seek to have unwilling witnesses testify under legal compulsion. (2) They may question the suspect under more or less coercive conditions. (3) They may view and analyze publicly available physical evidence or use legal authority to search in private places. (4) They may review publicly available or voluntarily produced records or seek to have other records produced under legal compulsion. (5) They may engage in physical or electronic surveillance of the suspect's activities. (6) They may use informants or offer rewards. (7) They may use undercover operations. This list would be about the same in any political system, because I have not yet included legal and administrative limitations in the analysis.

To develop an "investigative plan," the police must select particular techniques and arrange these choices into an orderly

sequence of steps for investigating a particular crime. Devising that plan requires considering several factors: the cost of the contemplated step, the resentment it provokes, the traces expected from the particular type of crime, the indications from information already available suggesting relatively likely traces not yet discovered, the risk of destruction of traces if the step is taken too early or too late, and so on.

Most important, legal rules limit and regulate the use of some investigative steps. Thus, an investigation generally may not *begin* with a search or electronic surveillance; these steps can only be taken after there is sufficient information, gathered in other ways, to constitute probable cause.

When a particular technique is forbidden, the police must develop an alternative investigative plan that incorporates a "second best" substitute for the prohibited technique. Investigative steps are often substitutes for each other. Officials can pursue traces of a crime held by a suspect's close associates by using informants or by calling those associates before a grand jury and by threatening them with contempt. Evidence can be obtained by a search warrant or by a grand jury subpoena. A suspect can be interrogated by the police or stimulated to speak on a phone that has been tapped.

Where there are not close substitutes for a forbidden investigative step, the chances of solving a crime will decrease unless more investigative resources are put into the effort. The probability of successfully solving a crime is initially a function of four factors: (1) the traces that are left by a particular crime and its perpetrator; (2) the willingness of private individuals to call these traces to the attention of the police; (3) the investigative resources (time, money, and equipment) devoted by officials to the particular crime and the intelligence with which the resources are used; and (4) the activities that the police are permitted to engage in while using these resources. Of these four factors, the police exercise some control only over the allocation and productive use of resources.

An important part of the investigative agency's job is to decide how much, if any, of its investigative resources to allocate to any particular investigation, in light of the estimated probability of success. If the probability is small and the crime ordinary, it may be wise to limit severely the investigative resources allocated to this crime. Investigations of other crimes will take priority over those formerly relying on the use of the forbidden

activity. Thus, forbidding an investigative technique means not only creating a need for new investigative plans for particular crimes, but also changing the investigative agency's broad strategy or portfolio of criminal investigations.

II. Four Problematic Investigative Situations

The hypothetical extortion of Jones the trash hauler, which I discussed earlier, exemplifies one particular, fortunate type of investigative situation. By "investigative situation," I mean the set of characteristics that defines the likelihood of successfully solving the crime and that determines the steps that must be taken to do so. The Jones example presents the ideal investigative situation because, from the beginning, the police possess everything necessary for a successful investigation. (1) The investigators had been told that a particular crime had been committed at a specific place and time. (2) They have a limited list of suspects. (3) They have witnesses who are willing to help solve the crime and to testify at trial. In this fortunate situation, the investigators can utilize most of the possible investigative techniques, as they see fit, to develop the necessary information.

Many other crimes create investigative situations where one or more of these characteristics is missing. In these problematic investigative situations the full array of possible investigative techniques for developing the information necessary to solve the crime is not available as a practical matter. A judicial or legislative restriction on the use of one of the relatively few remaining helpful techniques thus has far more serious consequences in these problematic investigative situations than in the fortunate ones.

There are four major investigative situations that depart significantly from the paradigmatic extortion case with which we began. One can identify particular types of crime with each of these problematic situations. The first two involve contexts in which investigators do know, as they did in the paradigmatic case, of a particular crime that was committed at a certain place and time. The problems arise elsewhere. In the first situation, the police cannot narrow the list of suspects in a way that will enable them to compare the personal histories of a limited number of suspects with the events known about the crime. This problem is endemic to burglary. In the second situation, the

investigators are aware of a crime and have a small list of possible suspects, but are unable to obtain witnesses who will help resolve the matter and testify at trial. A familiar example is an extortion case where the victim has been intimidated. Crimes committed by narcotics organizations or other organized crime will also frequently have this problematic quality.

The other two problematic situations are characterized by investigations that are initially undertaken without knowledge of a specific crime that took place at a particular place and time. In the third situation, the investigators suspect an organization or individual of ongoing criminal activity. Although they frequently know the general type of crime, they do not even know the place or time of any specific crime. This situation is also frequently found in investigations of organized crime or narcotics rings.

In the fourth and final problematic investigative situation, the investigators have neither an identified suspect nor knowledge of a specific crime. They have reason to believe that there is an ongoing type of crime involving some members of a particular group of people (for example, cocaine is being bought and sold at certain locations, or bribes are paid for construction contracts at others). But there is no one to report when a specific crime is committed, and there are no clearly identified suspects. Crimes of vice and corruption often present this problem.

A. *Investigative Situations Lacking a Relatively Narrow List of Suspects*

Most people and organizations categorize information and memories in terms of known, identifiable individuals. If one imagined this entire set of memories being collected and filed centrally, one might picture an immense inventory of detailed information about individuals. Under each heading, there would be significant information about the personal history and characteristics of an individual. By systematically going through the histories of possible suspects, the police, in this fantasy world, could theoretically solve any reported crime if they had only a fair description of its circumstances.

The problem lies elsewhere: no one collects all this information centrally and, of course, there are many very good reasons why we would not want it to be collected. Even if it were

collected, the task of reviewing all possibly pertinent files would be close to impossible if the police had merely a rough description of the perpetrator and the time and place of the crime.

A crucial step in solving any crime is to create "files" on suspected individuals that did not previously exist. To do so, investigators must isolate a relatively small group of suspects as to whom traces of personal history, around the time of the crime, can be collected. Investigators, hoping to find a suggestive match, will have to compare what they know about the crime with law enforcement records and the memories of private individuals. But, as a practical matter, they cannot create such files on more than a few individuals. In the easiest case, only a few individuals could have been present at the time of the crime or could have known of the particular criminal opportunity that was exploited. The case is far more difficult to solve and may not be worth pursuing if the number of individuals whose histories might match the traces of the crime is very large.

An identification of a short list of suspects will often accompany awareness of the crime. For example, a patrolling officer may witness a crime and apprehend a suspect; and information about offenders is often exchanged by one of their associates who is seeking a reduced sentence. But many burglary and robbery victims have no idea who the suspects are. In very important classes of crime the offenders will be unseen, disguised, or anonymous, even though the crime itself will be known. There will be no short list of suspects whose personal histories can be compared with what is known about the crime.

What are the alternative investigative strategies here? The police themselves may be able to generate a relatively narrow list of suspects out of collected intelligence regarding who has committed similar offenses in the past. Thereafter, the police may display the suspect's picture (a mug shot) or person (in a line-up) to victims or other witnesses. They may also check the suspect's activities around the time of the crime in an effort to find a suspicious mesh with what is already known. If police intelligence cannot produce a short list of suspects, it is occasionally possible to piece together from witnesses a description or drawing of the offender, which can then be widely circulated in the hope that citizens will recognize it.

If the crime appears to be one instance of a recurring activity, such as one of a number of burglaries taking place in the same neighborhood, secret or undercover patrol of the area can iden-

tify a suspect. Alarms or sophisticated viewing equipment would have the same effect. The chance of detecting the perpetrator can be further increased by making one target of crime unusually attractive and focusing observation on that target. For example, word could be let out that a particular house contained unusually valuable and transportable goods, and a window to it could be left open.

Investigators can also organize the use of informants around the particular crime or activities. Pawnshop owners may be enlisted as informants because they are likely business associates of persons attempting to sell recently stolen property. Other informants could be sent into any network of individuals who would be likely to know of an unusual robbery. If the robbery or burglary were exceptionally remunerative, one might ask informants to watch for a sudden increase in wealth among individuals who commonly engage in illegal activities. Finally, investigators can offer sizeable rewards, in the hope of inducing people who are not presently informants to bring forth relevant evidence. None of these strategies has a high chance of solving a crime committed by the unseen perpetrators of a burglary or by disguised robbers.

B. *Investigative Situations Lacking Willing Witnesses*

Witnesses who are willing to reveal the existence of a crime, to provide leads as to its major suspects, to help in the investigation, and to testify at trial, are without doubt the most crucial investigative resource. Without willing witnesses, it may be impossible to learn of certain crimes or to develop a limited list of suspects. Even if these needs can be met, identifying proper suspects and proving guilt at trial will generally require live witnesses to the criminal acts. When witnesses are unwilling to help, out of fear, indifference, or loyalty to the suspect, investigators have four principal ways to induce them to aid focused investigation or to testify at trial.

First, investigators can offer an unwilling witness protection or a reward for providing the requested information or assistance. Some rather mundane programs are designed to eliminate much of the personal irritation and inconvenience that come from testifying at trial. If fear discourages testimony, the witness can either be protected by guards or be given a change of

identity. Alternatively, officials may be able to use the witness's information without revealing his identity by obtaining a warrant to search or to wiretap or by introducing an undercover agent to the suspects and then relying on the agent to give needed testimony.

The second method for obtaining the cooperation of unwilling witnesses is coercion. A prosecutor can call the witness before a grand jury and require her to testify on pain of civil or criminal contempt. If the prosecutor has charged her with another crime, or if she is vulnerable to such charges, the witness can be threatened with more serious consequences if she does not testify.

Third, investigators can deceive an unwilling witness or even one of the suspects into providing the information. An informant or an undercover agent can elicit the necessary information from the witness or suspect. Electronic surveillance can accomplish the same results.

Finally, if the criminal conduct addressed by the investigation is planned for the future or is an ongoing activity, undercover agents can compensate for the absence of willing witnesses by offering to participate in the activity as victims, customers, or other business associates. In this case, conduct occurring after the introduction of the agents provides the evidence to support the criminal charges.

One type of criminal activity, organized crime, presents starkly the problem of the unwilling witness. Intimidation or organizational loyalty provides insulation for the higher levels of criminal organizations whose street level activities can be investigated with considerably less difficulty. Reaching the higher levels requires the use of one of the four investigative techniques just discussed.

C. *Groups Whose Specific Crimes Are Unknown But Which Are Believed To Be Engaged in Organized Criminal Activity*

A central function of law enforcement is to reduce the amount and impact of serious criminal behavior. To achieve this goal, the police do not rely solely on the general deterrent effect of investigating specific, discovered crimes. If a particular organization or group of offenders is believed to be committing a large

number of significant but unspecified crimes, the incidence of those crimes can be reduced substantially by incapacitating the offenders through imprisonment. The best strategy may be to monitor the activities of suspected offenders or to engage in observable criminal activities with the offenders in undercover operations.

The line between investigation of offenders and investigation of historic crimes is clearer in principle than in practice. For example, hiring an informant from among those closely connected with active criminals has the foreseeable effect of disproportionately patrolling these individuals' activities, even if the purpose of the investigation was only to discover who committed a particular past crime. Similarly, investigative activity focused on suspected career offenders furthers the solution of a particular crime when, having been alerted to a crime, law enforcement authorities must produce their own list of suspects to compensate for the lack of a list generated by witnesses.

If no particular crime is under investigation, the primary purpose of investigating criminal organizations or professional offenders is plainly incapacitative. In such a situation, because there is no specific crime to limit the amount of the suspect's recent history that is of interest, and because there is no probable cause, only the use of informants and undercover operations are promising investigative techniques. Wide-ranging review of records and sweeping grand jury inquiries are possible but unlikely substitutes.

D. *Investigative Situations Where There Are Neither Suspects Nor Complaining Parties To Alert Investigators That a Crime Has Been Committed*

Professor Mark Moore has identified four situations in which neither the victim nor another witness is likely even to notify the law enforcement authorities of a crime.[1] Frequently, the very existence of the criminal acts will be unknown to the authorities. First, the crime may involve only willing participants. Narcotics transactions, bribes of public officials, and

---

[1] Moore, *Invisible Offenses: A Challenge to Minimally Intrusive Law Enforcement*, in ABSCAM ETHICS: MORAL ISSUES AND DECEPTION IN LAW ENFORCEMENT 21–22 (G. Caplan ed. 1983).

much of organized criminal activity fall into this category. Second, a victim cannot be relied upon to play his customary role of giving notice of criminal acts if he is intimidated by the offender or, more rarely, fears the effects of revelation as they work their way through the law enforcement system. Intimidation of witnesses is common in the areas of extortion, loan sharking, and labor racketeering, particularly where traditional organized crime is involved. Victims may not complain of blackmail or sexual abuse because they fear the consequences of revelation or publicity.

Third, a victim may not know that he was harmed or that the harm was illegal. He may also doubt that law enforcement authorities will take the crime seriously enough to respond. The victim of a crooked gambling game or the victim of mob misuse of a union pension fund falls into this category. White collar offenses, such as insurance fraud, tax evasion, and counterfeiting, which distribute losses across a large number of individuals, also harm unknowing victims. Bribery and other political corruption involve illegal uses of public authority and resources that may not produce any victim (such as a competitor) who notices that he has been harmed. Environmental offenses such as the illegal disposal of hazardous wastes fall into the same category, because victims may notice their injury long after the incident. Finally, conspiracies to commit *future* criminal acts, even matters as serious as planned terrorist bombings, do not *yet* have victims to inform the police.

To be profitable over time, a highly organized and ongoing criminal enterprise requires some such means of invisibility. Complaining victims allow the police to patrol and to set traps and may themselves take protective steps, such as carrying weapons or installing alarms. For example, robbery is a dangerous enterprise for offenders as well as for victims. Invisible crimes, by contrast, do not lead to private precautions or police response. The perpetrator can look forward to a long career unless law enforcement authorities develop substitutes for willing complainants.

What unusual efforts can investigators take to deal with invisible crimes? Aside from the occasional "good citizen" report and prosecutorial deals, the only general strategy capable of detecting invisible crimes involves a broad notion of "patrol." The police will have to review the possible sources of the traces

of criminal activity, without any predicate basis for believing that a particular crime has been committed. The uniformed patrol, on foot or in squad cars, of crime prone areas is the most familiar example of this strategy. There is nothing essentially different about the patrolling done by the computers in the Department of Health and Human Services, which review doctors' bills to identify unusual Medicare and Medicaid claims.

The use of informants who can observe crimes in places where only the informants are trusted extends the reach of government patrol. This activity is actually similar to more familiar forms of patrol. For example, the street patrol of an officer becomes very much like that of an informant if the officer is out of uniform or undercover. Creating the opportunity to commit crimes on, or in association with, law enforcement officers also extends the reach of the patrol.

Review of third party records, the use of informants, and undercover operations are all forms of patrol that can compensate for the lack of willing complainants or other volunteers with information about ongoing criminal activity. Yet they rely on different traces of criminal activity. Private businesses that bribe government officials are likely to leave traces in their books, records, and tax returns. Organized crime groups that extort terrified victims are not likely to do so. The use of informants is crucial in the latter category and less essential in the former. In the case of public corruption, each of these three techniques is a substitute for the other. But only the use of informants and undercover operations are likely to produce evidence of a narcotics conspiracy.

III. The Practical Limitations Affecting the Choice of Investigative Steps

In analyzing these problematic investigative situations, we were tracing the effect of two factors on successfully solving a crime: namely, the type of traces left and the willingness of victims and witnesses to call these traces to the attention of the investigator. The chance that a crime will be solved also depends upon a third factor: the amount of time, energy, and other resources devoted to that particular case. Utilizing scarce resources too freely on one case means that investigators cannot

handle other cases effectively. Choosing investigative techniques necessarily entails judgment about these relative costs.

Because it is a good working assumption that *any* crime could be solved with unlimited resources, the problem for law enforcement officials is to allocate limited resources in fair, efficient, and justifiable ways. Inexpensive or plentiful resources can be used freely even where there is great uncertainty about their effectiveness. Expensive or scarce resources must be allocated to situations where success is more likely or more important. Responsible law enforcement authorities cannot ignore these allocative costs and probabilities.

Many techniques will be too costly for certain uses. Because there are no specific allegations of crime or suspects, the investigative strategy of patrol requires observing many places and many people. It cannot justify the use of expensive investigative techniques with regard to any single place or person. Thus, even if there were no relevant constitutional prohibitions, the widespread use of electronic surveillance to patrol for crime would be too costly. The same is true of large-scale undercover operations and detailed review of massive systems of records. Similarly, techniques such as the detailed auditing of financial records require resources (like trained financial investigators) that are often either unavailable or extremely scarce.

Decisions about the use of investigative techniques also have costs in terms of public support. Effective law enforcement requires more than dollars and personnel. Other necessary resources include citizen cooperation and legal authority. Legislatures provide the legal authority to use various investigative techniques. Voluntary private cooperation with law enforcement authorities bears importantly on how effectively alternative investigative techniques can be used. Wise law enforcement authorities recognize the importance of these resources and know that their availability depends, in turn, upon how well the public believes the authorities are managing two potentially conflicting goals: reducing or punishing crime, and alleviating the fears and sensitivities of citizens about government investigative power. The conflict is clearest when deeply intrusive techniques must be used to solve gravely serious crimes. Of course, the public may widely appreciate this problem. Officials may be unable to maintain support, however, when deeply intrusive techniques are used to investigate crimes the public does not consider serious.

IV. Assessing the Costs to Law Enforcement of a Particular Restriction on the Use of an Investigative Technique

The importance to law enforcement of a particular investigative technique depends upon two facts. First, it turns on the frequency of the particular investigative situations in which the technique is helpful and economically feasible, and in which there are few available substitutes. Second, the law enforcement need for the technique in terms of broader social purposes depends on the types of criminal activity that are associated with those investigative situations.

A similar analysis, modified only to reflect the continued availability of the technique in some investigative situations, helps to determine the impact on crime control of a restriction on a particular investigative technique. To illustrate how policymakers should utilize the foregoing analysis of the problematic investigative situation, I will take as an example undercover operations and the proposal for a restriction requiring that a factual predicate be established before the undercover technique can be used.

A. *The Civil Liberties Problem*

An undercover operation involves the encouragement, by either a disguised employee of an investigative agency (an undercover agent) or an informant, of other individuals to engage in criminal conduct for which they will then be prosecuted. It can take one of two basic forms. The undercover agent or informant may either pose as a potential victim of a particular crime or as a potential partner in a criminal enterprise. An example of the former is the policewoman disguised as an aged woman walking in a dangerous park area, ready to arrest anyone who attempts to rob her. An example of the latter is the agent disguised as a drug user who deals with a seller.

What makes the use of undercover operations problematic? They do not inherently intrude into areas of privacy or trusted associations. To the extent that they do, they are likely to be sharply focused on evidence of crimes. The central concern with undercover operations lies elsewhere: they are intended to encourage the commission of a particular crime. Believing that

the same suspects are committing similar crimes but leaving no traces for investigative use, investigators attempt to produce an observable instance of criminal activity, which will then be prosecuted. By definition, the operation increases the probability that a crime will take place on a particular occasion.

If a substantially similar crime would have taken place within some relatively short period of time without governmental instigation, it is unimportant that the government agent determined the timing and location of the crime. But if the actions of law enforcement authorities substantially increase the probability that the suspect will engage in *any* similar crime within a reasonable time, it is also more likely that an undeserving individual will be prosecuted and punished. Even if entrapment technically does not occur, this concern is inherent in undercover operations.

B. *A Proposed Restriction*

One proposed resolution of these competing concerns is to require some form of factual justification or predicate before an undercover operation can be approved. Modeled after the Fourth Amendment, this method of protecting civil liberties is familiar and time tested. The reasoning is straightforward.

Investigations involve an assertion of governmental authority and a power to invade the privacy of places, activities, and associations. They are therefore of paramount concern to any nation that watches with care the powers of its government. If one is to respect *both* the need to investigate crimes *and* a deep concern for the privacy of places, activities, and associations, a promising possibility is to require a factual justification or predicate before the government may take investigative actions of various types. One need only decide how much evidence of any crime or of some particular crimes should be required before the specific technique may be used.

A predicate is a factual basis for believing that an individual is involved in criminal activity or that evidence is in a particular place. The standard could be probable cause, reasonable suspicion, or something else. The requirement might involve a prior judicial determination that the predicate was met, or it may be left to the initial judgment of law enforcement officials. These variations make a difference, but all forms of a predicate requirement raise a common concern for those charged with carrying out successful investigations.

The matter can be stated generally. Depending upon the type and amount of evidence required, a predicate may completely preclude the use of an investigative technique for one or more of the four problematic investigative situations discussed above. If the predicate required is a "reason to believe" that a particular individual has committed a crime, the technique burdened with that requirement cannot be used to narrow the range of suspects. If the requirement is that there must be some evidence of particular crimes, the technique so burdened may be unavailable for investigating individuals or organizations that are suspected of pursuing crime as a full time business, but whose specific activities are as yet unknown.

The problem of investigating "invisible offenses" gives rise to the most serious situation in this regard. There is a special relationship between the predicate requirement and the discovery that invisible offenses have taken place. By definition, invisible offenses are crimes that are unusually susceptible to having their traces limited to private areas. It is precisely these areas, however, that a predicate requirement forbids investigators from entering until they already know something about the offense.

The system structured by the Fourth Amendment requirement of a predicate identifies two spheres of privacy: (1) places where one has a reasonable expectation of privacy; and (2) situations in which the speaker has a reasonable expectation that no one is listening to the conversation without the consent of any party to the conversation. This system forbids investigations into these private areas without some reason to believe that there is evidence of crime within them. The "reason to believe" in these situations must be established from traces of the crime that are available in nonprivate areas. To the extent that a criminal can limit the location of the traces of his crime to areas private to him or to his loyal (or intimidated) associates, a predicate requirement precludes successful investigation. In most cases, only the difficulty of keeping all traces in private areas and all associates loyal renders the predicate requirement compatible with adequate levels of law enforcement.

C. *The Costs To Law Enforcement*

Finally, I return to the present role of undercover operations in the four problematic investigative situations. Undercover is

most important to the fourth situation—discovering offenses that can be made invisible by eliminating any risk that witnesses will willingly report them. Corruption and narcotics trafficking are the most important crimes associated with this investigative situation. A predicate requirement such as that proposed by the Senate Judiciary Committee—demanding only reasonable suspicion that a particular type of criminal activity will be detected—would not prevent the use of undercover operations in these situations.[2] More rigorous predicate requirements, however, could altogether eliminate the use of undercover operations in even the most pressing situations.

For the purpose of patrolling for crimes that cannot be observed and that will not be reported by participants, there are few alternatives to undercover operations. The police will occasionally locate informants who are in a position to observe or hear reports of the crime. On other occasions, such as governmental corruption or fraud, the most promising substitute is a massive, time consuming, and expensive review of available records. Computer matching systems, pioneered by the then Department of Health, Education, and Welfare, economically patrol government records for evidence of program fraud. Narcotics investigations have turned, at greater cost, to bank records to identify likely drug transactions in situations identified as having a high probability of turning up criminal activity. At still greater expense, the Ward Commission in Massachusetts successfully conducted extensive, computer assisted analyses of financial records of particular classes of contractors dealing with the state to find evidence of suspicious cash movements. Again, the technique was practical only because other information pointed to a relatively narrow area for this relatively expensive patrol.

Insistence on a prior showing of probable cause or even reasonable suspicion that a particular individual was engaging in a particular form of criminal behavior would also foreclose crucial uses of the investigative technique in other problematic investigative situations. Familiar undercover operations include setting up a purported "fencing" operation to catch unknown prop-

---

[2] *See* SENATE SELECT COMMITTEE TO STUDY UNDERCOVER ACTIVITIES OF COMPONENTS OF THE DEPT. OF JUSTICE, FINAL REPORT, S. REP. NO. 682, 97th Cong., 2d Sess. 27–29 (1982); SUBCOMM. ON CIVIL AND CONSTITUTIONAL RIGHTS OF THE HOUSE COMM. ON THE JUDICIARY, 98TH CONG., 2D SESS., FBI UNDERCOVER OPERATIONS, 53–61, 75–85 (Comm. Print 1984).

erty thieves or a large warehouse to catch unknown hijackers. A restriction that forbade such encouragement to would-be thieves and hijackers would be costly in the first problematic investigative situation that I discussed, in which crimes are known but no "short list" of suspects exists. In the related form of undercover operation, in which police disguise themselves as attractive victims of predatory crimes, such a stringent predicate requirement would also deny law enforcement a technique with few substitutes. In none of these investigative situations is there a specific factual predicate in terms of particular suspects when the undercover operation begins; there is only a general pattern of criminal activity.

Undercover operations are important in the remaining problematic situation—compensating for the lack of a willing witness. In this context, however, a predicate requirement would not be disabling. Often a participant-informant will advise law enforcement officers that a particular organization is engaged in an ongoing criminal business; but danger to the informant, problems of hearsay, or witness credibility problems may preclude reliance on this source at trial. At that point, the only feasible option is an undercover operation. Specifically, the alternatives are electronic surveillance of, or monitored participation in, new instances of the criminal operation by the informant himself, or by involvement of an officer introduced to the business by the informant. In this situation, the predicate requirement is easily met.

In sum, striking a sensible balance between civil liberties and effective law enforcement requires understanding the impact that particular investigative restrictions will have on the solution of particular types of crime. Some types of crime are particularly difficult to investigate because their perpetrators are able to restrict sharply the traces of criminal activity they leave behind. From an investigator's point of view, the few techniques that can pick up traces of these crimes are unusually important, at least if the criminal activity cannot be safely ignored. In the four central problematic investigative situations, forbidding particular techniques may be tantamount to sanctioning high levels of the types of crime associated with those situations; alternatives may simply be unavailable, too costly, or too intrusive. Some forms of restriction will preclude essential uses of the technique.

Public policymakers, when considering proposals to restrict investigative techniques, must recognize that the effective so-

lution of particular crimes may be dependent upon the use of a particular technique. However difficult, the choice between restricting particular techniques and tolerating certain types of crimes must be made explicitly. Understanding the nature of individual investigative situations helps to clarify what is at stake. In the example of a predicate requirement for undercover operations, inquiry showed that an appropriately tailored restriction, such as that proposed by the Senate Judiciary Committee, would not unduly hamper law enforcement in certain crucial investigative situations. For most policymakers, the decision will be far more difficult where the proposed restriction leaves investigators without viable alternatives in the problematic investigative situations characteristic of dangerous or organized crime.

ARTICLE

SMALL STEPS ON THE LONG ROAD TO SELF-SUFFICIENCY FOR INDIAN NATIONS: THE INDIAN TRIBAL GOVERNMENTAL TAX STATUS ACT OF 1982

ROBERT A. WILLIAMS, JR.*

In enacting the Indian Tribal Governmental Tax Status Act of 1982, Congress attempted to eliminate economic obstacles to tribal self-sufficiency by granting Indian tribes a tax status similar to that of state and local governments. This status provides Indian tribes with the ability to implement previously unavailable fiscal initiatives. It is applicable, however, only on a limited basis.

In this Article, Professor Williams argues that while the Indian Tribal Governmental Tax Status Act is an important step toward self-sufficiency, Indian tribes should be given the same revenue raising authority and tax advantages that other governmental entities enjoy. He recommends that Congress amend the Act to make governmental tax status available to tribes under all sections of the Internal Revenue Code. He also recommends the creation of an American Indian financial institution and the grant of authority for Indian tribes to issue Industrial Development Bonds.

> [T]his Court has recognized the distinctive obligation of trust incumbent upon the Government in its dealings with [Indians] . . . . Under a humane and self-imposed policy which has found expression in many acts of Congress and numerous decisions of this Court, it has charged itself with moral obligations of the highest responsibility and trust. Its conduct . . . should therefore be judged by the most exacting fiduciary standards.[1]

The road to economic and social development for Indian Nations[2] in the United States is impeded by an intractable host of tangible and intangible barriers. Territorial remoteness, an

---

* Assistant Professor of Law, University of Wisconsin. A.B., Loyola College, 1977; J.D., Harvard University, 1980. Member, Lumbee Indian Tribe. The author gratefully acknowledges the helpful comments of Suzan Shown Harjo, Elmer Savilla, Tommy Thompson, Richard West, and William Whitford.

[1] Seminole Nation v. United States, 316 U.S. 286, 296–97 (1942) (citations and footnote omitted).

[2] In this Article, "Indian Nations" is used as the collective term for all Indian tribes, as this term better reflects their status as independent entities. "Indian Country" is used as the collective term for those areas where Indians live. The author is aware that some prefer other terms than these, but feels that these terms best describe the situation.

inadequate public infrastructure base, capital access barriers, land ownership patterns, and an underskilled labor and managerial sector combine with paternalistic attitudes of federal policymakers to stifle Indian Country development and investment.[3] The design of programs and policies to assist Indian people in successfully mitigating these barriers to economic and social self-sufficiency remains the greatest and most difficult

---

[3] On the barriers to Indian Country economic development, see *Oversight of Economic Development on Indian Reservations: Hearing Before the Senate Select Comm. on Indian Affairs*, 97th Cong., 2d Sess. 9–12 (1982) (statement of Joe DeLa Cruz, President and Chairman, Quinault Indian Nation) [hereinafter cited as *Oversight of Economic Development on Indian Reservations*]; TASK FORCE SEVEN, AM. INDIAN POLICY REVIEW COMM'N, 95TH CONG., 1ST SESS., REPORT ON RESERVATION AND RESOURCE DEVELOPMENT AND PROTECTION 1 (Comm. Print 1976) [hereinafter cited as TASK FORCE SEVEN]; Ickes, *Tribal Economic Independence—The Means to Achieve True Tribal Self-Determination*, 26 S.D.L. REV. 494 (1981). *See also* PRESIDENTIAL COMM'N ON INDIAN RESERVATION ECONOMIES, REPORT AND RECOMMENDATIONS TO THE PRESIDENT OF THE UNITED STATES (November 1984) (limited survey of conditions hindering private sector investment in Indian Country) (on file at HARV. J. ON LEGIS.) [hereinafter cited as PCIRE].

For Indian perspectives on the barriers to tribal development, see AMERICAN INDIAN ECONOMIC DEVELOPMENT (S. Stanley ed. 1978).

A survey of the above literature indicates that the major barriers to development for Indian communities, outside the control of tribes themselves, are those outlined in the text. This Article, therefore, focuses on these barriers as constituting the major obstacles that any comprehensive federal policy for Indian Country development must initially address before attempting to cope with more specific problems particular to one or more Indian communities.

Obviously, the problems of housing, education, nutrition, and health in Indian Country are immense. *See* Nat'l Tribal Chairmen's Ass'n, The State of the American Indian Nations: 1983 (June 1983), (position papers E, F, and G) (on file at HARV. J. ON LEGIS.) [hereinafter cited as NTCA Statement]. For example, the health of the Indian population is generally worse than that of the population as a whole. Indians suffer from tuberculosis, chronic liver disease, accidents, diabetes, pneumonia, influenza, cardiovascular disease, pulmonary diseases, and suicide at a far greater rate than the general population. Alcoholism related deaths are about five times the national average. PCIRE, *supra* pt. 2, at 85. These problems contribute to the structural underdevelopment that plagues Indian Country and require solutions specifically aimed at each set of problems. The barriers discussed in this Article are those that lend themselves to discussion under the traditional rubric of "economic development problems." They are thus theoretically amenable to solutions focusing specifically on devices traditionally understood as economic development incentives, such as those contained in the Indian Tribal Governmental Tax Status Act of 1982, Pub. L. No. 97-473, 96 Stat. 2607 (codified as amended at 26 U.S.C. § 7871 and in scattered sections of 26 U.S.C. (1982 & Supp. I 1983)) [hereinafter cited as the Tribal Tax Status Act].

The scope of this Article therefore has been arbitrarily limited in the interests of time and space. In order to understand fully Indian Country underdevelopment, the reader should refer to a comprehensive overview of the many structural barriers plaguing Indian Nations. *See, e.g.*, AM. INDIAN POLICY REVIEW COMM'N, 95TH CONG., 1ST SESS., FINAL REPORT (Comm. Print 1977) (surveying conditions in Indian Country and proposing broadly based recommendations) [hereinafter cited as FINAL REPORT]; PCIRE, *supra*. The criticisms and suggestions contained in this Article represent a purposely narrowed perspective on the immense dilemmas presented by Indian Nation underdevelopment, and should be understood and used as such.

challenge faced by the United States government in the execution of its trust responsibility to Indian Nations.[4]

In the past, federal policies for Indian Country development have either failed to recognize fully and to eliminate completely these interrelated barriers or have actually exacerbated the problems encountered by tribal governments in their efforts towards self-sufficiency.[5] Today, Indian Country development policies emanating from Washington, D.C., continue to follow similar misdirected paths. The Reagan administration's policy is premised principally upon the simplistic yet chimerical belief that the "avenue of development for many tribes" lies in the supposedly abundant natural resources underlying land that the

---

[4] The federal government's trust responsibility to Indian Nations arises from treaties, congressional statutes, and court decisions:

> The birth of the trust responsibility lies in the nature of treaties. In exchange for land concessions to the Federal Government, Indians were to receive protection from foreign nations, hostile Indian tribes, and individuals. Since only the Federal Government could extinguish aboriginal title, protecting Indian tribes came to be regarded as a trust; that is, the fiduciary responsibility to protect.
>
> Unfortunately, there was no one to protect the Indians from the Federal Government. The United States, by its own laws, illegally seized Indian lands and forced Indians onto what were then submarginal lands. Because others were prevented from extinguishing aboriginal title, essentially, the trust responsibility meant that the Indians were holding the land in trust for the United States, until it wanted it.
>
> Protection of Indian lands and people meant that the Federal Government should protect them from others. But in reality the Federal Government exercised control over Indian attempts to use what resources were left to them. Control of Indians could mean two things. It could mean that the Federal Government would benignly protect them from outsiders, leaving internal control with the Indians, or it could mean internally controlling Indian action. Title 25 of the United States Code, particularly section two, is the latter interpretation taken by the Executive and Congress. Only the Judiciary has consistently taken the other interpretation. The Judiciary has tried to hold the Government to its word.

TASK FORCE SEVEN, *supra* note 3, at 1.

[5] *See id.* at 1–2.

> [T]he current system of exercising the trust responsibility by the Federal Government perpetuates economic dependence. The three necessary conditions for economic development are control, capital, and management. Under the present system these three conditions are not being met.
>
> . . .
>
> . . . [T]he disastrous economic consequences suffered by Indian tribes [is] due to preemptive Federal control . . . . [T]he Federal Government has caused this state of affairs and is responsible for stultifying Indian action. The current interpretation of the trust responsibility has created barriers to reservation development.

Id.

nineteenth century American government regarded as useless.[6] Unfortunately, such a policy ignores the fact that fewer than one in eight Indian Nations has energy and mineral reserves that can be developed.[7] In the minds of many Indian leaders and their people, federal policies premised upon such a belief implicitly condone a neocolonial status for those tribes fortui-

---

[6] Statement by the President on Indian Policy, 1 PUB. PAPERS 96 (Jan. 24, 1983) [hereinafter cited as Reagan Indian Policy Statement].

At the heart of the Reagan Indian Policy Statement rests a belief in the ability of private market forces to solve the myriad problems of Indian Country underdevelopment:

> With regard to energy resources, both the Indian tribes and the Nation stand to gain from the prudent development and management of the vast coal, oil, gas, uranium and other resources found on Indian lands. As already demonstrated by a number of tribes, these resources can become the foundation for economic development on many reservations, while lessening our nation's dependence on imported oil. The Federal role is to encourage the production of energy resources in ways consistent with Indian values and priorities. To that end, we have strongly supported the use of creative agreements such as joint ventures and other nonlease agreements for the development of Indian mineral resources.
>
> It is the free market which will supply the bulk of the capital investments required to develop tribal energy and other resources.

Id. at 98.

Indian tribal government leaders and spokespeople have universally rejected the thrust of the Reagan Indian Policy Statement. *See, e.g.*, NTCA statement, *supra* note 3; *see also infra* notes 7–9 and accompanying text.

Though estimates differ, there is universal agreement that Indian resource holdings are varied and extensive. Coal, molybdenum, oil, gas, and uranium are heavily concentrated on Indian reservations. *See* COMPTROLLER GENERAL'S REPORT TO THE SENATE COMM. ON INTERIOR AND INSULAR AFFAIRS, 94TH CONG., 2D SESS. MANAGEMENT OF INDIAN NATURAL RESOURCES 78 (Comm. Print 1976) [hereinafter cited as COMPTROLLER GENERAL'S REPORT]; TASK FORCE SEVEN, *supra* note 3, at 46–47. One commentator claims that over 60% of the nation's energy resources may be located on Indian controlled lands in the western part of the United States. Means, *Reagan Policies Force Indians to Give Up Traditional Views*, INDIAN TRUTH, Apr. 1983, at 7, 7 (available from the Indian Rights Ass'n, 1505 Race St., Philadelphia, Pa. 19102). Others cite lower, though still significant, percentages. *See, e.g.*, Ruffing, *Fighting the Substandard Lease*, AM. INDIAN J., June 1980, at 2, 7 (estimating that Indian Nations control 11% of all uranium, 8% of all coal, and 21% of all strippable coal in the United States). A major problem in estimating Indian resource holdings is that the Bureau of Indian Affairs, charged with the trust responsibility for managing Indian resources, has never completed an accurate resource inventory. *See* PCIRE, *supra* note 3, pt. 2, at 47.

[7] Williams, *Reagan's Initiatives Lead to More Questions Than Answers*, INDIAN TRUTH, Apr. 1983, at 4, 15; *see also* PCIRE, *supra* note 3, pt. 1, at 29 (stating that the majority of Indian mineral royalties are received by less than 10% of the nation's Indian tribes). Estimates of Indian mineral holdings vary according to source, see *supra* note 6, but those Indian Nations that do possess mineral and energy reserves control large quantities of these resources, disproportionate to their numbers as members of the general population of the United States. This fact alone, along with the history of non-Indian exploitation and appropriation of tribal resources, gives those few tribes sitting atop these stockpiles ambivalent feelings about their mineral wealth. *See generally* Owens, *Can Tribes Control Energy Development?*, AM. INDIAN J., Jan. 1979, at 3 (tribal efforts to control the development and exploitation of their resources are hindered by the federal government's administrative structure, their own inexperience, and the advantageous bargaining position of more knowledgeable corporations).

tously sitting atop strategic mineral and energy stockpiles.[8] The majority of tribes either lack a substantial developable resource base or have chosen not to develop. These tribes believe that policies based on such an assumption encourage relegation of their sovereign territories to a permanent Third World status, blamable either on failed socialism or capitalism, according to one's ideological viewpoint.[9] In short, a federal Indian policy that focuses upon the exploitation of tribal natural resources, and not on the development of tribal economies, is doomed to resistance and failure. To satisfy the "moral obligations of the highest responsibility and trust"[10] incumbent upon the United States in its dealings with Indian Nations, federal Indian Country development policy must address itself to the structural barriers currently preventing tribal economic and social self-sufficiency.

At the close of its 1982 session, Congress enacted legislation designed to address the barriers to Indian Country development.[11] In the Indian Tribal Governmental Tax Status Act of 1982 (the Tribal Tax Status Act) Congress provided Indian tribes a favorable tax status, similar to that now enjoyed by state and local governments.[12] According to its sponsors, the Tribal Tax Status Act is intended to "strengthen tribal governments significantly by providing additional sources of financing and by eliminating the unfair burden of taxes Indian tribal governments must now pay."[13]

This Article examines the Tribal Tax Status Act and its asserted potential for significantly strengthening the ability of Indian tribes to address the complex obstacles to economic and social development. Part I of the Article provides a brief expla-

---

[8] *See, e.g.*, Means, *supra* note 6, at 7; *see also* COMPTROLLER GENERAL'S REPORT, *supra* note 6, at 1–41; TASK FORCE SEVEN, *supra* note 1, at 46–47.

[9] *See* Winslow, *Speaking With Forked Tongue*, INDIAN TRUTH, Apr. 1983, at 8.

[10] Seminole Nation v. United States, 316 U.S. 286, 297 (1942).

[11] Tribal Tax Status Act, Pub. L. No. 97-473, 96 Stat. 2607 (codified as amended at 26 U.S.C. § 7871 and in scattered sections of 26 U.S.C. (1982 & Supp. I 1983)).

[12] *Id.*

[13] 127 CONG. REC. S5666, S5667 (daily ed. June 2, 1981) (remarks of Sen. Wallop (R-Wyo.)).

> Efforts of Indian tribal governments to levy their own taxes, however, have been only partially successful because the Internal Revenue Code does not extend to Indian tribes the same treatment it accords other state and local governments. This difference in treatment undermines the tax initiatives of tribal governments and seriously interferes with their efforts to improve the conditions of life in Indian country.

Id.

nation of the structural underdevelopment that has historically plagued Indian Country. By identifying the barriers to tribal development, it then becomes possible to analyze the likely effectiveness of the Tribal Tax Status Act in overcoming those barriers. Furthermore, a clear vision of the obstacles to tribal economic and social development is needed to devise effective initiatives if those contained in the Tribal Tax Status Act are inadequate for the task or require revision to become adequate.

Part II of the Article briefly describes the tax status of state and local governments under the Internal Revenue Code and compares their status to the treatment historically accorded tribal governments. Part II includes the legislative history of the Tribal Tax Status Act, beginning with a bill introduced in Congress in 1975. Part III of the Article analyzes the provisions of the Tribal Tax Status Act as finally enacted by Congress in 1982 and suggests the new opportunities for economic development made available to tribal governments by the Act. In granting tribes a tax status similar to states, Congress, at least in theory, has provided Indian tribes the ability to implement previously unavailable fiscal initiatives.[14]

Part IV of the Article argues that although the Tribal Tax Status Act is indeed an important step for Indian Nations on their long road to economic and social self-sufficiency, this new statute represents only a limited victory. The Act grants tribes a tax status equal to states only in *selected* areas of the Internal Revenue Code and still denies tribes the ability to utilize many of the important fiscal development tools now used by state and local governments.[15] Without these tools, Indian governments will not be able to stimulate development and to provide a tax base capable of sustaining needed governmental services for their people. Until Congress fully discharges its trust obligation to provide tribal governments with the equal ability to serve those they govern, tribal tax status will continue to be an area of conflict in the many sided battle waged by Indian people to exercise their treaty guaranteed rights of sovereignty and self-

---

[14] For a discussion of these initiatives, see *infra* notes 162–95 and accompanying text.

[15] *See infra* notes 228–34 and accompanying text (describing the limited authority granted to Indian tribal governments to issue tax-exempt bonds under the Tribal Tax Status Act).

sufficiency.[16] Part V makes specific recommendations that will allow Indian people to exercise more effectively those rights.

I. THE BARRIERS TO RESERVATION DEVELOPMENT AND TRIBAL SELF-SUFFICIENCY

A. *Structural Barriers: Remoteness, Infrastructure, Capital Access, Land Ownership, and Labor Force*

A host of interrelated structural barriers to tribal development exists that the expenditure of countless billions of federal dollars has yet to alleviate.[17] Although differing in degree from tribe to tribe, several critical barriers to development are shared by virtually every Indian Nation in the United States. These obstacles include the following: the remoteness of many Indian Nations from large regional markets; inadequate public services and physical infrastructure base; limited access to the capital needed for public services and other public/private and private sector economic initiatives; the trust and fractionated title status

[16] For accounts of the more significant legal battles for greater self-determination waged by tribes in the past two decades, see Ickes, *supra* note 3; Erlich, *Sovereignty and the Tribal Economy*, AM. INDIAN J., Nov. 1980, at 21; Barsh, *Issues in Federal, State, and Tribal Taxation of Reservation Wealth: A Survey and Economic Critique*, 54 WASH. L. REV. 531 (1979); Chambers & Price, *Regulating Sovereignty: Secretarial Discretion and the Leasing of Indian Lands*, 26 STAN. L. REV. 1061 (1974); Comment, *The Indian Battle for Self-Determination*, 58 CALIF. L. REV. 445 (1970).

[17] The budgetary outlay for federally funded Indian programs in fiscal year 1984 alone equaled nearly $2.7 billion. The figure was higher in 1983, exceeding $2.9 billion, compared with $3.4 billion in 1982. *See* PCIRE, *supra* note 3, pt. 2, at 93; Williams, *supra* note 7, at 15. *See generally* A. SORKIN, AMERICAN INDIANS AND FEDERAL AID (1971) (detailing federal programs and initiatives benefiting Indians).

The major federal Indian economic development programs include the following: The Snyder Act of 1921, 25 U.S.C. § 13 (1982) (providing funding authority for industrial and irrigation assistance, and the development of water supplies, employment, and administration); The Adult Indian Vocational Training Act of 1956, 25 U.S.C. §§ 309–309a (1982) (providing funding authority to develop skilled labor); The Indian Self-Determination and Education Assistance Act of 1975, 25 U.S.C. §§ 450–450n (1982) (providing funding authority for contracts, and grants for planning, training, and operating tribal businesses and federally funded programs administered by the Sec'y of Interior); The Indian Reorganization Act of 1934, 25 U.S.C. § 470 (1982) (establishing a revolving economic development loan fund); The Indian Financing Act of 1974, 25 U.S.C. §§ 1451–1543 (1982) (continuing Indian Reorganization Act of 1934 loan fund, adding interest subsidies, loan guarantees, and insurance). *See generally* Ickes, *supra* note 3, at 500–501 (discussing reasons for failure of federal Indian assistance programs to develop reservation economic potential).

of much land within Indian Country; and an unskilled and inexperienced labor and managerial force.[18]

These basic impediments to development act and interact to hinder indigenous economic and social development in Indian communities.[19] Furthermore, these barriers contribute to the private sector's lack of awareness or negative attitude toward investment opportunities in Indian Country.[20]

1. Remoteness and Lack of Infrastructure

Surreptitiously quartered in the nether reaches of the American outback,[21] many Indian Nations find themselves victims of past government policies designed to relocate them on lands regarded either as useless or as isolated from the advancing westward path of white colonization.[22] Today, the physical remoteness and the lack of modern infrastructure facilities—inevitable results of these past relocation policies—pose formidable entry barriers to Indian Country development initiatives.[23]

Physical remoteness can present a particularly intense and intractable problem for a tribe seeking to create or to expand business opportunities within its borders. An Economic Development Administration study of Indian projects found "a direct correlation between successful economic development programs on Indian reservations and their closeness to markets and large cities. The further away . . . from a large market, the more difficult it was to accomplish economic development."[24]

---

[18] *See supra* note 3.

[19] *See supra* note 3.

[20] *See Oversight of Economic Development on Indian Reservations*, *supra* note 3, at 6–21 (statement of Joe DeLa Cruz, President and Chairman, Quinalt Indian Nation).

[21] The typical Indian Nation has been likened to a high plains Third World ghetto:

> [The reservation resembles] an open-air slum. It has a feeling of emptiness and isolation. There are miles and miles of dirt or gravel roads without any signs of human life. The scattered Indian communities are made up of scores of tarpaper shacks or log cabins with one tiny window and a stovepipe sticking out of a roof that is weighted down with pieces of metal and automobile tires. These dwellings, each of them home for six or seven persons, often have no electricity or running water—sometimes not even an outhouse.

A. SORKIN, *supra* note 17, at 1.

[22] *See* R. BERKHOFER, JR., THE WHITE MAN'S INDIAN 145–69 (1978); U.S. COMM'N ON CIVIL RIGHTS, INDIAN TRIBES: A CONTINUING QUEST FOR SURVIVAL 18–20 (1981).

[23] *See supra* note 3.

[24] *See Indian Economic Development Programs: Oversight Hearings Before the House Comm. on Interior and Insular Affairs*, 96th Cong., 1st Sess. pt. 1, at 45 (1979) (testimony of Harold Williams, Deputy Ass't Sec'y, Economic Dev. Admin. (EDA)) [hereinafter cited as *Hearings on Indian Economic Development Programs*]. *But cf.* TASK FORCE SEVEN, *supra* note 3, at app. II 149–67 (exploring the role of the EDA in assisting tribes to achieve economic self-sufficiency, and noting the inconsistencies between EDA's legislative mandate and its actual actions in Indian Country).

Numerous other surveys analyzing industrial location decisions in the general economy have found that proximity to an identifiable market was one of the most important factors influencing a firm's decision to locate in a particular area.[25] Distance from such markets represents an insurmountable barrier to Indian Country economic development, stifling local development initiatives and strongly discouraging non-Indian investment.[26] Even when a tribe is fortunate enough to possess energy and mineral resources, the higher transportation costs associated with greater distance may mean that development will not occur for many years, while more accessible reserves are developed first.[27]

The usual lack of an adequate public infrastructure in Indian Country exacerbates the problems caused by remoteness.[28] Transportation networks, sewer and water systems, health facilities, public housing, schools, sanitation, and every other type of facility or service traditionally provided by government are woefully substandard throughout Indian Country.[29]

Few businesses are willing to invest capital in a location devoid of even the most basic public services and facilities, and the lack of infrastructure support in Indian communities strangles most indigenous development initiatives. Surveys of business location decisions indicate that governmental support and service facilities are a prime consideration in a firm's choice to expand, to move, or to open its doors for business in an area.[30] Few Indian tribes can presently compete with other governmental units in providing infrastructure bases for businesses within their jurisdictional borders.[31] Hence, any successful federal Indian development policy must initially address and offer

---

[25] *See* Meadows & Mitrisin, *A National Development Bank: Survey and Discussion of the Literature on Capital Shortages and Employment Changes in Distressed Areas*, in NEW TOOLS FOR ECONOMIC DEVELOPMENT: THE ENTERPRISE ZONE, DEVELOPMENT BANK, AND RFC 84–143 (G. Steinleib & D. Listokin eds. 1981) (citing several studies) [hereinafter cited as NEW TOOLS FOR ECONOMIC DEVELOPMENT].

[26] *See* Ickes, *supra* note 3, at 509–10; S. LEVITAN & W. JOHNSTON, INDIAN GIVING PROGRAMS FOR NATIVE AMERICANS 31 (1975).

[27] *See* S. LEVITAN & W. JOHNSTON, *supra* note 26, at 31.

[28] *See* TASK FORCE SEVEN, *supra* note 3, at 105–10; Barsh & Henderson, *Tribal Administration of Natural Resource Development*, 52 N.D.L. REV. 307, 308–12 (1975).

[29] TASK FORCE SEVEN, *supra* note 3, at 105–10.

[30] SUBCOMM. ON FISCAL AND INTERGOVERNMENTAL POLICY OF THE JOINT ECONOMIC COMM., 95TH CONG., 2D SESS., CENTRAL CITY BUSINESS—PLANS AND PROBLEMS (Comm. Print 1979); SUBCOMM. ON THE CITY OF THE HOUSE COMM. ON BANKING, FIN. AND URBAN AFFAIRS, 95TH CONG., 2D SESS., CITY NEED AND THE RESPONSIVENESS OF FEDERAL GRANTS PROGRAMS (Comm. Print 1978).

[31] *See* Barsh & Henderson, *supra* note 28, at 312.

solutions not only to the problem of remoteness, but also to the severe infrastructure deficiencies that exist throughout Indian Country.[32]

2. Capital Access and Tribal Land Ownership Patterns

Two critical and interrelated barriers to tribal social and economic development are the ownership patterns of a large portion of the land base within Indian Country and the problems of obtaining capital to finance development activities on this land.[33] Originally, legal title to all lands reserved to Indian Nations by treaty resided in the United States.[34] The tribe was said to have an equitable title, meaning that the federal government, acting as trustee, held title to the lands entirely for the benefit and use of the Indian tribe.[35]

Under the Indian General Allotment Act of 1887 (the Allotment Act), however, the United States government instituted a policy of dividing tribal land holdings among individual Indians.[36] The policy rested on a belief that the Indian's "savagism" was related to the communal ownership of tribal property.[37]

---

[32] *See Oversight of Economic Development on Indian Reservations*, *supra* note 3, at 5 (statement of Alan R. Parker, Director and Chairman of Strategic Planning Comm., Am. Indian Nat'l Bank):

> [W]ithout Federal assistance to provide basic infrastructure to stimulate the development of the private sector economy, to enable Indian tribes themselves to operate governments, that can, in fact, create an environment that is conducive to the development of a private sector economy; without that kind of Federal assistance, it is clear that the Indian tribes themselves will continue to be severely crippled and you will not see progress toward economic self-sufficiency.

Id.

[33] *See supra* note 3.

[34] *See* TASK FORCE ONE, AM. INDIAN POLICY REVIEW COMM'N, 94TH CONG., 2D SESS., REPORT ON TRUST RESPONSIBILITIES AND THE FEDERAL INDIAN RELATIONSHIP 47–68 (Comm. Print 1976).

[35] *See id.*; U.S. COMM'N ON CIVIL RIGHTS, *supra* note 22, at 25; Chambers, *Judicial Enforcement of the Federal Trust Responsibility to Indians*, 27 STAN. L. REV. 1213 (1975).

[36] The Indian General Allotment Act of 1887 (Dawes Act), 25 U.S.C. §§ 331–334, 336, 339, 341–342, 348–349, 381 (1982) [hereinafter cited as Allotment Act].

[37] *See* Getches, *Water Rights on Indian Allotments*, 26 S.D.L. REV. 405, 412–18 (1981); *see also* Ickes, *supra* note 3, at 498. Ickes notes that:

> The General Allotment Act was . . . an attempt to teach the Indians the Anglo-Saxon concept of private ownership and to further integrate them into white society. The Act largely failed to accomplish its objectives, partly due to a failure on the part of the federal government and Congress to recognize the cultural attitudes of the Indians toward private land ownership and toward communal development instead of individual land development . . . .

Id.

The notion that Indian land ownership patterns encouraged "savagism" presents a

Under the Allotment Act, each Indian was to be allotted an individual tract of tribal land—80 acres if the land was suited for agricultural purposes and 160 acres if it was suited for grazing purposes.[38] The trust status of the land continued for twenty-five years after the allotment, during which time it remained tax-exempt and inalienable.[39] After that period, or if it was found by the Secretary of Interior that the allottee was "competent and capable of managing his or her affairs," a patent in fee simple was issued to the allottee, free of "all restrictions as to sale, incumbrance, or taxation."[40]

Under the allotment policy, large amounts of treaty guaranteed land passed out of tribal ownership.[41] Conveyances, bequests, and sales for the benefit of Indian and non-Indian heirs and a provision in the Allotment Act permitting the Secretary of Interior to purchase "surplus" Indian land for non-Indian homesteaders[42] drastically reduced the tribally held land base.[43] In fact, by the time Congress finally repudiated the allotment policy in 1934,[44] tribes had lost ninety million acres, or two-thirds of all lands held prior to passage of the Allotment Act in 1887.[45]

The effects of the Allotment era are manifested today in the checkerboard pattern of land tenure in most of Indian Country.[46]

---

curious historical paradox. The right to own property in conglomerate form for the benefit of individual interests was the very basis of the American corporate system and gained prominence throughout the nineteenth century, *see* A. PAUL, CONSERVATIVE CRISIS AND THE RULE OF LAW 55–58 (1976), while a similar form of ownership when practiced by Indians was considered an indicium of savagery. *See* U.S. COMM'N ON CIVIL RIGHTS, *supra* note 22, at 34.

[38] 25 U.S.C. § 331 (1982).

[39] *Id.* § 348.

[40] *Id.* § 349.

[41] Getches, *supra* note 37, at 413–14.

[42] 43 U.S.C. § 1195 (1982).

[43] Getches, *supra* note 37, at 413–15.

[44] The Indian Reorganization Act of 1934, ch. 576, 48 Stat. 984–988 (codified at 25 U.S.C. §§ 461–479 (1982 & Supp. I 1983)).

[45] Getches, *supra* note 37, at 415.

[46] By statute (25 U.S.C.A. §§ 331–416 (1984)), tracts are subject to different terms and conditions of resource leasing, taxation, and development depending on their status as tribal land, allotted tracts, or fee simple tracts. This checkerboard ownership pattern leads to high costs when attempting to assemble a parcel of land. Further complications arise due to a number of jurisdictional concessions made to the states by the allotment acts. *E.g.*, 25 U.S.C. §§ 349, 357, 348a (1982). As a result, either tribal or state laws, in addition to federal regulation, may control any one tract. Integrated development requires that the tribe buy back diverse individual interests granted away in the past. The resulting economic effect is that development in Indian Country suffers a competitive disadvantage compared to development outside the borders of Indian Nations. Barsh & Henderson, *supra* note 28, at 320.

Ownership of much allotted land has become fractionated due to inheritance, with many tracts having six or more owners of undivided interests.[47] Non-Indians also control large tracts of land within Indian Nations, either through leases, prior transfers of "surplus" lands, or purchase or inheritance from Indian allottees.[48] Thus a majority of land in any Indian community is likely to be owned either in trust by the United States for the benefit of a tribe or an Indian allottee,[49] or by multiple Indian or non-Indian owners through the effects of the Allotment Act.[50]

Indians have historically encountered immense difficulty in acquiring access to development capital and conventional sources of commercial credit to finance investment initiatives on these trust lands and tracts affected by allotment.[51] The trust status of tribally held land prevents its use as collateral to finance a loan or mortgage for a development project on the tract.[52] Indian owned institutions, such as the American Indian National Bank, have been successful in creating alternatives to conventional types of financing, including such innovations as providing security for loans through assignment of leasehold interests and tribal grazing or timber fees.[53] The non-Indian

---

[47] H. HOUGH, DEVELOPMENT OF INDIAN RESOURCES 48 (1967). The National Congress of American Indians found that 48% of the allotments in a sample of 12 reservations had 6 or more owners of undivided interests, 29% had more than 10 owners of undivided interests, and 14% had more than 20. *See* Barsh & Henderson, *supra* note 28, at 320–21.

[48] *See* Getches, *supra* note 37, at 413–15. Getches notes that non-Indians presently cultivate approximately 63% of all Indian agricultural lands. *Id.* at 415. *See also* FINAL REPORT, *supra* note 3, at 318 (noting control by non-Indians of large amounts of land in Indian Country).

[49] Congress has extended the trust status on remaining Indian owned allotments through numerous legislative acts. *See, e.g.*, Indian Reorganization Act of June 18, 1934, ch. 576, § 2, 48 Stat. 984 (codified at 25 U.S.C. § 462 (1982)); Appropriations Act of June 21, 1906, ch. 3504, 34 Stat. 325, 326 (codified at 25 U.S.C. § 391 (1982)). *See generally* F. COHEN, HANDBOOK OF FEDERAL INDIAN LAW 619–21 (3d ed. 1982) (discussion of statutes extending the trust status on remaining Indian owned allotments).

[50] *See supra* notes 47–48 and accompanying text.

[51] *See Oversight of Economic Development on Indian Reservations*, *supra* note 3, at 5, 16–17 (statement of Alan R. Parker); TASK FORCE SEVEN, *supra* note 3, at 33–34.

The tribe's sovereign immunity from suit can be perceived as a factor contributing to poor capital access of Indians and their governments. However, sovereign immunity can be impliedly waived to the extent of liability insurance coverage, thereby improving relative access to capital. Note, *In Defense of Tribal Sovereign Immunity*, 95 HARV. L. REV. 1058, 1073–74 (1982).

[52] 25 U.S.C. §§ 202, 483a (1982) (immunity of trust land from mortgages).

[53] The American Indian National Bank was federally chartered by the Comptroller of the Currency in 1973 as an attempt to create "a national Indian financial structure, wholly owned by American Indian organizations and individuals, that would assist Indian communities to establish and develop a strong economic base." *Oversight of Economic Development on Indian Reservations*, *supra* note 3, at 91 (statement of R. Conley Ricker, Chief Executive Officer of the Am. Indian Nat'l Bank).

The Bank, a member of the Federal Reserve System, is subject to the normal regu-

business community, however, generally views the trust status of tribal lands as a nearly impenetrable barrier to business and capital investment.[54] Furthermore, because the trust status of tribal lands precludes non-Indian ownership of trust lands,[55] businesses that prefer fee ownership of a building site for tax or other financial considerations are discouraged from locating in Indian Country.[56]

Lands held by non-Indian owners, or by Indian allottees in fractionated interests, present similar financing barriers. Transaction costs in assembling sites or pooling rights for natural resource development and other types of projects may be insurmountable, thereby thwarting efforts to obtain commercial credit.[57]

The trust status and the checkerboard ownership pattern of tribal lands work to hinder capital access and financing opportunities for tribal development, thereby significantly inhibiting tribal self-sufficiency.[58] A federal development policy that ignores these barriers, or underestimates their significance, holds little potential for success.

3. Unskilled Labor Force and Inexperienced Managerial and Technical Personnel

Two other major factors contributing to tribal underdevelopment are the lack of a skilled labor force and the lack of an experienced nucleus of technicians and managers. Both prob-

---

latory requirements of the Comptroller of the Currency, and is insured by the Federal Deposit Insurance Corporation. The Bank was initially capitalized by the stock purchases of several Indian tribes and Alaska Native corporations. *Id.* at 91–93, 95.

[54] *Oversight of Economic Development on Indian Reservations*, *supra* note 3, at 5, 16–17 (statement of Alan Parker).

[55] *See* 25 U.S.C. §§ 202, 483a (1982).

[56] *See Hearings on Indian Economic Development Programs*, *supra* note 24, pt. 1, at 51 (testimony of Neil Daniel, Indian Desk, EDA). For example, the disappointing performance of the EDA's Indian industrial park location program has been attributed, in part, to the limitations on flexible ownership options posed by the trust status of tribal lands. *Id.* at 51. Daniel testified that of 48 industrial parks in Indian Country, built at a cost of over $25 million, 25 were without occupants. *Id.* at 52.

[57] *See* Barsh & Henderson, *supra* note 28, at 320. "Checkerboarding of reservations results in high transaction costs of orchestrating different classes of ownership all pertaining to, say, the same stream or coal seam." *Id.* Barsh and Henderson note that many tribes have begun comprehensive purchase-leaseback programs for purposes of more orderly ownership and development. *Id.* *See also* H. Hough, *supra* note 47, at 38–44 (detailing one such land consolidation program, administered by Cheyenne River Sioux).

[58] *See* S. Langone, The Heirship Land Problem and its Effect on the Indian, the Tribe, and Effective Utilization (1969), *reprinted in* Subcomm. on Economy in Gov't of the Joint Economic Comm., 91st Cong., 1st Sess., Toward Economic Development for Native American Communities, at 519 (Comm. Print 1969).

lems reflect the human dimension of the structural underdevelopment affecting virtually every Indian Nation in this country.[59]

Indian unemployment has historically run at rates from *six to eight times greater* than the general population.[60] For example, in the midst of the depression between 1981 and 1983 that engulfed every Indian Nation, Indian Country unemployment averaged between seventy-five and ninety-five percent, while the national rate hovered below fifteen percent.[61] As a result of this long-term structural unemployment, large numbers of Indians are without significant work experience or marketable job skills.[62] Even those Indians who have found employment in their

---

[59] *See generally supra* note 3. The 1980 census estimated there were approximately 408,000 Indians who were living below the poverty level, or approximately 26.6% of the total Indian population. This figure compares with 12.4% of the general population living below the poverty level. For Indians living in Indian Country, the poverty rate has been estimated to be twice as high as the national average for all Indians. PCIRE, *supra* note 3, pt. 2, at 81–85.

[60] *See* A. SORKIN, *supra* note 17, at 12; Schifter, *Indian Reservation Development: Reality or Myth?*, 9 CAL. W.L. REV. 38, 39–40 (1972).

[61] *See* Hertzberg, *Reaganomics on the Reservation*, THE NEW REPUBLIC, Nov. 22, 1984, at 15–18; Winslow, *supra* note 9, at 8; *Oversight of Economic Development on Indian Reservations*, *supra* note 3, at 4 (statement of Alan R. Parker) (citing Bureau of Indian Affairs estimates of 80% unemployment on some reservations in 1983); NTCA Statement, *supra* note 3, position paper B, at 6.

Reporting the true rate of unemployment in Indian Country is a difficult, if not impossible task. Given the long-term structural unemployment present in Indian Country, a large number of Indians would fall into the category of permanently discouraged and no longer actively seeking work. The BIA's official 1983 unemployment estimates were arrived at by tabulating the percentage of unemployed Indians between the ages of 16 and 65 years old who were actively seeking employment. Using this method, the BIA estimated that nationally Indian Country unemployment was over 21%. The unemployment rate among Indians reached 54% in one state, and exceeded 27% in 12 states. *See* PCIRE, *supra* note 3, pt. 2, at 82.

> Even before the current economic recession set in, Indian Country was historically experiencing high unemployment. For those Indian tribes without natural resources for development, unemployment ran as high as 75 percent. The largest employer on those locations were the federal programs and the TRIBAL GOVERNMENT. Since the onset of the recession the figures for unemployment in those same locations now range up to 95 percent. Most locations are now experiencing high increases in social problems of all kinds. In short, all of those same social afflictions as exists [sic] in all other non-Indian communities are on the rise in Indian Country and is [sic] aggravated by the federal neglect.
>
> The strategy of Reaganomics dictated that those most in need would be taken care of by a never-explained "safety net." Community problems would be taken care of by "volunteerism," and economic problems would be taken care of by attracting private industry. Early on, the Indian leaders pointed out that these prescriptions would not work in Indian Country, to no avail. No one would listen.

NTCA Statement, *supra* note 3, position paper B, at 6 (emphasis in original).

[62] *See* PCIRE, *supra* note 3, pt. 2, at 84–85.

national homelands or in nearby communities are generally employed in unskilled or semi-skilled jobs.[63]

Business and industry surveys consistently cite the availability of a skilled and experienced labor force as a significant determinant in any location or expansion decision.[64] The long-term structural unemployment affecting many Indian Nations therefore represents an obvious entry barrier to a non-Indian employer considering an Indian Country location.[65] For the individual Indian entrepreneur, or an indigenous tribal enterprise, the costs of training an Indian labor force may represent yet another almost insurmountable barrier to establishing a business.

The lack of job opportunities at entry level positions for tribal members inevitably translates into a lack of experienced managers and technicians among the general Indian population.[66] Furthermore, the Bureau of Indian Affairs (BIA) within the United States Department of the Interior, which historically controlled the majority of federally financed Indian development projects, for years systematically promoted a neocolonialist policy that assumed Indian incapacity to manage tribal development. The BIA routinely assigned non-Indians to establish and to run new enterprises or developments in Indian Country.[67]

---

[63] *Id.* at 84; S. LEVITAN & W. JOHNSTON, *supra* note 26, at 32. At the same time, there are numerous, though underpublicized, examples of highly successful businesses in Indian Country which demonstrate that Indians can be trained for high-skilled work such as electronics assembly and specialized manufacturing processes. *See* Ickes, *supra* note 3, at 503 (Yankton Sioux electronics factory, Greenwood, South Dakota); Choctaw Indian Tribe, Choctaw Industrial Park (Mar. 1982) *reprinted in Oversight of Economic Development on Indian Reservations*, *supra* note 3, at 59–85 (information booklet published by the Mississippi band of Choctaw Indians describing several successful specialized manufacturing enterprises owned and managed by the tribe).

[64] *See supra* notes 25 & 30; Ickes, *supra* note 3, at 508–09.

[65] *See supra* note 61. Other reasons cited for the reluctance of private industry to locate in Indian Country include the lack of business traditions among tribal members, an outgrowth of the "sub-culture of poverty," and long-term unemployment that plagues most Indian Nations. Ickes, *supra* note 3, at 508–09. Overcoming such obstacles, however, will involve addressing issues of cultural identity and preference and working towards solutions that both red and white races can support. *Id.*

[66] *See, e.g., Miscellaneous Tax Bills: Hearings Before the Subcomm. on Miscellaneous Revenue Measures of the House Comm. on Ways and Means*, 95th Cong., 1st Sess. 118 (1977) (statement of Ernest Clark, Chairman of the Fin. Comm., Colville Confederated Tribes) [hereinafter cited as *Hearings Before the Subcomm. on Miscellaneous Revenue Measures*].

For example, only 7.8% of the Indian male labor force and 7.3% of the Indian female labor force are employed in managerial and professional occupations, compared with 22.5% and 20.6% for males and females respectively in the national labor force. PCIRE, *supra* note 3, pt. 2, at 84.

[67] *See* Barsh & Henderson, *supra* note 28, at 309–19. The authors attribute the Bureau's negative attitude toward Indian control of development as emanating from

Because of this paternalistic policy, valuable training opportunities for potential Indian managers and technicians were lost.

Without experienced management and technical personnel among the Indian population, indigenous business efforts are not likely to thrive. Furthermore, Indians will continue to be relegated to the lowest positions in any non-Indian owned and supervised enterprise that might locate in Indian Country.[68]

The Indian manpower problem provides a striking example of the interrelated nature of the barriers to Indian Country self-sufficiency. Long-term structural unemployment has created an unskilled and inexperienced labor force, thus presenting an additional entry barrier to Indian and non-Indian business ventures. Obviously, federal Indian development policy must aggressively pursue solutions that can assist tribal governments in resolving the Indian manpower dilemma. Increased job opportunities, both in entry level and management and technical positions, are prerequisites for attaining tribal economic and social self-sufficiency.

B. *Cultural Attitudes*

Perhaps the greatest barrier to Indian Country development is the cultural impasse that has historically existed between

---

historical federal policies of "leadership-deficit-thinking" concerning tribal managerial abilities:

> "Leadership-deficit-thinking" describes a theory of policy that justifies federal administration of tribal resources and simultaneously seeks to excuse the United States from responsibility for continuing Indian poverty. It is characterized by two axiomatic assertions about Indians: (1) they are anti-developmental and ignorant, and (2) these failings are the result of their cultural differentness.

Id. at 311. *See also* PCIRE, *supra* note 3, pt. 2, at 47–54 (Presidential Comm'n Report emphasizing BIA management deficiencies); Ickes, *supra* note 3, at 506–508 (citing congressional studies and hearings on the BIA's own deficiencies in handling tribal business affairs).

> The BIA's supervisory staff, however, rarely has any business expertise itself. There is a notable absence of managerial and organizational capacity throughout the BIA. Decisions are made on a day-to-day basis with little or no long-range planning. Communication among organizational levels is poor and there are few, if any, performance standards or evaluations, even of key positions. A number of government-sponsored industrial projects have failed because the so-called government experts were not up to the job of giving the Indians sound economic advice. The result is often the loss of limited tribal capital, as well as the loss of government funds.

Id. at 506–507.

[68] *See, e.g.*, *Hearings on Indian Economic Development Programs*, *supra* note 24, pt. 2, at 9–11 (statement of John Navarro, Director, Tribal Employment Rights Program, Papago Tribe) (detailing discriminatory treatment of Indians by non-Indian employers and the difficulties in promoting development in Indian Country).

white and red America. This unbridged gulf has prevented the formulation of policies that both cultures can support and administer with commitment, and it accounts for the long history of failed Indian Country federal development initiatives.[69]

An unquestioning Eurocentric bias has historically hindered white America in developing an appreciation of the Indian's unique, tribally centered vision of the meaning and goals of development.[70] Instead, United States Indian policy, in both its rhetoric and substance, has generally reflected the white man's idealized image of his own role in the taming of the frontier—a paradigmatic collective experience that each generation of white Americans mythologically reconstitutes for itself.[71] For example, the seventeenth century New England Puritans structured their relations with neighboring Indian Nations upon Biblically inspired notions of America as an untamed chosen land.[72] The Puritans regarded the Indians as prime antagonists to their predestined mission in this divinely orchestrated plan. Referring to surrounding New England tribes, the Puritan leader and clergyman Cotton Mather declared, "The Wilderness thro' which we are passing to the Promised Land, is all over fill'd with Fiery flying Serpents. . . . [T]here are incredible droves of devils in our way."[73]

---

[69] The historical gulf between red and white cultures is an ingrained feature of the relationship between the original inhabitants of the Western Hemisphere and the European colonists. White culture, aided by force of arms and superior destructive technology, has historically viewed Indian culture as inherently inferior and therefore properly subject to assimilative and appropriative policies and laws, dictated by the European conquerors without benefit of Indian consent. On the history of the stance taken by Europeans, see generally R. BERKHOFER, JR., *supra* note 22.

[70] *See* V. DELORIA, JR. & C. LYTLE, AMERICAN INDIANS, AMERICAN JUSTICE 80 (1983) (the view of government held by most Indians was measurably different from that held by white men until white pressure and influence led to Indian change):

> The whites' inability to understand the problems and behavior of American Indians is not a new phenomenon. Indeed, this difficulty in appreciating and comprehending "foreign" cultures and traditions seems to be the peculiar burden that western European peoples carry into their encounters with peoples of different traditions. This blinded vision has affected all dealings with non-Westerners.

Id. See also Williams, *The Medieval and Renaissance Origins of the Status of the American Indian in Western Legal Thought*, 57 S. CAL. L. REV. 1 (1983) (discussing earliest Eurocentric conceptions of Indian societies as backward and primitive).

[71] *See generally* R. BERKHOFER, JR., *supra* note 22, at 27: "[M]any commentators on the history of White Indian imagery see Europeans and Americans as using counterimages of themselves to describe Indians . . . ." *Id.* On the significance of the frontier paradigm in American history, see F. TURNER, THE SIGNIFICANCE OF THE FRONTIER IN AMERICAN HISTORY (1893).

[72] *See generally* C. SEGAL & D. STINEBACK, PURITANS, INDIANS AND MANIFEST DESTINY (1977) (discussing Puritan views on surrounding Indian tribes).

[73] C. MATHER, THE WONDERS OF THE INVISIBLE WORLD 63 (Boston 1677).

Given such a Biblically oriented world view, it is not surprising that the authorities of New Plymouth declared that "there was no dealing with the Indians . . . above board . . . ."[74] Wars, chicanery, threats and "miraculous" plagues, which forced the Indian to leave his ancestral lands, were regarded as providential opportunities for the Puritans to claim legal title to Indian property by virtue of the natural law doctrine of *vacuum domicilium*.[75] Thus, one of the earliest policies of whites toward the tribal peoples of this continent viewed Indian Nations as the ultimate barrier to development of the frontier, to be removed by any and all means.[76]

This same frontier paradigm shaped the removal and allotment policies of the nineteenth century, as non-Indian policymakers attempted to remold the conquered, but still savage, Indian into the revered image of that tamer of the American wilderness, the yeoman farmer.[77] The failure of these policies has been attrib-

---

[74] I. MATHER, A RELATION OF THE TROUBLES 16 (Boston 1677).

[75] C. SEGAL & D. STINEBACK, *supra* note 72, at 31.

[76] *Id.* at 25–49.

John Cotton, one of the great preachers of the Puritan era, fashioned the following Biblical argument as justification for dispossessing American Indians:

> *The placing of a people in this or that countrey [sic] is from the Appointment of the Lord.*
>
> This is evident in the Text; and the Apostle speaks of it as grounded in nature, *Act. 17. 26. God hath determined the times before appointed, and the bounds of our habitation. Deut. 2 chap. 5. 9.* God would not have the *Israelites* meddle with the *Edomites*, or the *Moabites*, because he had given them their Land for a possession. God assigned out such a Land for such a Posterity, and for such a time.
>
> . . .
>
> Now God makes room for a People three ways:
>
> 1. When he casts out the Enemies of a people before them, by lawful War with the Inhabitants, which God calls them unto, as in *Psal.* 44. 2. *Thou didst drive out the Heathen before them.* But this course of Warring against others, and driving them out without provocation, depends upon special Commission from God, or else it is not imitable.
>
> 2. When he gives a *forreign* [sic] People favour in the eyes of any *native* People to come and sit down with them; either by way of purchase, as *Abraham* did obtain the field of *Machpelah*: or else when they give it in courtesie [sic], as *Pharoah* did the Land of *Goshen* unto the sons of *Jacob*.
>
> 3. When he makes a country, though not altogether void of Inhabitants, yet void in that place where they reside. Where there is a vacant place, there is liberty for the Son of *Adam* or *Noah* to come and inhabit, though they neither buy it, nor ask their leaves. *Abraham* and *Isaac*, when they Sojourned amongst the *Philistines*, they did not *buy* that Land to feed their cattel [sic], because they said There is room enough.
>
> And so did *Jacob* pitch his Tent by *Sechem*, Gen. 34, 21.

J.COTTON, GOD'S PROMISE TO HIS PLANTATIONS 3–4 (2d ed. Boston 1686) (1st ed. London 1634) (footnote omitted) (emphasis in original).

[77] *See supra* notes 36–45 and accompanying text.

uted in large part to the Indian's own lack of reverence or interest in the white man's idealized image of himself. The Indian, as always, simply wanted to remain and to rule himself, as an Indian.[78]

During the 1930's, Indian tribes were "reorganized" according to the institutionalized democratic structures favored by the New Deal's social engineers.[79] For these policymakers, Indian Nations offered a unique laboratory for testing the theories of their new and evolving social sciences—theories holding the promise of rationally reordering American economic and political life and creating brave new worlds.[80] For the Indian, such theories, grounded in alien traditions and political institutions, only succeeded in creating strange new worlds. The infusion of non-Indian political and social values and institutions into tribal life created tensions and destructive factions, which continue to impede Indian progress today.[81]

In the 1950's, federal policymakers and members of Congress sought to "terminate" the nation's Indian tribes by ending the trust relationship between the federal government and America's indigenous nations.[82] Despite the New Deal's efforts at

---

[78] *See generally*, W. WASHBURN, THE ASSAULT ON INDIAN TRIBALISM: THE GENERAL ALLOTMENT LAW (DAWES ACT) OF 1887 (1975). Washburn notes that although "the overwhelming majority of Indians opposed the breakup of the tribal system, the Indian voice was either not heard, not heeded, or falsely reported." *Id.* at 8.

[79] *See* G. TAYLOR, THE NEW DEAL AND AMERICAN INDIAN TRIBALISM: THE ADMINISTRATION OF THE INDIAN REORGANIZATION ACT 1934–1945 (1980). The principal tool used for this reorganization effort was the Indian Reorganization Act of 1934, 25 U.S.C. §§ 461–479 (1982 & Supp. I 1983), which focused on tribal culture and institutions as a means for the assimilation of Indian tribes into American society. *See* K. PHILIP, JOHN COLLIER'S CRUSADE FOR INDIAN REFORM, 1920–1954 (1977); Comment, *Tribal Self-Government and the Indian Reorganization Act of 1934*, 70 MICH. L. REV. 955 (1972).

[80] R. BERKHOFER, JR., *supra* note 22, at 176–86. *See generally* K. PHILIP, *supra* note 79.

[81] *See* K. PHILIP, *supra* note 79, at 162–67; *see also* D. GETCHES, D. ROSENFELT & C. WILKINSON, FEDERAL INDIAN LAW: CASES AND MATERIALS 84–85 (1979) (describing intense factionalism between traditionalists and assimilated elements of the Hopi tribe, brought about in large part by the introduction of a non-Indian constitutional form of government under the Indian Reorganization Act).

[82] The key component of the termination policy was House Concurrent Resolution 108, introduced and approved on June 9, 1953:

> *Resolved by the House of Representatives (the Senate concurring)*, That it is declared to be the sense of Congress that, at the earliest possible time, all of the Indian tribes and the individual members thereof located within the States of California, Florida, New York, and Texas, and all of the following named Indian tribes and individual members thereof, should be freed from Federal supervision and control and from all disabilities and limitations specially applicable to Indians: [list of individual tribes omitted] . . . . It is further declared to be the sense of Congress that the Secretary of the Interior should examine all existing legislation dealing with such Indians, and treaties between

Indian reorganization, many whites at the dawn of the McCarthy era felt that tribes still exhibited "communistic traits"—traits regarded as a corrosive and intolerable influence in a free and democratic society.[83] Even before Indians could effectively organize themselves politically to protest what they regarded as an inherently genocidal federal policy, white guilt over the rapid cultural and social disintegration of tribal communities led to the policy's quick demise in the early 1960's.[84]

Since the 1960's, white policymakers have evidenced a seeming willingness at least to listen to Indian desires in the formulation of Indian development policy.[85] Indian leaders, however, still assert that those in Washington, D.C., consistently fail to recognize adequately and to address effectively any tribally centered vision of the meaning and goals of Indian economic and social self-sufficiency.[86]

> the Government of the United States and each such tribe, and report to Congress at the earliest practicable date, but not later than January 1, 1954, his recommendations for such legislation as, in his judgment, may be necessary to accomplish the purposes of this resolution.

H.R. Con. Res. 108, 67 Stat. B132 (1953).

[83] *See Investigate Indian Affairs: Hearings Before a Subcomm. of the House Comm. on Indian Affairs Pursuant to H.R. Res. 166*, 78th Cong., 2d Sess. 1054 (1944). On the history of the termination policy, see generally Wilkinson & Boggs, *The Evolution of the Termination Policy*, 5 AM. INDIAN L. REV. 139 (1977); G. Orfield, A Study of the Termination Policy (1966), *reprinted in* STAFF OF SUBCOMM. ON INDIAN EDUCATION OF THE SENATE COMM. ON LABOR AND PUBLIC WELFARE, 91ST CONG., 1ST SESS., THE EDUCATION OF AMERICAN INDIANS 674–90 (Comm. Print 1970).

[84] *See, e.g.*, Preloznik & Felsenthal, *The Menominee Struggle to Maintain Their Tribal Assets and Protect their Treaty Rights Following Termination*, 51 N.D.L. REV. 53 (1974) (detailing effects of termination on Menominee tribe of Wisconsin); *see also* Menominee Tribe of Indians v. United States, 391 U.S. 404 (1968) (holding that the tribe's hunting and fishing rights under treaty survived the Termination Act of 1954).

[85] *See, e.g.*, MESSAGE OF THE PRESIDENT OF THE UNITED STATES TRANSMITTING RECOMMENDATIONS FOR INDIAN POLICY, H.R. DOC. No. 363, 91st Cong., 2d Sess. 1 (1970):

> It is long past time that the Indian policies of the Federal government began to recognize and build upon the capacities and insights of the Indian people. Both as a matter of justice and as a matter of enlightened social policy, we must begin to act on the basis of what the Indians themselves have long been telling us. The time has come to break decisively with the past and to create the conditions for a new era in which the Indian future is determined by Indian acts and Indian decisions.

Id. (statement by Pres. Nixon) [hereinafter cited as NIXON MESSAGE]. *See also* Israel, *The Re-emergence of Tribal Nationalism and Its Impact on Reservation Resource Development*, 47 U. COLO. L. REV. 617 (1976) (the 1960's civil rights movement sensitized the Indian community to the vulnerability of tribal existence, resulting in the re-emergence of tribal nationalism supported to a certain extent, indirectly, by the Federal government).

[86] *See* NTCA Statement, *supra* note 3, position paper B, at 3–13, position paper C, at 1–9; *see also* Reagan Indian Policy Statement, *supra* note 6, at 96. Since the announcement by Pres. Nixon of his recommendations for Indian policy in 1970, *see* NIXON MESSAGE, *supra* note 85, "there has been more rhetoric than action," on the

This historic inability to comprehend the Indian's vision of the meaning and goals of development accounts for the history of failure of federal Indian policy. Indian people have demonstrated that they will not support federal initiatives incorporating meanings and goals alien to their world view.[87] In refusing to incorporate Indian preferences, non-Indian policymakers have systematically ignored the most basic requirements of a successful economic development policy—strong, idea oriented local leadership and the commitment of the people directly affected by the policy.[88]

Creating and sustaining the conditions whereby America's indigenous tribal cultures can thrive as a people, although surrounded by a dominant society animated by an individualistic world view radically divergent from the Indian world view, remains as the fundamental challenge of federal Indian policy. History teaches, however, that the cultural impasse existing between red and white America has prevented the formulation of policies capable of meeting this challenge. Perhaps the most

---

part of the federal government in fostering and encouraging tribal self-government. Reagan Indian Policy Statement, *supra* note 6, at 96. "Instead of fostering and encouraging self-government, Federal policies have by and large inhibited the political and economic development of the tribes. Excessive regulation and self-perpetuating bureaucracy have stifled local decisionmaking, thwarted Indian control of Indian resources, and promoted dependency rather than self-sufficiency." *Id. But see* Means, *supra* note 6, at 7:

> [T]he Reagan Administration's proposed solution to the social problems of the reservations . . . promises instead the further undermining of an already battered Native American culture. Administration policy calls for the reduction of funding for the reservations and the invitation of private interests to assume the responsibility for economic support and development of Indian communities.
>
> . . .
>
> . . . [B]y catering to corporate interests [in developing natural resources], . . . the Administration's proposal does not offer Indian people the opportunity to rebuild independent communities and to develop a viable economic strategy consistent with the nonexploitive attitude of traditional American Indian culture.

Id.

[87] *See infra* notes 88–89.

[88]

> Tribal experience has shown us that the cornerstone of true economic development on Indian reservations is through the establishment of stable tribal governments with the management and structural capacities to carry out their responsibilities effectively. The importance of strong and effective tribal governments associated to [sic] long-term economic advancement cannot be overemphasized. In fact, we believe that without federal support to help tribal governments build strong foundations, all other policies attempting to foster economic development will be destined to fail.

Oversight of Economic Development on Indian Reservations, *supra* note 3, at 35 (material submitted by Harry Early, Chairman, Council of Energy Resource Tribes).

important barrier to be removed from the road to tribal self-sufficiency, therefore, is the failure to consider seriously the views of Indian people on the meaning and goals of federal Indian development policy.[89]

C. *Summary*

In sum, the barriers to tribal development remain immense. Remoteness of many Indian communities remains an intractable problem, which perhaps only the most perfectly conceived and executed development policy can ever hope to resolve. Inadequate infrastructure, manpower, and management skills and the ownership patterns of tribal lands inhibit the growth of both Indian and non-Indian development opportunities. The historical cultural impasse that exists between Indians and non-Indians prevents the formulation of effective development strategies that both Indians and non-Indians can support and administer. These major barriers to tribal self-sufficiency stand as ominous challenges to any new federal Indian policy initiative. The remainder of this Article considers how the recently enacted Tribal Tax Status Act measures up to that challenge.

II. THE LONG ROAD OF THE TRIBAL TAX STATUS ACT: 1976–1982

Like so many other issues in federal Indian law, attempts to define the tax status of Indian tribal governments under the Internal Revenue Code have been marked by inconsistency and

---

89

> [W]hen we talk about the development of reservation resources, that has to be done consistent with the wishes and desires of that Indian community or the Indian people that reside on that reservation.
>
> There are certain areas of reservations that are sacred to a tribe, and that tribe is not going to go in and tear that piece of reservation up for coal development or any other kind of development . . . .
>
> . . . [T]he point is that we have to . . . give substantial consideration to those cultural values as we develop an approach for resource development on reservations.

Hearings on Indian Economic Development Programs, *supra* note 24, at 10–11 (statement by Martin Seneca, Acting Deputy Comm'r of Indian Affairs).

contradiction.[90] During the closing days of its 1982 session, Congress sought to resolve that status by passage of the Indian Tribal Governmental Tax Status Act.[91] This legislation, the product of nearly a decade of lobbying activity by Indian governments and their supporters, provides a favorable tax status under the Internal Revenue Code for Indian tribes, similar to the favorable tax status enjoyed by state and local governments. The Tribal Tax Status Act is intended to facilitate tribal efforts to provide the types of governmental services and facilities that can create the development environment necessary for true economic and social self-sufficiency.[92]

A. *The Tax Status of States and Their Subdivisions*

State and local governments have historically received favorable treatment under federal tax laws.[93] Even prior to the establishment of a national income tax in the early years of this century, the Supreme Court in *Pollock v. Farmers' Loan & Trust Co.*[94] held that the doctrine of intergovernmental immunity barred the federal government from imposing a tax on the interest paid on state and municipal debt obligations.[95] Although most commentators today would question the continuing vitality of *Pollock*'s basic principle that no level of government within the federal system may tax the income received from activities of another level,[96] Congress has provided numerous tax benefits to states and their subdivisions within the Internal Revenue Code. Aside from certain narrowly defined exceptions,[97] reve-

[90] *See generally* Pelcyger, *Justices and Indians: Back to Basics*, 62 OR. L. REV. 29, 30 (1983). "In this very vital area—governmental authority over reservation activities of non-Indians—the [Supreme] Court has been anything but consistent and predictable." *Id. See also* Clinton, *State Power Over Indian Reservations: A Critical Comment on Burger Court Doctrine*, 26 S.D.L. REV. 434 (1981) (criticizing recent Supreme Court rulings involving the extent of state power in Indian Country).

[91] Tribal Tax Status Act, Pub. L. No. 97-473, 96 Stat. 2607 (codified as amended at 26 U.S.C. § 7871 and in scattered sections of 26 U.S.C. (1982 & Supp. I 1983)).

[92] *See* 127 CONG. REC. S5666–67 (daily ed. June 2, 1981) (remarks of Sen. Wallop).

[93] *See, e.g.*, Act of Aug. 16, 1954, ch. 736, 68A Stat. 29, 35 (codified as amended at 26 U.S.C. §§ 103, 115 (1982 & Supp. I 1983).

[94] 157 U.S. 429 (1895), *aff'd on reh'g*, 158 U.S. 601 (1895).

[95] *Pollock*, 157 U.S. at 585–86.

[96] *See, e.g.*, F. MICHELMAN & T. SANDALOW, GOVERNMENT IN URBAN AREAS 770–71 (1970).

[97] For instance, interest on certain classes of industrial development bonds and arbitrage bonds issued by states is not tax-exempt. I.R.C. §§ 103(b)(1), (c)(1) (1982). *See generally* Perkins and Goss, *Taxation of State and Local Obligations*, 11 URB. LAW. 488, 498 (1979) (describing the principal features of rules and regulations limiting state and municipal investment).

nues of state and local governments and the interest that those governments pay on their debt obligations are not subject to federal taxation.[98]

Thus, under the Internal Revenue Code, income accruing to state and local governments is not taxable.[99] These governments are also exempt from most federal excise taxes, such as the retailer's excise tax[100] and the federal motor vehicle tax.[101] In addition to the exclusion from gross income of interest paid on state and municipal debt obligations,[102] numerous other transactions between private parties and states and their subdivisions receive favorable tax treatment under the Code. Gifts and bequests to these governments are treated as tax-deductible charitable contributions,[103] and taxes paid by businesses and individuals to state and local governments are also allowed as deductions for federal tax purposes.[104]

Principles of federalism, together with practical financial considerations, dictate that "the capability of state and local governments to raise and use revenue should be facilitated and enhanced whenever possible in order that they may better serve the needs of their people."[105] Despite the fact that the federal government has expressly recognized tribal governments as sovereign entities for more than 150 years,[106] with responsibilities to constituents equal to that of state and local governments,[107] Indian tribes have historically been denied an equal federal tax status. Rather, Indian governments have occupied an anomalous niche within the structure of federal tax laws, enjoying some of the privileges afforded states, while at the same time being

---

[98] I.R.C. §§ 103, 115 (1982 & Supp. I 1983). *See Tax Reform: Hearings Before the House Comm. on Ways and Means*, 91st Cong., 1st Sess. 2196–97 (1969) (statement of Northcutt Ely, General Counsel, Am. Pub. Power Ass'n).

[99] I.R.C. § 115(1) (1982) (exemption for income derived from a public utility or other essential governmental function and accruing to a state or its subdivisions).

[100] *Id.* § 4222.

[101] *Id.* § 4483.

[102] *Pollock*, 157 U.S. at 450–51.

[103] I.R.C. § 170(c)(1) (1982).

[104] *Id.* § 164(a).

[105] *Hearings Before the Subcomm. on Miscellaneous Revenue Measures*, *supra* note 66, at 109 (statement of William E. Sudow, counsel for various Indian tribes).

[106] *See* Worcester v. Georgia, 31 U.S. (6 Pet.) 515, 559–60 (1832).

[107] *See* Merrion v. Jicarilla Apache Tribe, 455 U.S. 130, 137–41 (1982). In *Merrion*, non-Indians who produced oil and gas within the tribe's jurisdictional territory pursuant to leases granted under the auspices of the Secretary of the Interior sued to prohibit enforcement of the tribe's oil and gas severance tax. The Court, in an opinion by Justice Marshall, held that the tribe had the inherent power to impose a severance tax on mining activity as part of its power to govern and to pay for the costs of self-government. *Id.* at 159.

subjected to many of the burdens borne by ordinary individual taxpayers.[108]

This peculiar tax status of tribal governments was created principally by the Internal Revenue Service (IRS) in a series of revenue rulings during the 1960's and 1970's respecting transactions involving Indian governments.[109] The IRS's most comprehensive statement on Indian tax status is contained in Revenue Ruling 67-284, which sets forth "the general principles applicable to the federal income tax treatment of income paid to or on behalf of enrolled members of Indian tribes."[110] In a section devoted specifically to the "tax status of tribes,"[111] the ruling declares, "Income tax statutes do not tax Indian tribes. The tribe is not a taxable entity."[112]

Under Revenue Ruling 67-284, the income of an Indian tribal government is exempt from federal taxation, despite the total absence of any express congressional authority for such an exemption.[113] On the other hand, the IRS has pursued the dia-

---

[108] *See* Barsh, *supra* note 16, at 544–55.

[109] *See, e.g.*, Rev. Rul. 56-342, 1956-2 C.B. 20 ("[i]ncome . . . derived directly from [taxpayer's own] allotted and restricted Indian lands" exempted from taxation); Rev. Rul. 67-284, 1967-2 C.B. 55 (limited Indian income exemption). "Tax law is said to be a matter of statutory interpretation, of locating and construing code provisions and regulations. The distinguishing feature of Indian tax law, however, is the absence of statutes. Indians are not mentioned in the Internal Revenue Code." Barsh, *supra* note 16, at 544.

[110] Rev. Rul. 67-284, 1967-2 C.B. 55, 55.

> The Service will . . . recognize the exempt status of income received by an enrolled member of an Indian tribe where each of the following tests are [sic] met: (1) The land in question is held in trust by the United States Government; (2) such land is restricted and allotted and is held for an individual noncompetent Indian, and not for a tribe; (3) the income is "derived directly" from the land; (4) the statute, treaty or other authority involved evinces congressional intent that the allotment be used as a means of protecting the Indian until such time as he becomes competent; and (5) the authority in question contains language indicating clear congressional intent that the land, until conveyed in fee simple to the allottee, is not to be subject to taxation.

Id. at 56–57.

[111] *Id.* at 58.

[112] *Id.* The ruling goes on to state that "[a]bsent a provision in a treaty or statute to the contrary, income directly derived by a member of an Indian tribe from unallotted Indian tribal lands is subject to Federal income tax." *Id.* (citing Rev. Rul. 58-320, 1958-1 C.B. 24).

[113] The IRS did not attempt to harmonize its ruling with the line of Indian tax cases holding that exceptions to the Internal Revenue Code must be explicit. *See* Superintendent of Five Civilized Tribes v. Commissioner, 295 U.S. 418, 420 (1935); Big Eagle v. United States, 300 F.2d 765, 769 (Ct. Cl. 1962). Congress has explicitly exempted the income derived from a public utility or other essential government function by a state or its subdivisions, I.R.C. § 115(1) (1982); the income of any United States possession or its subdivisions, *id.* § 115(2) (1982); and income of certain federal corporations, *id.* § 501(c)(1) (1982). Given these explicit exemptions, the failure to grant Indian tribes an express tax exemption in the Code casts doubts on the validity of this portion of Revenue

metrically opposite policy of strict statutory construction in its rulings on all other aspects of Indian tribal governmental tax status. For example, section 103 of the Internal Revenue Code specifically provides a tax exemption for interest paid to an individual on state or local government debt obligations.[114] The IRS has ruled that tribal government debt obligations are not obligations of a "State, a Territory or a possession of the United States, or any political subdivision of any of the foregoing" as defined by section 103 of the Internal Revenue Code, and therefore interest paid to an individual on bonds issued by a tribe is not tax-exempt.[115] Critics have argued that the reasoning supporting this ruling is blatantly inconsistent with the IRS's prior ruling (enunciated in Revenue Ruling 67-284), holding that tribes themselves are tax-exempt despite a similar absence of express statutory language granting exemption.[116]

In a 1974 revenue ruling involving the Zuni Indian Pueblo, the IRS continued its policy of limiting the special treatment accorded tribes solely to their own income tax liability and refused to allow a decedent's bequest to the Zuni Pueblo as a governmental charitable deduction under section 2055 of the Code.[117] Similarly, the IRS concluded that the sale of an auto-

---

Ruling 67-284. Fortunately for many tribes, Congress included express statutory exemptions from taxation in many treaties, and courts have shown a willingness to find implied tax exemptions in a statute or treaty if evidence points to a federal policy encouraging tribal Indians to become self-sufficient in the development and commercial use of the resources of the reservation. *See, e.g.*, Squire v. Capoeman, 351 U.S. 1, 10 (1956); *see also* Comment, *Indian Taxation: Underlying Policies and Present Problems*, 59 CALIF. L. REV. 1261, 1268–71 (1971) (real property tax exemption).

[114] I.R.C. § 103(a)(1) (1982).

[115] Rev. Rul. 68-231, 1968-1 C.B. 48, 49:

> [T]he powers of the Tribal Community are delegated to it by the United States Government rather than the State of Washington in which it is located. The Tribe is not a division of the State, and since it exercises its governing powers by virtue of Federal, rather than State authority, the bonds in question are not issued on behalf of the State within the meaning of [§ 103(a)].

Id. at 49–50.

[116] *See, e.g.*, Barsh, *supra* note 16, at 553:

> It seems strange that the tribe's bonds are taxable for lack of express statutory language when the tribe itself is exempt by analogy at best. It is stranger still because it should be in the United States' own financial interest to exempt tribal bond interest from income taxation, since tribes are more dependent on federal financial support when they cannot compete with tax-exempt state and municipal bonds in private capital markets.

Id.

[117] *See* Rev. Rul. 74-179, 1974-1 C.B. 279:

> The powers of self-government which Indian tribes or pueblos exercise stem from their original sovereignty. The Supreme Court has recognized that although these powers are subject to the paramount authority of Congress, such powers do not spring from the United States . . . . [Therefore] it is concluded

mobile to an Indian tribe for governmental or police purposes is subject to applicable federal excise taxes.[118]

Each of these rulings on the taxation of transactions involving Indian tribal governments has recognized the similarity between tribal governments and states with respect to their sovereign powers and responsibilities to provide needed government services to their people.[119] Yet, despite its statutorily unsupportable position in Revenue Ruling 67-284 that Indian tribes themselves are not taxable entities, the IRS in all other transactions has found that language in the Code extending favorable tax treatment to "the United States, any State, [or] any political subdivision thereof"[120] cannot be extended to include tribes. The Internal Revenue Service, like other branches of the federal government, has difficulty placing Indian tribes within such neat conceptual pigeonholes.[121] Instead, tribes are treated as anom-

---

> that the Zuni Indian Pueblo is not a political subdivision of the United States [or any state] as that term is used in section 2055(a)(1) of the Code.

Id. at 280.

[118] Rev. Rul. 58-610, 1958-2 C.B. 815. "[A] tribe is not a political subdivision or agency of a state . . . within the scope of the exemption provided by section 4224 of the Code pertaining to sales made for the exclusive use of a state or political subdivision thereof." *Id.* at 816.

The IRS's excise tax branch has been vigilant in enforcing this particular ruling, sending letters to tribes, such as the Navajo Nation, reminding the tribal council of its liability to pay excise taxes on the purchase of tribal vehicles. *Hearing on Miscellaneous Tax Bills Before the House Comm. on Ways and Means*, 94th Cong., 2d Sess. 23 (1976) (statement of Richard Schifter, member of Fried, Frank, Harris, Shriver & Kampelman, on H.R. 8989) [hereinafter cited as *Tax Bills Hearing*].

[119] *See, e.g.*, Rev. Rul. 74-179, 1974-1 C.B. 279, 280; *Tax Bills Hearing*, *supra* note 118, at 23 (statement of Richard Schifter).

[120] I.R.C. § 2055 (a)(1) (1982); *see also id.* § 103(a)(1) ("a State, a Territory, or a possession of the United States, or any political subdivision of any of the foregoing").

[121] *See, e.g.*, Cherokee Nation v. Georgia, 30 U.S. (5 Pet.) 1 (1831). This Supreme Court case represents one of the first attempts by a branch of the federal government to define the status of indigenous American tribal nations within the federal system. In *Cherokee Nation*, the State of Georgia was attempting to impose its laws on the tribe. The Cherokees objected to the state's assertion of jurisdiction and filed suit with the Supreme Court under Article III of the U.S. Constitution, which gives the Court original jurisdiction in cases and controversies involving states and foreign nations. U.S. CONST. Art. III, § 2. The Court, in an opinion by Chief Justice Marshall, held that the Cherokees and other tribes were not foreign nations as understood in Article III, but rather "domestic dependent nations." *Cherokee Nation*, 30 U.S. (5 Pet.) at 17. They were therefore not within the Court's original jurisdiction. *Id.* at 18. *But see id.* at 33, (Thompson, J., dissenting) (the Cherokees were a foreign nation within the meaning of Article III and thus within the Court's original jurisdiction).

Marshall wrote at length of the "special relationship" that existed between Indian tribes and the federal government. *Id.* at 15–17. Attempting to define this unique relationship, Marshall asked the rhetorical question:

> Do the Cherokees constitute a *foreign* state in the sense of the constitution?
>
> . . . The condition of the Indians in relation to the United States, is perhaps, unlike that of any other two people in existence. In general, nations not owing a common allegiance are foreign to each other. The term *foreign* nation is,

alies, enjoying some of the tax benefits of states and localities and suffering many of the burdens of ordinary taxpayers. As in so many other areas of federal Indian law, because of the unique status of Indian tribes within our federal system, tribal governments find themselves lost within the seams of a complex and ill-suited statutory web.[122]

B. *Legislative History of the Tribal Tax Status Act: 1975–1982*

In response to the inconsistent positions taken by the Internal Revenue Service with respect to the tax status of Indian tribes, legislation was introduced in Congress in 1975.[123] The legislation, H.R. 8989, specifically proposed amending the Internal Revenue Code to provide "recognized Indian tribes (and their subdivisions) . . . performing governmental functions" a tax status similar to that granted to state and local governments.[124]

> with strict propriety, applicable by either to the other. But the relation of the Indians to the United States is marked by peculiar and cardinal distinctions which exist nowhere else . . . . Though the Indians are acknowledged to have an unquestionable, and heretofore unquestioned, right to the lands they occupy, until that right shall be extinguished by a voluntary cession to our government; yet it may well be doubted, whether those tribes which reside within the acknowledged boundaries of the United States can, with strict accuracy, be denominated foreign nations. They may, more correctly, perhaps, be denominated domestic dependent nations . . . . These considerations go far to support the opinion, that the framers of our constitution had not the Indian tribes in view, when they opened the courts of the Union to controversies between a state or the citizens thereof and foreign states.

Id. at 15–17 (emphasis in orginal).

[122] *See* Cohen, *The Spanish Origin of Indian Rights in the Law of the United States*, 31 GEO. L.J. 1 (1942).

> To trace the origins of our Federal Indian law is a difficult task. The law of the United States with respect to Indian tribes is a curious historical patchwork in which may be found the product of many looms and many weavers. One may divide this strange patchwork into its component patches and find nearly four hundred federal treaties with Indian tribes, about four thousand federal statutes, and an even larger number of judicial and administrative decisions which, by and large, attempt to interpret and to apply these treaties and statutes.

Id.

[123] *See* H.R. 8989, 94th Cong., 1st Sess., 121 CONG. REC. H25,304 (1975).

[124] *Tax Bills Hearings*, *supra* note 118, at 19 (summary of bill).

The purpose and intent of the legislation was stated in section two of the bill, where it was noted that Congress, the executive, and the courts, "from the earliest days of the Republic," had repeatedly recognized and affirmed "the governmental status and powers of Indian tribes." H.R. 8989, *supra* note 123, § 2. Despite this previous recognition, Indian tribal governments had been subject to various federal taxes not applicable to states and local governments. Legislation exempting Indian tribal governments from such taxes would, therefore, "be consistent with the Federal laws and treaties recognizing the governmental status of Indian tribes and with the role of Indian tribes in providing public service to Indians." *Id.*

Under the bill, taxes on real property and income paid to tribal governments would be deductible for federal tax purposes.[125] Charitable contributions to a tribal government would be tax-deductible.[126] The bill also proposed treating tribes similarly to states with respect to exemptions and credits for various federal excise taxes, contributions to political candidates for tribal office, retirement income derived from tribal employment, tribal contributions for employee annuities, and tribal scholarship and fellowship grants.[127] Finally, H.R. 8989 proposed similar treatment of the unrelated business income of a tribal government or subdivision to that accorded state and local government.[128]

Perhaps the most important aspects of the bill were provisions enabling Indian tribes to issue tax-exempt debt obligations to finance their governmental functions under section 103 of the Internal Revenue Code.[129] The bill also granted tribes the authority to issue private activity Industrial Development Bonds (IDBs), but prohibited a tax exemption for the interest on a tribal IDB "used in any trade or business carried on outside of the area reserved by federal Statute or treaty to the Indian tribe issuing the bond."[130]

Testimony on H.R. 8989 included strongly supportive statements by the Treasury Department, tribal government leaders, and national Indian rights groups.[131] Witnesses stressed the need

---

[125] H.R. 8989, *supra* note 123, § 7.

[126] *Id.* § 8.

[127] *Id.* §§ 4, 6, 9, 13, 14.

[128] *Id.* § 10. The statutory test applicable to states and municipalities concerning their income distinguishes between essential government functions and proprietary activities. Income is exempt only when derived from a public utility or an essential government function. I.R.C. § 115(1) (1982). *See* Barsh, *supra* note 16, at 554.

[129] H.R. 8989, *supra* note 123, § 5.

[130] *Id.* Industrial Development Bonds are also sometimes referred to as Industrial Revenue Bonds (IRB's). While minor mechanical distinctions between the two terms are recognized by the municipal bond community, *see infra* notes 203–04, the term Industrial Development Bond will be used throughout this Article to refer to a tax-exempt state or local government debt obligation, the proceeds of which are used as an economic development incentive to benefit a private business.

[131] *See Tax Bills Hearings*, *supra* note 118, at 20–43.

The Treasury Department gave its support to the legislation, provided that the bill's definition of "recognized Indian tribes" enjoying favored tax status be "restricted to those tribes performing *substantial* governmental functions." *Id.* at 21 (statement of Richard Schifter) (emphasis in original). The Treasury Department pointed to revenue sharing legislation definitions contained in the State and Local Fiscal Assistance Act of 1972, § 108, 31 U.S.C. § 6701(c) (1982), as an appropriate guide for determining which groups should be eligible for treatment under the bill. *Id.* at 20. Under § 108(b)(4) of the State and Local Fiscal Assistance Act, a separate entitlement is provided to any "Indian tribe or Alaskan Native village which has a recognized governing body which performs substantial governmental functions." State and Local Fiscal Assistance Act

for the legislation and pointed to the many laws enacted by Congress that had given Indian tribes "the substance as well as the form of self-government," which necessitated the same tax treatment for tribes provided to other governmental units.[132] According to one tribal attorney, enactment of H.R. 8989:

> would be one of the single most positive steps taken by Congress to better the condition of American Indian tribes in the United States since the Indian Reorganization Act of 1934, and could be the first step towards the long sought realization by Indian tribes of the chance for economic independence and an opportunity to strengthen self-government and provide the many services that are needed by members of the Tribes and which are not met through federal appropriations and available tribal income.[133]

Despite favorable House committee action,[134] H.R. 8989 was never considered by the full House prior to adjournment of the

---

of 1972, § 108(b)(4), 31 U.S.C. § 6701(a) (1982). *See infra note* 134 (describing subsequent amendments to H.R. 8989, including the adoption of the Treasury's recommended "substantial governmental functions" language).

[132] *See, e.g., Tax Bills Hearings, supra* note 118, at 24 (statement of Richard Schifter); *see also Hearings Before the Subcomm. on Miscellaneous Tax Bills, supra* note 66, at 109 (statement of William E. Sudow); *id.* at 114 (statement of Owen M. Panner, representative of Confederated Tribes of the Warm Springs Reservation of Oregon) (discussing a substantially similar bill which was later introduced in the House during the 95th Congress, *see infra* note 135). The statutes cited as furthering Indian self-government included: The Indian Self-Determination Act, § 2, 25 U.S.C. § 450 (1982); The Indian Reorganization Act of 1934, §§ 1–19, 25 U.S.C. §§ 461–79 (1982); The State and Local Assistance Act of 1980, § 108(d)(1), 31 U.S.C. § 6701(a)(5)(B) (1982) (specifically including Indian tribes in the revenue sharing program); The Comprehensive Employment and Training Act of 1973, § 2, 29 U.S.C. § 801 (1976) (repealed 1982); The Housing and Urban Development Act of 1969, § 414, 40 U.S.C. § 484(b) (1982) (repealed 1983); The Public Works and Economic Development Act of 1965, § 101, 42 U.S.C. § 3131(a) (1982); The Economic Opportunity Act of 1964, § 104, 42 U.S.C. § 2790(f) (1976) (repealed 1981).

The Indian Reorganization Act suspended the allotment of tribal lands, *see supra* notes 41–45 and accompanying text, and provided for financing repurchases of tribal lands. Indian Reorganization Act of 1934, § 5, 25 U.S.C. § 463(b) (1982). It also established a revolving credit fund to promote economic development on reservations. *Id.*, § 10, 25 U.S.C. § 470. In addition, the Indian Reorganization Act laid the foundation for a more modern framework for tribal self-government by allowing tribes to adopt constitutions and corporate charters, *id.*, §§ 16–17, 25 U.S.C. §§ 476–77, thus giving greater recognition to the sovereign status of Indian tribes. *But see supra* notes 79–81 and accompanying text (discussing the problems of imposing non-Indian political and social values and institutions into tribal life).

[133] *Tax Bills Hearings, supra* note 118, at 35–36 (statement of Richard A. Baenen, member of Wilkinson, Cragun & Barker).

[134] H.R. 8989 was reported favorably out of the House Ways and Means Committee, and recommended for passage to the full House with minor amendments. H.R. REP. No. 1693, 94th Cong., 2d Sess. 1 (1976). Included among the amendments by the House Ways and Means Committee was the suggestion requested by the Treasury Department, *see supra* note 131, tightening the definition of "recognized Indian tribe" to include any "tribe, band, community, village, or group of Indians or Alaska Natives which is

94th Congress. Nonetheless, Indian tribes and their supporters continued their lobbying efforts on behalf of the legislation in subsequent sessions of Congress.[135]

In the 97th Congress, hearings were held[136] in the Senate before the Subcommittee on Taxation and Debt Management of the Senate Committee on Finance on a version of the Tribal

---

recognized by the [Treasury] Secretary, after consultation with the Secretary of the Interior, as performing governmental functions." H.R. 8989, *reprinted in* H.R. REP. No. 1693, *supra*, at 2–3.

Also, the House Ways and Means Committee amended the section of the original bill pertaining to IDBs, clarifying that the interest exemption on tribal IDBs would still apply even when activities supported by the bond issue were carried on beyond the reservation boundaries, so long as substantially all of those activities related to "purchasing, marketing, or similar activities directly related" to the reservation IDB financed trade or business. *Id.* § 4, *reprinted in* H.R. REP. No. 1693, *supra*, at 2–3. Furthermore, tax-exempt status would not be available for interest on any tribal obligation, other than an IDB, if all or a major portion of the proceeds were to be used directly or indirectly in an industrial or commercial activity. *Id.*

In stating its reasons for supporting the legislation, the Committee noted:

> Indian tribes and Alaskan villages perform many of the same functions as do municipal and State governments . . . includ[ing] law enforcement, water, sewage, and garbage services, business licensing and regulation, land use planning, housing, social and health programs, legal services, natural resource development and other activities . . . . [I]n order to assist the Indian tribes and Alaskan native villages in carrying out their self-governing responsibilities, those tribes which perform substantial governmental functions should be provided with substantially the same tax status as that enjoyed by other governmental units.

H.R. REP. No. 1693, *supra*, at 4.

[135] The Indian Tribal Governmental Tax Status Act was reintroduced in the House of Representatives during the 95th Congress by Rep. Ullmann (D-Or.) and Rep. Udall (D-Ariz.), among others. *See* H.R. 4089, 95th Cong., 1st Sess., 123 CONG. REC. H5372 (1977). H.R. 4089 was identical to the version of H.R. 8989, reported out of the House Ways and Means Committee during the previous congressional session, and witness testimony on the substance of the bill was once again universally favorable. *See, e.g.*, *Hearings Before the Subcommittee on Miscellaneous Revenue Measures*, *supra* note 66, at 103–28 (statement of Albert G. Fiedler, Jr., Manager, Communications Services, Standard Oil Co. (Indiana)) (describing the tax treatment of Indian tribes and Alaskan villages). Again, the bill passed the House Ways and Means Committee. *See* H.R. REP. No. 843, 95th Cong., 1st Sess. 1 (1978) (accompanying H.R. 4089). In its report, the House Ways and Means Committee stated its intention that, for the specified sections of the Internal Revenue Code:

> Indian tribal governments are to be treated the same as States or similar to States. The committee intends that generally the provisions of the bill are to be applied in such a manner as to achieve parity under these provisions between Indian tribal governments (and their subdivisions) and State governments (and their political subdivisions).

Id. at 4. Again, however, Congress adjourned before taking final action on the bill. *See* 127 CONG. REC. S5666 (daily ed. June 2, 1981) (remarks of Sen. Wallop). The Indian Tribal Governmental Tax Status Act was resubmitted in the 96th Congress, but no action was taken on it during that session. *See id.* at S5667.

[136] *See 1981-82 Miscellaneous Tax Bills, XVI, Hearing Before the Subcomm. on Taxation and Debt Management of the Senate Comm. on Fin. on S. 1298, S. 2197, and S. 2498*, 97th Cong., 2d Sess. (1982) [hereinafter cited as *1981-82 Miscellaneous Tax Bills Hearing*].

Tax Status Act[137] virtually identical to that first introduced in 1975.[138] As in past sessions, testimony from tribal leaders, Indian support groups, and administration officials was universally supportive of the bill's passage.[139] William McKee, Tax Legislative Counsel for the Department of the Treasury, noted that the Treasury supported the bill on the simple "principle that similarly situated taxpayers should be treated alike."[140] He stated:

> Indian tribes and tribal members occupy a unique role in our scheme of government, and the taxation of Indian tribal

---

[137] S.1298, 97th Cong., 1st Sess., 127 CONG. REC. S5666 (daily ed. June 2, 1981). Sen. Wallop, a cosponsor of the bill in the Senate, noted in his comments submitting the legislation that tribal governments had assumed greater responsibility for financing their own governmental functions and services, but that these efforts to raise funds through tribal taxes have "been only partially successful" due to the differential tax treatment of tribes by the Internal Revenue Service. 127 CONG. REC. S5666-67 (daily ed. June 2, 1981) (remarks of Sen. Wallop). "This difference in treatment undermines the tax initiatives of tribal governments and seriously interferes with their efforts to improve the conditions of life in Indian country." *Id.* Sen. Wallop also remarked:

> This legislation would enhance the opportunities of Indian tribal governments to provide essential governmental services to these tribal members by according Indian tribal governments the same tax status which is currently accorded state, county and municipal governments. Enactment of this bill would strengthen tribal governments significantly by providing additional sources of financing and by eliminating the unfair burden of taxes Indian tribal governments must now pay.

Id. at S5666.

In the House, Rep. Jones (D-Okla.), a cosponsor of the bill, stated in his introductory comments that Indian tribes and Alaskan Native villages provided many of the same essential public services as states and municipalities. 127 CONG. REC. H2571 (daily ed. June 2, 1981) (remarks of Sen. Wallop). Despite their exercise of these important governmental functions, Rep. Jones noted that Indian tribal governments had been unfairly denied numerous forms of favorable tax treatment enjoyed by other governmental units:

> [T]his distinction in treatment clearly is at odds with Federal Indian Policies promulgated for the past several administrations. Specifically, for the past half century, the Federal government virtually without exception has sought to strengthen tribal governments politically and economically . . . . [This difference] seriously interferes with their efforts to improve conditions of life in Indian country.

Id.

[138] *See supra* notes 123–30 and accompanying text.

[139] *See, e.g., 1981-82 Miscellaneous Tax Bills Hearing, supra* note 136, at 70–146.

> It should be noted that this legislation would not empower tribal governments to exercise any governmental powers which they now do not have, neither would it extend to tribal governments any benefits now not extended to other governments; it merely would end a discriminatory application of the Internal Revenue Code toward tribal governments.

Id. at 78 (statement of Judy Knight, Tribal Treasurer, Ute Mountain Tribe). "It is a good bill. It is one Indians have worked for for years. There is no local opposition to it. It will help us and the federal government because the stronger the tribal governments are, the less the United States must do itself." *Id.* at 85 (statement of Alfred Ward, Co-Chairman of the Business Council of the Shoshone Indian Tribe of the Wind River Reservation, Wyo.).

[140] *Id.* at 58 (statement of William McKee, Tax Legislative Counsel, U.S. Dep't of Treas.).

> members and tribal governments is a matter of some complexity. Treasury believes, however, that if Indian tribal governments perform the same essential governmental functions for their members as State and local governments perform for their residents, then the fundamental tax policy interest in promoting horizontal equity suggests that tribal governments should be treated for Federal tax purposes in the same manner as State and local governments. Although the status of Indian tribes in our scheme of government is unique, there is sufficient analogy between the status of states and the status of Indian tribes to support treating tribal governments on the same basis as State and local governments.[141]

Roy H. Sampsel, the Interior Department's Deputy Assistant Secretary for Indian Affairs, stated that:

> [the bill] would, at very little cost to the federal treasury, provide badly needed benefits to Indian tribes. Tribal governments today are more dependent on federal funds than are most State and local governments. It is a goal of this Administration and of many tribal leaders that this dependency be lessened. As the federal budget becomes more restricted, tribal governments, like State and local governments, will have to find new ways to support the delivery of essential government services. But without enactment of the bill, tribes will be very hard pressed to do this.[142]

The Committee on Finance's report to the full Senate recommended passage of the bill,[143] which was incorporated as Title II of the House originated Periodic Payment Settlement Act (H.R. 5470).[144] Disagreement in the House as to the Senate version of H.R. 5470, which now included the Tribal Tax Status

---

[141] *Id.* at 57 (statement of William McKee). McKee qualified the Treasury's endorsement of S. 1298 by stating:

> S. 1298 would extend benefits only to tribal governments which do exercise substantial governmental functions. Thus, insofar as Congress exercised its power to defease tribes of sovereignty, and hence to make those tribes less like States, the tribes so defeased would not be entitled to the bill's benefits . . . . In endorsing S. 1298, we neither endorse nor question the desirability of the provisions of the tax law whose benefits are extended to tribal governments by this bill. Treasury does, however, strongly endorse the principle that taxpayers who are similarly situated should be treated alike for tax purposes if the law is to be applied fairly and equitably. That principle underlies our support of this bill.

Id. at 58.

[142] *Id.* at 61–62 (statement of Roy Sampsel, Deputy Ass't Sec'y of Indian Affairs, U.S. Dep't of Interior). Sampsel further stated that the Interior Department was "not suggesting that this bill alone would be a fast solution to the economic problems on Indian reservations, but it would be an important step toward removing some of the impediments to tribal economic development." *Id.* at 62.

[143] S. REP. No. 646, 97th Cong., 2d Sess. 8–17 (1982).

[144] H.R. 5470, 97th Cong., 2d Sess. §§ 201–04 (1982).

Act, however, led to a request for a House-Senate conference.[145] The subsequent conference agreement resulted in several important substantive changes to the original version of the Tribal Tax Status Act.[146]

The most important of the changes was the addition of a provision substantially modifying the authority of tribal governments to engage in tax-exempt financing. This provision limited the tax exemption for interest paid on tribal bonds to instances in which the proceeds of such bonds "are used in an essential governmental function (such as schools, streets, and sewers)."[147] Furthermore, the conference agreement, without explanation, specifically prohibited tribal governments from issuing any "private-activity bonds," such as industrial development bonds.[148] This limitation had never been included in any of the previous versions of the Tribal Tax Status Act.[149] In fact, authority to

---

[145] 128 CONG. REC. S15,199 (daily ed. Dec. 16, 1982, Part III).

[146] *See* H.R. REP. No. 984, 97th Cong., 2d Sess., 11–15 (1982).

Actually, the Senate-added Tribal Tax Status Act was deleted from the Periodic Payment Settlement Act on the House floor. The Senate disagreed with this deletion and a conference was called. 128 CONG. REC. S15,199 (daily ed. Dec. 16, 1982, Part III). Rep. Gibbons (D-Fla.) had objected to the Tribal Tax Status Act because of concerns over Indian bingo operations, and the possible use of the Tribal Tax Status Act's IDB provisions in support of such tribal enterprises. The conference agreement resulted in the deletion of the IDB section in the Tribal Tax Status Act, and the inclusion of short-term sunset provisions with respect to various benefits conferred to tribes under this Act. H.R. REP. No. 984, 97th Cong., 2d Sess. at 17 (1982). Letter from Suzan Shown Harjo, former legislative liaison, Fried, Frank, Harris, Shriver & Kampelman, Washington, D.C., to author (Sept. 22, 1983) [on file at HARV. J. ON LEGIS.] [hereinafter cited as "Letter from Shown Harjo"].

[147] H.R. REP. No. 984, 97th Cong., 2d Sess. 17 (1982).

[148] *Id.* In addition, the Tribal Tax Status Act as passed included sunset provisions limiting exercise of tribal public activity bonding and other important benefits to taxable years beginning before January 1, 1985. Indian Tribal Government Tax Status Act of 1982, Pub. L. No. 97-473, § 204, 96 Stat. 2607, 2611 (codified as amended at 26 U.S.C. § 7871 note). An amendment to the Tribal Tax Status Act passed by Congress in July, 1984 eliminated the sunset provisions. Tax Reduction Act of 1984, Pub. L. No. 98-369, § 1065(a), *reprinted in* 1984 U.S. CODE CONG. & AD. NEWS (98 Stat.) 494, 1048 (to be codified at 26 U.S.C. § 7871 note).

[149] Every prior version of the Tribal Tax Status Act had included specific authority for tribes to engage in at least some limited form of tax-exempt IDB financing. *See supra* notes 130, 134–35 and accompanying text. Nor had any witness or member of the Senate or House, in printed hearings, committee reports, or debate, ever voiced objection to the principle of providing Indian tribes a tax status similar to states and local governments with respect to IDB financing. In fact, the Treasury Department, which sought to curtail and limit the IDB tax exemption generally, specifically endorsed all earlier versions of the Tribal Tax Status Act providing Indian tribes with IDB tax-exempt authority. *See, e.g.*, *Tax Bills Hearing*, *supra* note 118, at 20–21, 124, 125. The Treasury Department's support was based on "the principle that taxpayers who are similarly situated should be treated alike for tax purposes." *1981-82 Miscellaneous Tax Bills Hearing*, *supra* note 136, at 58 (statement of William McKee).

issue such bonds was regarded by tribal governments as one of the most important features of the proposed Tribal Tax Status Act.[150] Despite this important substantive modification of the Tribal Tax Status Act, prompted apparently in part by concerns of some members of Congress over Indian bingo operations and the use of IDBs to finance such enterprises,[151] the House and Senate both accepted the conference report on H.R. 5470,[152] and the bill as amended received final House and Senate approval in the closing days of the 1982 legislative session.[153]

As signed into law by the President in January 1983,[154] the most significant tribal economic development initiatives contained within the Tribal Tax Status Act include the following: first, the grant of a deduction from federal income tax for taxes paid to an Indian tribe; second, a provision that charitable contributions to Indian tribal governments are deductible for income, estate, and gift tax purposes; third, an exemption for Indian governments from a variety of federal excise taxes; and fourth, the grant of an income tax exemption on the interest of certain limited classes of Indian tribal government debt obligations.[155]

---

[150] *See, e.g., 1981–82 Miscellaneous Tax Bills Hearing, supra* note 136, at 71–74, 83–85, 135–40 (statements of tribal leaders emphasizing importance of tribal tax-exempt bonding authority to finance economic development projects).

[151] *See supra* note 146. The inclusion of this limitation may also have been due in large part to the inhospitable attitude that generally existed in the 97th Congress towards IDBs. *See infra* notes 218–26 and accompanying text (discussion of criticism of IDBs and congressional response, including significant restrictions on IDB financing contained in the Tax Equity and Fiscal Responsibility Act of 1982).

[152] 128 CONG. REC. H10,672–73, S15,987–88 (daily ed. Dec. 21, 1982).

[153] *Id.* at H10,720, S16,069.

[154] 19 WEEKLY COMP. PRES. DOC. 96 (Jan. 24, 1983).

[155] Indian Tribal Governmental Tax Status Act of 1982, Pub. L. No. 97-473, § 202(a), 96 Stat. 2607, 2608–09 (codified as amended at I.R.C. § 7871). The Tribal Tax Status Act grants important tax advantages to Indian tribes, including the same preferential treatment as states respecting:

(1) The tax on unrelated business income of tribal higher educational institutions;

(2) The taxes imposed on certain prohibited transactions by tribal charities and foundations;

(3) The income tax credit for individuals who receive retirement income from tribal retirement systems;

(4) Eligibility for certain tax-deferred annuities;

(5) The income tax credit for tribal campaign contributions; and

(6) The exclusion of tribally-sponsored scholarships and fellowships.

Id. In the 1984 amendments to the Tribal Tax Status Act, Congress also extended equivalent tax treatment to tribes respecting:

(1) Treatment of payments under certain state accident and health plans;

III. THE TRIBAL TAX STATUS ACT AND ITS IMPLICATIONS FOR ECONOMIC DEVELOPMENT IN INDIAN COUNTRY

A. *Eligibility for Tax Benefits*

Under the Tribal Tax Status Act, Indian tribal governments are extended numerous tax advantages currently enjoyed by states and their subdivisions. In order to qualify for these benefits, however, the Indian government must meet the definitional standard set out in section 203 of the Tribal Tax Status Act: "The term 'Indian tribal government' means the governing body of any tribe, band, community, village, or group of Indians, or (if applicable) Alaska Natives, which is determined by the [Treasury] Secretary, after consultation with the Secretary of the Interior, to exercise governmental functions."[156]

In addition, subdivisions established by an Indian tribal government may also be accorded the same favorable tax treatment received by subdivisions of a state, if the Treasury Secretary "determines (after consultation with the Secretary of the Interior) that such subdivision has been delegated the right to exercise one or more of the substantial governmental functions of the Indian tribal government."[157]

In order to qualify for the tax status provided under the Tribal Tax Status Act, therefore, a tribal government, or its subdivision, must first be recognized by the Treasury Department, after consultation with the Interior Department, as exercising "substantial governmental functions." Legislative history on the meaning of the term "substantial governmental functions," indicates that only those tribes exercising sovereign powers, such as taxation, eminent domain, zoning, police protection, and fire protection, are to be recognized by the Treasury Department as

---

(2) Deductibility of certain costs associated with appearances before the state legislatures; and

(3) Rules relating to treatment of original issue discount on certain state obligations.

Tax Reduction Act of 1984, Pub. L. No. 98-369, § 1065(b), *reprinted in* 1984 U.S. CODE CONG. & AD. NEWS (98 Stat.) 494, 1048 (to be codified at I.R.C. § 7871(a)(6)). *See* H.R. REP. No. 861, 98th Cong., 2d Sess. 1272–73 (1984), *reprinted in* 1984 U.S. CODE CONG. & AD. NEWS 751, 1266–67. As these important benefits are not likely to have a significant impact on major tribal economic development initiatives, they are not analyzed in this Article.

[156] Indian Governmental Tribal Tax Status Act of 1982, § 203, I.R.C. § 7701(a)(40) (1982 & Supp. I 1983).

[157] *Id.* § 202(a), I.R.C. § 7871(d).

eligible for the Act's benefits.[158] Similarly, a tribal subdivision will qualify for favorable tax status only if the tribe has delegated to it the right to the sovereign exercise of one or more such powers.[159] Ultimately, it can be expected that the Internal Revenue Service will continue to publish periodically lists of tribes and subdivisions that have been recognized for purposes of the Tribal Tax Status Act, similar in form and function to the IRS's "blue book" of organizations eligible to receive deductible charitable contributions.[160] Such a list of tribes and subdivisions eligible for the benefits afforded under the Tribal Tax Status Act will enable the tribe and those contemplating transactions with a tribal government to determine in advance the potential tax consequences of their activities.[161]

B. *Deductibility of Tribally Imposed Taxes*

Under the Tribal Tax Status Act, if a tax imposed by an Indian tribe falls under any of the categories of state and local taxes that may be deducted under section 164 of the Internal Revenue Code, then the tribal tax is also deductible.[162] The provision of a federal income tax deduction for taxes paid to tribal governments is one of the most significant features of the Tribal Tax

---

[158] *See Tax Bills Hearing*, *supra* note 118, at 19 (report from the U.S. Treas. Dep't on H.R. 8989).

[159] *See* 128 CONG. REC. H10,650 (daily ed. Dec. 21, 1982); *see also* H.R. REP. No. 843, *supra* note 135, at 5.

> It is intended that [the] essentially equivalent criteria be used in making determinations as to delegations of sovereign powers by Indian tribal governments as to their subdivisions and delegations of sovereign powers by States to their political subdivisions.
>
> If such a determination is made, that is to be sufficient to cause the subdivision to be treated as a political subdivision of a State. It is not necessary for a whole range of sovereign powers to be delegated, in order for the subdivision to be treated as a political subdivision of a State; it is sufficient if at least one sovereign power has been so delegated. Also, it is not necessary that the subdivision in fact exercise that power at any given time, so long as the Indian tribal government has in fact delegated to the subdivision the right to exercise the power.

Id.

[160] *See id.*

[161] *See, e.g.*, BUREAU OF INDIAN AFFAIRS, DEP'T OF THE INTERIOR, INDIAN TRIBAL ENTITIES THAT EXERCISE GOVERNMENTAL FUNCTIONS FOR PURPOSES OF P.L. 97-473—INDIAN TRIBAL GOVERNMENTAL TAX STATUS ACT OF 1982 (BIA's list of 251 federally recognized tribes, including 62 Alaskan villages, which exercise substantial governmental functions, transmitted to the IRS) [on file at HARV. J. ON LEGIS.]. *See also* Rev. Proc. 84-37, 1984-1 C.B. 513; Rev. Proc. 84-36, 1984-1 C.B. 510; Rev. Proc. 83-87, 1983-2 C.B. 606.

[162] Indian Governmental Tribal Tax Status Act, § 202(a), I.R.C. § 7871(a)(3) (1982).

Status Act. In conjunction with the Supreme Court's recent affirmation of the sovereign right of Indian tribes to tax in *Merrion v. Jicarilla Apache Tribe*,[163] this section affords Indian governments a greater degree of flexibility in devising revenue raising tax initiatives than was previously available.

Prior to the Tribal Tax Status Act, taxes imposed by a tribal government could only be deducted, if at all, under section 162 of the Code, which allows deductions in computing federal taxable income only for expenses paid or incurred in a business context.[164] If an individual subject to a tribal tax could not establish that the tax was incurred in a business context as defined by section 162, as was usually the case with tribally imposed personal income taxes, then the taxpayer could not deduct the tax for federal income tax purposes.[165]

Under the Tribal Tax Status Act, the ability to deduct all tax payments to an Indian tribe will enhance tribal taxpayer acceptance of this important tribal power. Although certainly no taxpayer will enjoy paying newly imposed tribal taxes, the ability to deduct tribal taxes from federal income taxes renders such taxes no more onerous than similar state and locally imposed taxes.[166]

In essence, the Tribal Tax Status Act provides tribal governments the opportunity to devise numerous broadly based taxing schemes in order to raise revenues for essential governmental services. Because tribal taxes are now included under section 164, tribally imposed real property taxes, personal property taxes, income taxes, excess profits taxes, and general sales taxes

---

[163] 455 U.S. 130 (1982). *See supra* note 107; Note, Merrion v. Jicarilla Apache Tribe: *Tribal Power to Tax Non-Indian Lessees Who Exploit Reservation Natural Resources*, 26 S.D.L. REV. 595 (1981).

[164] *See* I.R.C. § 162(a) (1982). This section provides a deduction for "all ordinary and necessary expenses paid or incurred during the taxable year in carrying on any trade or business, including:

> (1) a reasonable allowance for salaries or other compensation for personal services actually rendered;
>
> (2) travelling expenses . . . while away from home in the pursuit of a trade or business;
>
> (3) and [certain trade or business rental payments]."

Id.

[165] *See 1981-82 Miscellaneous Tax Bills Hearing*, *supra* note 136, at 73 (statement of Delbert Frank, Chairman of the Tribal Council of the Confederated Tribes of the Warm Springs Reservation, Wash.).

[166] *Id.* at 81 (statement of Burton Hutchinson, Chairman, Arapahoe Business Council). Of course, if a taxpayer does not itemize deductions, then he or she cannot take advantage of this benefit. Unfortunately, many tribal members subject to a tribal tax will find themselves in this situation, given the fact that many of these individuals earn relatively low wages.

are now deductible for federal income tax purposes.[167] For example, under the Tribal Tax Status Act, a tribal tax on income earned on the reservation would be fully deductible. The tribe can structure such an income tax so as to tax only those individuals earning above a set income amount. Poorer wage earners who probably do not itemize deductions will not be subject to the tax, while those earning higher wages will be required to pay the tribal income tax. Ideally, tribes can now impose taxes at rates proportionate to states and their subdivisions, and can devote such tax revenues to providing the type of public infrastructure necessary to support a diversified tribal economy.

C. *Deductibility of Charitable Contributions to Indian Tribal Governments*

The Internal Revenue Code provides that charitable contributions to or for the use of a state or political subdivision are generally deductible for federal income,[168] estate,[169] and gift tax[170] purposes, but only if the contribution is made for "exclusively public purposes."[171] The Tribal Tax Status Act amends the Code to include charitable contributions, gifts, bequests, and donations made to an Indian tribal government as deductible expenses.[172] This portion of the Tribal Tax Status Act specifically overrules the Internal Revenue Service's previous ruling, which held that a bequest to an Indian tribe did not qualify for the estate tax charitable contribution deduction under section 2055(a) of the Code.[173]

Numerous supporters of the Tribal Tax Status Act pointed out the advantages of this favorable tax treatment of contribu-

---

[167] *See* I.R.C. § 164(a) (1982) (list of state and local taxes deductible under that section).

[168] I.R.C. § 170(a) (1982).

[169] I.R.C. § 2055(a)(1) (1982); 26 U.S.C.A. § 2106(a)(2) (Dec. 1984 Supp.).

[170] I.R.C. § 2522(a)(1) (1982).

[171] *Id.* § 170(c)(1) (definition of charitable contribution).

[172] Indian Tribal Governmental Tax Status Act, § 202(a), I.R.C. § 7871(a)(1) (1982 & Supp. I 1983). An Indian tribal government shall be treated as a state for purposes of determining the tax-deductibility of any contribution or transfer to or for the use of such government or its political subdivision under:

> (A) section 170 (relating to income tax deduction for charitable, etc., contributions and gifts),
> (B) sections 2055 and 2106(a)(2) (relating to estate tax deduction for transfers of public, charitable, and religious uses), or
> (C) section 2522 (relating to gift tax deduction for charitable and similar gifts).

Id.

[173] Rev. Rul. 74-179, 1974-1 C.B. 279; *see supra* note 117 and accompanying text.

tions to tribal governments.[174] Tribal leaders have consistently noted the reluctance of individuals and private organizations to make needed contributions to Indian tribes in the absence of assured federal tax-deductibility for such contributions.[175] The Tribal Tax Status Act clearly makes such contributions eligible for favorable tax treatment.[176] Thus, tribes can now seek contributions for numerous purposes such as promoting higher education for tribal members.[177] Also, an individual desiring to bequest property to a tribal government by his or her will is assured of a favorable federal estate tax treatment for such a contribution.[178] Such tax-deductibility will greatly facilitate the current efforts of several tribes to consolidate the reservation land base by encouraging bequests of fractionalized allotment interests to the tribe itself.[179]

D. *Tribal Excise Tax Exemptions*

State and local governments have historically benefited from a variety of federal excise exemptions.[180] The Tribal Tax Status Act extends many of these same exceptions to Indian tribal government, so that tribes will be able to realize significant tax savings in expenditures for their capital budgets.[181] Specifically, the Act provides tribal governments an excise tax exemption on purchases of items such as gasoline and other fuel oil products, communications equipment, firearms, and, ironically, bows and arrows.[182] Tribal governments are also exempted from certain highway use taxes under the new legislation.[183] In order to qual-

---

[174] *See, e.g., 1981-82 Miscellaneous Tax Bills Hearing, supra* note 136, at 81 (statement of Burton Hutchinson).

[175] *See, e.g., id.* at 82–83 (statement of Burton Hutchinson) (discussing the creation of the Arapahoe Educational Trust, which seeks tax-deductible contributions to promote higher education of members of the Arapahoe tribe).

[176] Indian Tribal Governmental Tax Status Act of 1982, § 202(a), I.R.C. § 7871(a)(1) (1982 & Supp. I 1983).

[177] *Id.* § 7871(a)(1)(A).

[178] *Id.* § 7871(a)(1)(B).

[179] *See generally supra* notes 36–58 and accompanying text (describing problems encountered by tribes because of fractionated land interests on reservation).

[180] *See, e.g.,* I.R.C. § 4041(g)(2) (1982) (relief from special fuels taxes); *id.* § 4221(a)(4) (relief from manufacturer's excise taxes); *id.* § 4483(b) (relief from motor vehicle taxes); and *id.* § 4253(i) (relief from excise taxes on communications services and facilities).

[181] S. REP. No. 646, 97th Cong., 2d Sess. 15–16 (1982), *reprinted in* 1982 U.S. CODE CONG. & AD. NEWS 4580, 4593.

[182] Indian Tribal Governmental Tax Status Act of 1982, § 202(a), I.R.C. § 7871(a)(2), (b) (1982 & Supp. I 1983).

[183] *Id.* § 7871(a)(2)(D), (b).

ify for these excise tax exemptions, the Act imposes the additional requirement that the transaction involve "the exercise of an essential governmental function of the Indian tribal government."[184]

1. Special Fuels Excise Tax Exemptions

The Tribal Tax Status Act grants to Indian governments the same exemption now enjoyed by state and local governments from the federal special fuels excise tax.[185] Hence, under the Act, excise taxes on liquids sold for use as fuels do not apply to sales for the exclusive use of any Indian tribal government or subdivision.[186] Where fuel is sold to a tribe tax-paid instead of tax-free, the tribe will receive the same benefits of the Code's refund and credit provisions that presently apply to state and local governments.[187]

2. Manufacturer's Excise Tax Exemptions

Under section 4221(a)(4) of the Code, a sale by a manufacturer of any article subject to the federal manufacturer's excise tax is exempt if the sale is to a state or local government for "the exclusive use" of that government.[188] Where the tax has been prepaid by the state or subdivision, Subtitle F of the Code provides for a credit or refund of the manufacturer's excise tax.[189] The Tribal Tax Status Act extends this manufacturer's excise tax exemption to tribal governments and their subdivisions.[190] This provision expressly overrules another Internal Revenue Service ruling which held that Indian tribes were not state or local governments and that sales to Indian tribes were therefore not exempt from the manufacturer's tax under the Code.[191]

Henceforth, all purchases in the exercise of an essential government function by a tribal government previously subject to the federal manufacturer's excise tax are now exempt. A tribe

---

[184] *Id.* § 7871(b).

[185] *Id.* § 7871(a)(2)(A), (b).

[186] *Id.*

[187] *See* 26 U.S.C.A. §§ 6416(b), 6427(g) (West Supp. 1984).

[188] I.R.C. § 4221(a)(4) (1982). "The term 'State or local government' means any State, any political subdivision thereof, or the District of Columbia." *Id.* § 4221(d)(4).

[189] *See id.* § 6416.

[190] Indian Tribal Governmental Tax Status Act of 1982, § 202(a), I.R.C. § 7871(a)(2)(B), (b) (1982 & Supp. I 1983).

[191] Rev. Rul. 58-610, 1958-2 C.B. 815; *see supra* note 118 and accompanying text.

purchasing equipment needed for the provision of governmental services, such as police vehicles, can realize the same dollar savings on such purchases as state and local governments can.

3. Communications Excise Tax and Highway Use Taxes

The Tribal Tax Status Act also extends to Indian governments the same tax status as states with respect to federally imposed excise taxes on communications services and facilities and highway use taxes.[192] Thus tribal purchases of communications equipment, such as telephone systems, are exempt from federal taxation.[193] Likewise, the use of any highway motor vehicle by the tribal government or its subdivision is also now exempt from the federal excise tax.[194] Both exemptions, however, are restricted to uses in the exercise of essential government functions.[195]

E. *Tax-Exempt Tribal Public Activity Bonds*

1. The Background: Differential Treatment of State and Local Governmental Debt Obligations and Tribal Debt Obligations Under Section 103 of the Code

Prior to the passage of the Tribal Tax Status Act in 1982, the Internal Revenue Service maintained that interest paid to an individual on debt obligations of an Indian tribe was not tax-exempt under section 103(a) of the Code.[196] The IRS's position placed Indian tribes at a severe financial disadvantage compared to state and local governments in the tribes' efforts to raise capital for financing needed governmental services and infrastructure.[197]

Under section 103(a) of the Internal Revenue Code, interest paid on state and municipal debt obligations is excluded from income for federal tax purposes as long as the particular obli-

---

[192] Indian Tribal Governmental Tax Status Act of 1982, § 202(a), I.R.C. § 7871(a)(2)(C)–(D), (b) (1982 & Supp. I 1983).

[193] *Compare* 26 U.S.C.A. § 4253(i) (West Supp. 1984) (state exemption) *with* 26 U.S.C.A. § 7871(a)(2)(C) (West Supp. 1984) (Indian tribal exemption).

[194] *Compare* 26 U.S.C.A. § 4483(a) (West Supp. 1984) (state exemption) *with* 26 U.S.C.A. § 7871(a)(2)(D) (West Supp. 1984) (Indian tribal exemption).

[195] 26 U.S.C.A. § 7871(b) (West Supp. 1984).

[196] Rev. Rul. 68-231, 1968-1 C.B. 48; *see supra* note 115 and accompanying text.

[197] *See* Barsh, *supra* note 16, at 553–54.

gation is issued to finance a customary governmental function or program, such as waterworks, schools, or streets.[198] Specifically, section 103 applies to interest on the obligations of "a State, a Territory, or a possession of the United States, or any political subdivision of any of the foregoing, or of the District of Columbia."[199] The Internal Revenue Service has ruled that the exclusion of interest contained in section 103(a) also extends to obligations issued "on behalf of" these state or local governmental units.[200] For instance, where a separate corporate body such as a turnpike or port authority is created to carry out a recognized governmental program, interest on the debt obligations issued by such a public corporation is also exempt for federal income tax purposes.[201]

The interest exemption on state and local government debt obligations permits these governmental units to borrow money[202] at below-market interest rates[203] to finance such vital public improvements as roads, bridges, water and sewer systems, and public buildings. Investors in high tax brackets, seeking to earn tax-free income, readily put their investment dollars into state and local government bonds at yields lower than other types of taxable investment opportunities. For instance, a taxpayer in the fifty percent tax bracket would pay no federal income tax on interest earned on a one-year, government bond of $1,000 bearing ten percent interest. He would receive a $100 tax-free gain on his investment. To earn the same amount of income on a taxable investment, such as a private corporation's one-year debt obligation, he would have to earn twenty percent on that same $1,000 investment because even though he would receive a $200 return, his federal income tax liability on that earned income would equal fifty percent, or $100.[204]

---

[198] *See* I.R.C. § 103(a)(1) (1982); Rev. Rul. 54-106, 1954-1 C.B. 28.

[199] I.R.C. § 103(a)(1) (1982).

[200] Rev. Rul. 54-106, 1954-1 C.B. 28.

[201] *See* Commissioner v. Shamberg's Estate, 144 F.2d 998 (2d Cir. 1944), *cert. denied*, 323 U.S. 792 (1945) (Port Authority of New York); Rev. Rul. 61-181, 1961-2 C.B. 21 (Los Angeles Metropolitan Transit Authority); Rev. Rul. 55-76, 1955-1 C.B. 239 (Maine Turnpike Authority).

[202] S. SATO & V. ALSTYNE, STATE AND LOCAL GOVERNMENT LAW 640, 710 (1970); Forbes, *State and Local Debt*, in FINANCIAL HANDBOOK §§ 3.16–.18 (I. Altman 5th ed. 1981).

[203] *See* M. MENDELSON & S. ROBBINS, INVESTMENT ANALYSIS AND SECURITIES MARKETS 513–17 (1976); Forbes, *supra* note 202, at §§ 3.23–.27 .

[204] For a discussion of the tax advantages of municipal bonds, see generally M. MENDELSON & S. ROBBINS, *supra* note 203; S. SATO & A. VAN ALSTYNE, *supra* note 202, at 709 (1970).

In addition to the yield on his or her investment, the municipal or state bond investor will also be concerned with the security of that investment. In other words, the bondholder wants some assurance that the principal and interest owed on the obligation by the state or local government will be repaid. Principally, these governments utilize two types of security instruments that guarantee the bond obligation.

Traditionally, the most prominent type of security device offered by state and local governments is the general obligation bond.[205] The state or local government offering its bonds for sale pledges its "full faith and credit" behind such bonds.[206] This pledge signifies to the investor that the governmental unit will do everything in its power to honor its obligation. In essence, the pledge is backed by the issuing government's power to raise taxes in order to pay off the bond. The buyer of the general obligation bond, then, has security in the fact that the tax base of the issuer backs up the debt obligation.[207]

The second type of security instrument utilized by state and local governments is the revenue bond. The security for this type of debt obligation is the issuer's pledge of specified revenues from a public project to pay off the obligation.[208] Toll bridges, airports, roads, and other types of revenue generating public activities are usually financed by revenue bonds.[209] The advantage of such a device is that the state or municipality does not have to pledge its "full faith and credit" as security for the bond.[210] The tax base of the government is thus insulated from securing the debt. Rather, the investor relies solely on the revenues generated by the public project as security for the tax-exempt investment.

By virtue of the section 103(a) exclusion, therefore, state and local governments can enter the national money markets and offer bonds bearing interest rates lower than those that private, non-public borrowers can offer. Clearly, the section 103(a) exclusion on state and local government interest payments accords these governments an important financial benefit in raising

---

[205] *See* Forbes, *supra* note 202, at §§ 3.13–.15.
[206] *Id.*
[207] *See* M. MENDELSON & S. ROBBINS, *supra* note 203, at 528–37.
[208] *See id.* at 537–40.
[209] *Id.*
[210] *See generally* Note, *Small Issue Industrial Development Bonds: The Growing Abuse*, 29 WASH. & LEE L. REV. 223, 223-24 (1982) [hereinafter cited as Note, *Small Issue IDBs*].

money to construct needed infrastructure and public improvements.

Section 103(b) of the Code affords state and local governments the opportunity to issue another type of tax-exempt obligation, known as an Industrial Development Bond (IDB).[211] The proceeds of these tax-exempt issues can be made available in certain defined circumstances to a private business for either quasi-public or purely private ventures.[212] The private business receiving the proceeds from the IDB normally provides security for the bond.[213] A state or local government issuing an IDB can therefore provide a source of capital financing to encourage economic development without pledging its "full faith and credit" as security for the bond. Because of their tax-exempt nature, IDBs have been an important economic development tool for state and local governments that seek to attract and to retain businesses by offering those businesses below market rate financing.[214]

Typically, a government desirous of attracting new private business ventures, or providing aid to existing businesses that wish to expand, will offer to finance private projects for the businesses by issuing tax-exempt IDBs.[215] The government can loan the proceeds of the bond issue directly to the private con-

---

[211] *See* 26 U.S.C.A. § 103(b) (West Supp. 1984).

[212] CONGRESSIONAL BUDGET OFFICE, SMALL ISSUE INDUSTRIAL REVENUE BONDS 18–20 (1981).

[213] *Id.* at 1–2.

[214] *See id.*; Note, *State and Local Industrial Location Incentives—A Well-Stocked Candy Store*, 5 J. CORP. L. 517, 532 (1980) [hereinafter cited as Note, *Industrial Location Incentives*]. Many development economists and experts, however, have questioned the effectiveness of the use of tax incentives, such as industrial development bonds to influence firm location decisions. An excellent survey of this literature is contained in Meadows & Mitrisin, *supra* note 25, at 108–22. As to the question of "why do these subsidy schemes remain so popular with most state and local officials in the face of a large literature that concludes overwhelmingly that they do not work," *id.* at 119, Meadows and Mitrisin suggest that:

> [m]ost state-local public officials may believe the primary contribution of these measures is in the promotion of a "pro-business" image for the state or locality . . . . Indeed, from another perspective it could be argued that state-local officials may feel that they have no choice in the matter; if other neighboring jurisdictions offer fiscal inducements, they will be forced to offer the same benefits in order to remain competitive.

Id. at 119–20.

This argument, aside from debates as to IDB effectiveness as a locational incentive, is a principal reason for supporting IDB authority. Tribes are at a competitive disadvantage in that they presently cannot offer IDB bonds. *See also infra* text accompanying notes 256–58. (recommending that Congress grant tribal governments tax-exempting status).

[215] Note, *Small Issue IDBs*, *supra* note 210, at 223–25.

cern at below market interest rates, or it can use the bond proceeds itself to construct a facility for the private user. In the latter situation, the user then rents the facility from the governmental unit at a below market rental rate, sufficient to retire the total debt obligation owed to the IDB bondholders.[216] Once the obligations of the IDB have been satisfied, the governmental issuer can either sell the facility at a nominal cost to the user or continue to rent the facility.[217]

Although all IDBs were originally tax-exempt,[218] the tremendous growth in IDB volume and perceived abuse of the IDB tax exemption by large corporations led Congress to narrow the IDB tax exemption in 1968.[219] The tax exemption on IDBs was limited to certain specified activities[220] or to "small issues" under specified dollar amounts.[221] Despite these attempts at reform, the volume of IDB issues continued to increase, particularly IDBs issued under the small issue exemption.[222] Critics contended that tax-exempt IDB financing contributed to a significant erosion of the federal tax base.[223] Furthermore, increased IDB volume was cited as contributing to greater competitive pressures on the municipal bond market for funds normally devoted to traditional municipal projects, such as schools and public improvements.[224] In the Tax Equity and Fiscal Responsibility Act of 1982 (TEFRA), Congress responded to these criticisms and substantially modified the small issue exemption,

---

[216] *Id.* at 223–24, n.6. *See* Note, *Industrial Location Incentives*, *supra* note 214 at 532.

[217] *See* Note, *Small Issue IDBs*, *supra* note 210 at 223–25.

[218] I.R.C. § 103(a)(1) (1964) (current version at 26 U.S.C.A. § 103(b)–(n) (West Supp. 1984 & Dec. 1984 Supp.)). *See* CONGRESSIONAL BUDGET OFFICE, *supra* note 212 at 7–12.

[219] Revenue and Expenditure Control Act of 1968, Pub. L. No. 90-364, § 107, 82 Stat. 251, 266–68 (current version at 26 U.S.C.A. § 103 (b) (West Supp. 1984 & Dec. 1984 Supp.)); *see also* Note, *Small Issue IDBs*, *supra* note 210, at 227–29.

[220] *See* I.R.C. § 103(b)(4) (1982). Under the 1968 Act, an IDB issue was considered tax-exempt only if substantially all of the proceeds were used to finance activities such as: projects for residential real property; sports, convention, or trade show facilities; airports, wharves, mass commuting facilities, or parking facilities; training facilities; sewage, waste, air, or water pollution control facilities; water, electric, or gas utilites; and purchase of land for industrial parks. Revenue and Expenditure Control Act of 1968, Pub. L. No. 90-364, § 107, 82 Stat. 251, 266–68 (current version at 26 U.S.C.A. § 103(b) (West Supp. 1984 & Dec. 1984 Supp.)).

[221] *See* 26 U.S.C.A. § 103(b)(6) (West Supp. 1984 & Dec. 1984 Supp.). Small issue bonds that were issued as part of an issue of one million or less were tax-exempt under the 1968 Act. Revenue and Expenditure Control Act of 1968, Pub. L. No. 90-364, § 107, 82 Stat. 251, 266–68 (current version at 26 U.S.C.A. § 103(b) (West Supp. 1984 & Dec. 1984 Supp.)).

[222] *See* Note, *Small Issue IDBs*, *supra* note 210, at 231–32.

[223] *Id.* at 225–26.

[224] *See, e.g.*, *id.* at 225–26.

tightening restrictions on the dollar amounts and types of facilities financed by IDBs.[225] Also, as of December 31, 1988, the small issue exemption will be terminated.[226] Until that time, however, state and local governments can continue to issue tax-exempt industrial development bonds. Officials of these governments can also be expected to intensify their lobbying efforts in Congress to maintain the Code's support for what they regard as a vital economic development tool.

2. Tribal Tax-exempt Bonding Authority Under the Tribal Tax Status Act

Prior to the Tribal Tax Status Act, tribal governments were denied the benefits extended to state and local governments

---

[225] Tax Equality and Fiscal Responsibility Act of 1982, Pub. L. No. 97-248, §§ 214–221, 96 Stat. 324, 466–78 (current version at 26 U.S.C.A. § 103 (1982) (West Supp. 1984 & Dec. 1984 Supp.)) [hereinafter cited as TEFRA]. In TEFRA, Congress significantly amended the IDB provisions of the Code. For example, issuers of IDB bonds must now file quarterly reports to the IRS and provide information respecting users of facilities and descriptions of property financed from bond proceeds. *Id.* § 215(b), I.R.C. § 103(l) (1982). New formal requirements respecting issuance of an IDB have also been instituted, requiring approval by applicable *elected* officials and public hearings before a bond can be issued. *Id.* § 215(a), I.R.C. § 103(k) (1982). One of the most dramatic changes brought about by TEFRA is a provision that mandates that only straight-line depreciation, rather than accelerated depreciation, will be available for most IDB financed property. *Id.* § 216, I.R.C. § 168(f)(12) (1982). In addition, new restrictions have been placed on the small issue exemption portion of I.R.C. § 103(b), the most important being the enumeration of certain types of facilities that may not be built with the proceeds of a small issue IDB. *Id.* § 214(e), I.R.C. § 103(b)(6)(O) (1982). For the most part, these prohibited facilities are recreation and sports related enterprises, such as golf courses, country clubs, massage parlors, racket sport facilities, and racetracks. *Id.* Furthermore, no small issue IDB can be used to finance more than 25% of a retail food or automobile sales or service industry. *Id.* On reactions to the TEFRA proposals see Pierson, *Tax-Exempt Bonds: A Status Report*, 15 TAX NOTES 435 (1982).

[226] Tax Reform Act of 1984, Pub. L. No. 98-369, § 630, *reprinted in* 1984 U.S. CODE CONG. & AD. NEWS (98 Stat.) 494, 933–34, (codified at 26 U.S.C.A. § 103(b)(6)(N) (West Supp. 1984)). One of the most important changes by TEFRA was a provision stating that the small issue exemption for IDB financing was to terminate as of December 31, 1986. TEFRA, § 214(c), 96 Stat. 324, 467 (current version codified as amended at 26 U.S.C.A. § 103(b)(6)(N) (West Supp. 1984)). Pressure from state and local governments, however, led Congress in 1984 to extend the termination deadline for IDBs used to finance manufacturing to December 31, 1988. Tax Reform Act of 1984, § 630, Pub. L. No. 98-369, *reprinted in* 1984 U.S. CODE CONG. & AD. NEWS (98 Stat.) 494, 933–34 (codified at 26 U.S.C.A. § 103(b)(6)(N) (West Supp. 1984)). Additionally, under the Tax Reform Act of 1984, Congress established a limit on the aggregate amount of private activity bonds that could be issued by a state or state agency. *Id.* § 621, *reprinted in* 1984 U.S. CODE CONG. & AD. NEWS (98 Stat.) at 915–18 (codified at 26 U.S.C.A. § 103(n) (West Supp. 1984)), and Congress also moved to deny the tax exemption for obligations issued with a Federal guarantee. *Id.*, §§ 621–48, *reprinted in* 1984 U.S. CODE CONG. & AD. NEWS (98 Stat.) at 915–41 (codified at 26 U.S.C.A. § 103 (West Supp. 1984)). Other minor restrictions were also imposed. *See* H.R. REP. No. 861, 98th Cong., 2d Sess. 1199–1218 (1984), *reprinted in* 1984 U.S. CODE CONG. & AD. NEWS 751, 1193–1212.

under sections 103(a) and 103(b) of the Internal Revenue Code.[227] In fact, in order to raise capital to finance public improvements or job creating economic development initiatives, Indian governments would have had to offer bonds at interest rates competitive with private borrowers, such as General Motors or Chrysler—a debt burden which few if any tribes could afford. Under the Tribal Tax Status Act, Congress finally granted tribal governments the authority to issue tax-exempt debt obligations.

At least with respect to this one highly controversial area of the Code (IDBs), however, Congress substantially departed from its stated goal of providing Indian tribes a tax status similar to that granted states and local governments.[228] The Tribal Tax Status Act authorizes Indian tribal governments to issue tax-exempt bonds "only if such obligation is part of an issue substantially all of the proceeds of which are to be used in the exercise of any essential governmental function."[229] Furthermore, the Act specifically denies a tax exemption on interest paid for any tribally issued industrial development bond,[230] thus effectively denying Indian tribes the ability to utilize the IDB financing device.[231]

Under the Tribal Tax Status Act, therefore, the debt obligations of Indian tribal governments are treated similarly to those of state and local governments, subject to the essential governmental function requirement. The conference report accompanying the final version of the Tribal Tax Status Act defines essential governmental functions as including projects such as "schools, streets, and sewers."[232] That same report states that

---

[227] *See supra* notes 115–16 and accompanying text.

[228] *See* S. REP. No. 646, *supra* note 143, at 11.

[229] Indian Governmental Tribal Tax Status Act of 1982, § 202(a), I.R.C. § 7871(c)(1) (1982).

[230] *Id.*, I.R.C. § 7871(c)(2).

[231] H.R. REP. NO. 984, *supra* note 146, at 17.

[232] *Id.* Treasury's regulation accompanying the Tribal Tax Status Act sets out three tests to determine whether a tribal activity qualifies as an "essential governmental function" and is therefore eligible for tax-exempt financing treatment. *See* Treas. Reg. § 305.7871-1(d) (temporary regulation) (1984). The first two tests refer specifically to congressionally created funding programs designed exclusively for tribal governments. Under the regulations, if a tribal function qualifies for funding, grants, or contracts under 25 U.S.C. §§ 13, 450f–450h (1982), then that function is regarded as an essential governmental function for purposes of the Tribal Tax Status Act. A wide array of activities, including health, education, and welfare activities, as well as basic infrastructure and community development, are eligible for funding under those sections. *See* Treas. Reg. § 305.7871-1(d)(1)–(2) (temporary regulation) (1984).

The third test for an essential governmental function under the Regulation refers

tribal governments are not permitted under the Tribal Tax Status Act to issue tax-exempt private activity bonds, such as IDBs.[233] As a result, Indian tribal governments now have the legal authority to enter the money markets and to offer bonds bearing interest rates lower than those that private borrowers can offer, only so long as the proceeds are used to finance a limited set of projects. This decision to provide tribes such a narrow tax exemption for their debt obligations raises the important question of whether Congress, in claiming to supplement the self-governing powers of Indian tribes, has in reality provided the tribes with additional powers that exist only in theory.[234]

IV. Implications of the Limited Exemption for Tribal Debt Obligations

A federal Indian policy initiative that pledges tribal self-sufficiency as its goal must first recognize the barriers contributing to Indian Country underdevelopment. Though no single initiative can substantively address each of these major impediments, any potentially successful initiative must at least recognize and work within the constraints imposed by the economic and social realities present in Indian Country. Meaningful economic and social development cannot occur in Indian Country absent ef-

---

expressly to § 115 of the Internal Revenue Code as an additional test to determine whether a tribal function should be accorded tax-exempt status. Treas. Reg. § 305.7871-1(d)(3) (temporary regulation) (1984). Section 115 of the Code excludes from gross income the income derived from a public utility or "the exercise of any essential governmental function and accruing to a state." I.R.C. § 115(1) (1982). Thus, if a function of a state would be regarded as an essential governmental function under § 115 of the Code, then a tribal government performing that same function would be regarded as performing essential governmental functions eligible for tax-exempt financing status. *See* Treas. Reg. § 305.7871-1(d)(3) (temporary regulation) (1984).

Legislative history and the provisions of the Tribal Tax Status Act, however, seem to indicate a clear congressional intent to limit tribal tax-exempt financing authority to areas traditionally regarded as public functions. *See, e.g.*, H. Rep. No. 984, *supra* note 146, at 16–17. "Indian tribal governments are to issue public activity bonds, the proceeds of which are used in an essential governmental function (such as schools, streets, and sewers). Therefore, tribal governments are not permitted to issue private activity bonds (i.e., industrial development bonds . . .)." *Id.* Given the strong expression of congressional intent, courts and the IRS can be expected to monitor closely any tribal bond issue used to finance activities other than school construction, transportation, or other projects traditionally regarded as public works. Tribes can also expect legal challenges to their authority to issue any type of tribal tax-exempt revenue bond whose proceeds are utilized either directly or indirectly to finance activities other than traditional government revenue producing projects, such as roads, airports, or water and sanitation facilities.

[233] H.R. Rep. No. 984, *supra* note 146, at 17.

[234] *See supra* notes 136–42 and accompanying text.

fective mitigation of the obstacles to growth that confront Indian Nations.

Congress's decision in the Tribal Tax Status Act to provide only a limited tax exemption for tribal debt obligations seriously undermines this initiative's capacity to strengthen tribal self-governing powers. Denying tribes the authority to issue tax-exempt IDBs effectively denies tribal governments the ability to raise the large amounts of capital needed to finance private and public/private sector development initiatives. These types of development projects are desperately needed in Indian Country to create and to sustain the economic and social environment necessary for meaningful job creation and lasting economic growth. In denying tribes the ability to engage in creative tax-exempt IDB financing arrangements, Congress has failed to provide tribal governments with the tools to address the structural barriers impeding tribal self-sufficiency, thereby undermining the other initiatives contained within the Tribal Tax Status Act.

Furthermore, broader notions of equitable tax policy[235] and Congress's treaty based trust responsibility to Indian tribes[236] argue strongly for amendment of the Tribal Tax Status Act to provide tribes full tax-exempt bonding authority. The possible contours of such revisions are suggested in the final section of this Article.

One of the most important initiatives in the Tribal Tax Status Act in terms of providing tribal governments with the ability to address the structural barriers impeding economic and social development is the provision permitting full federal income tax deductions for tribal taxes.[237] In theory, the Tribal Tax Status Act therefore provides tribal governments far greater flexibility

---

[235] *See supra* note 149.

[236] *See supra* note 4.

[237] *See supra* notes 162–67 and accompanying text. For tribal members, however, many of whom will not earn a sufficient income to justify itemizing deductions, this new benefit extended to their tribal governments may have a negative impact on their own financial picture. If tribes now move to impose new taxes because of the Tribal Tax Status Act, the Act ironically will decrease the take-home income of tribal members who work in Indian Country. For those who do not itemize deductions on their federal income tax forms, their take-home income will be reduced by the full amount of the new tribally imposed tax. Perhaps a federal tax credit for tribally imposed taxes would be a more equitable device for these types of taxpayers. Given the difficulty that tribes already have in attracting private sector businesses to locate in their jurisdictions, a tribal tax, even if deductible for federal income tax purposes, may represent another entry barrier. Again, a more fair and effective method from a development incentive perspective would be to treat tribally imposed taxes as eligible for a tax credit toward federal income taxes.

in devising broadly based revenue raising tax initiatives. It is unlikely, however, that many Indian governments will ever be able to benefit fully from this increased flexibility provided under the Tribal Tax Status Act.

The ability to devise broadly based tax measures presumes sources of economic activity upon which to impose a tax. States and their subdivisions, which have always enjoyed full federal income tax-deductibility of their levies, can usually draw upon a diverse and stable tax base in devising taxing measures to raise revenues, such as employment and wage taxes, corporation and business taxes, real and personal property taxes, and sales taxes. Other than the relatively small number of Indian Nations aggressively developing their mineral and energy reserves, few Indian communities enjoy the thriving economic environment necessary to sustain a stable tax base. This lack of economic activity, of course, is directly attributable to the interrelated structural barriers to development previously noted.[238]

The Tribal Tax Status Act does not provide a mechanism that would enable tribes to create a thriving economic environment within Indian Country. Instead, it offers tribes only the theoretical ability to exercise broadly based taxing authority over a nonexistent tax base. Absent substantial, long-term growth in tribal economies, the majority of Indian Nations will find themselves in the same predicament as before the Act in terms of their tax base and revenue raising capacity.

The second major development initiative contained within the Tribal Tax Status Act affords tribal governments the ability to offer tax-exempt bonds, permitting them to raise funds at interest rates lower than those that private borrowers must pay in the money markets. Tribes can issue these bonds, however, only so long as the proceeds of a tribal debt obligation "are to be used in the exercise of any essential governmental function," such as schools, roads and sewers.[239] This limited authority to issue tax-exempt public activity bonds for essential public infrastructure needs provides tribes, in theory at least, with an important governmental power to reduce barriers to development and self-sufficiency. As discussed above, an inadequate public infrastructure base significantly impedes the ability of tribes to

---

[238] *See supra* note 3.
[239] *See supra* notes 229–33 and accompanying text.

attract and promote business investment.[240] But once again, Congress has undermined a power vested in tribes under the Tribal Tax Status Act by failing to recognize that the structural underdevelopment that plagues Indian Country will frustrate the exercise of the tribe's public activity tax-exempt financing power.

As with the limits on a tribe's ability to devise broadly based revenue raising tax initiatives, tribal bonding capacity is functionally limited by the reservation's tax base. A stable, diverse tax base is essential to assure potential investors that their investment is secured by sufficient revenue raising capacity.[241] Currently, few tribes can offer investors the security of a stable tax base, sustained by a healthy economic and social environment. Absent such security, investors in capital markets will show little interest in tribal bonds that are intended to finance the public infrastructure needed to facilitate economic development.[242]

Congress, in failing adequately to address these structural impediments, has provided tribes important governmental powers in theory but not in fact. While purporting to confer upon tribes a status equal to that of states and local governments, Congress has ignored the fundamental factual distinctions between these governments and Indian governments. State and local governments have highly diversified, dynamic economies, providing stable tax bases upon which to enact revenue raising measures and to borrow for capital investment. Indian Nations, because of their remoteness from major markets, poor infrastructure environments, land ownership patterns, and other structural deficits, lack such diverse and dynamic economies and accompanying tax bases.

Ironically, the IDB tax-exempt financing authority, which Congress denied Indian tribes in the Tribal Tax Status Act, could have provided an important vehicle for Indian Nations to stimulate and to develop new economic activity and the tax base within their borders. Specifically, IDB authority could have assisted tribes in acquiring previously unavailable development capital for private and joint venture investment initiatives on trust or allotment lands, thus assisting tribal efforts to overcome

---

[240] *See supra* notes 28–32 and accompanying text.
[241] *See supra* notes 204–07 and accompanying text.
[242] *See supra* notes 205–07 and accompanying text.

this obstacle to self-sufficiency.[243] In those few instances where tribes have been able to utilize the IDB financing device through agreements with nearby local governments, they have been very successful in generating jobs and capital investment in Indian Country.[244]

Because Industrial Development Bonds are backed by the credit of the business or industry receiving the proceeds of the issue,[245] IDB financing bypasses many of the impediments to development caused by Indian Country land ownership patterns. Specifically, IDB financing backed by the credit of the benefited enterprise expands development opportunities in tribally owned trust or allotment lands, which generally cannot be pledged as security for a loan or mortgage.

As an example of the flexibility of this economic development tool,[246] a tribe could recruit a manufacturing concern in exchange for below market rental rates on a tribally owned and IDB financed building site. So long as a tribe could induce a credit worthy business to enter its borders, the trust or fractionated status of the tribal land base would no longer have to represent a nearly insurmountable impediment to development and investment. To encourage indigenous business growth, a tribe could translate its low interest financing advantage into a

---

[243] *See supra* notes 33–58 and accompanying text. Recently, the United South and Eastern Tribes (USET), an organization representing many of the recognized tribal governments in the eastern half of the United States, reached an agreement with the city of Nashville, Tennessee, whereby the city agreed to issue an IDB on behalf of USET to finance a major part of a $5.2 million development project. The project will feature a $2.8 million, 30,000 square foot commercial office complex and a 10 acre tourist center complex that will generate annual estimated revenues of $300,000. *See* Williams, *Capital Investment: USET Seeks Private Partners for an Indian Cultural Center in the Heart of Nashville*, INDIAN TRUTH, Apr. 1984, at 9.

[244] For instance, the Mississippi Band of Choctaw Indians has financed a successful industrial park within its borders through an Industrial Development Bond issued by the city of Philadelphia, Mississippi. Earnest Tiger, Economic Development Specialist for the Choctaws, explained the progress of this project:

> The project at this point has 219 employees. It has been open since August 17. By midsummer we plan to have 350 people working. It has a 120,000-square-foot facility, 21 acres leased to the city of Philadelphia, and many tribes see this as a barrier because they have to lease the land which means the possibility of waiving sovereign immunity to a municipality or a county, which puts them back into a relationship with the State, and in many areas that has some serious drawbacks and consequences. But it is proceeding and we have created jobs and we are making money.

Oversight of Economic Development of Indian Reservations, *supra* note 3, at 54 (statement of Earnest Tiger, Economic Development Specialist, Miss. Band of Choctaw Indians).

[245] *See supra* notes 211–14 and accompanying text.

[246] *See generally* Note, *Industrial Location Incentives*, *supra* note 214 (comprehensive catalog of the various IDB incentives offered by state and local governments).

low interest capital loan program for member owned Indian enterprises. Even though these new Indian owned businesses might lack credit records sufficient to secure adequately a tribally issued IDB, the tribe could provide additional guarantees to potential IDB investors by pledging, for instance, a portion of its grazing or timber fees, or rents and royalties from present and future natural resource development projects.[247] Risk-pooling, royalty sharing, and other types of creative financing arrangements can also be explored if a tribe wishes to adopt a development policy of encouraging new member owned businesses in Indian Country.[248] Without the authority to engage in creative tax-exempt financing activities, however, tribes cannot even explore these possible development options.

Absent the authority to engage in creative tax-exempt IDB financing, tribes stand at a competitive disadvantage compared to state and local governments, which have numerous IDB locational incentive programs. In its present form, the Tribal Tax Status Act fails to recognize this disadvantage suffered by tribal governments. The Act will not provide the tools needed by tribes in overcoming the barriers to development posed by Indian Country land ownership patterns and inadequate capital access. Furthermore, denying tribal governments IDB authority hinders tribes in promoting indigenous economic growth,[249] which could create entry level and management jobs, all too scarce in Indian Country.[250] Member owned small businesses, created and supported by tribal tax-exempt IDB financing, could serve as a breeding ground for the development of Indian entrepreneurial and technical expertise.

Without the ability to promote development initiatives that IDB financing authority has provided to many states and local governments, the typical tribal government will continue to lack the self-governing capacity to provide for the needs of its people and to pursue goals of social and economic self-sufficiency. That Congress ignored the expressed desires of the Indian tribes

---

[247] *See supra* note 53 and accompanying text.

[248] Risk-pooling spreads the risks of a number of investments among a group of investors. A tribe could issue a tax-exempt IDB bond, the proceeds of which are secured by a number of business ventures, thereby reducing the risk of total default on a bond. Royalty financing provides the investor with a royalty on items sold that were produced through the investment. *See* Williams, *State and Local Development Incentives for Successful Enterprise Zone Initiatives*, 14 RUTGERS L.J. 41, 88 n.137.

[249] *But see supra* note 214 (economists' questioning of the effectiveness of IDBs in influencing firm location decisions).

[250] *See supra* notes 59–68 and accompanying text.

which anxiously desired IDB authority,[251] indicates that non-Indian policymakers are still insensitive to Indian preferences—an attitude that has characterized the long and dismal history of federal Indian development policy.[252]

In short, the Tribal Tax Status Act does indeed represent an important, but largely symbolic, milestone for Indian tribes in their continuing struggle for economic and social self-sufficiency. Under the Act, Indian tribes now receive many of the same favorable tax advantages afforded states and their subdivisions. By denying tribes the ability to utilize the same tax-exempt fiscal development tools available to other governments to stimulate growth and to sustain a tax base, however, Congress has extended merely theoretical benefits. Furthermore, in its decision to deny Indian Nations the ability to engage in creative tax-exempt financing, Congress has failed to recognize a unique opportunity to assist tribal governments in addressing the structural barriers to self-sufficiency. In ignoring the realities of Indian Country underdevelopment, realities that Indians themselves attempted to express in their testimony supporting the Tribal Tax Status Act,[253] Congress has enacted an Indian development initiative which has little hope of having a significant impact on the social and economic environment of many Indian Nations. As presently constituted, the Tribal Tax Status Act represents but a few, nearly meaningless steps for Indian Nations on their long road to self-sufficiency.

V. Recommendations

A. *Suggested Amendments to the Tribal Tax Status Act*

Now that Congress has recognized that tribal governments today provide the same services and functions for their constituents as other governmental units,[254] "the fundamental tax policy interest in promoting horizontal equity [between similarly situated taxpayers]"[255] argues strongly in favor of Congress

---

[251] *See supra* note 150.
[252] *See supra* notes 69–89 and accompanying text.
[253] *See supra* note 150.
[254] *See supra* notes 124–30 and accompanying text (statements of congressional intent in enacting the Tribal Tax Status Act).
[255] *1981-82 Miscellaneous Tax Bills Hearing, supra* note 136, at 57 (statement of William McKee).

amending the Tribal Tax Status Act. Tribal governments should be provided the same authority to raise revenues and to promote development that are provided to those other governmental units in the Internal Revenue Code.

Congress, therefore, should grant tribal governments the authority to issue tax-exempt Industrial Development Bonds as proposed in all earlier versions of the Tribal Tax Status Act.[256] Congress must recognize that in denying tribes the ability to provide private sector financing incentives at least equal to that of state and local governments, it has created an almost insurmountable competitive disadvantage for Indian Nations in their promotion of economic development.

As a second step, Congress must recognize that the federal government's trust responsibility to Indian Nations should include repairing the damage inflicted upon Indian Country by the past discriminatory tax treatment of tribal governments. State and local governments have enjoyed a favorable tax status permitting them to engage in tax-exempt economic development financing for nearly five decades. Indian Nations, meanwhile, have endured Third World conditions of poverty and underdevelopment, created and sustained in large part by federal neglect.[257] Therefore, Congress should amend the Tribal Tax Status Act to provide tribes with IDB financing authority beyond the December 1988 termination date set for all state and local government IDB tax-exempt financing issues.[258]

Such an addition to the Tribal Tax Status Act is suggested as an important remedial measure. Extending to tribal governments a period of time during which they *alone* can offer significant tax-exempt financing incentives to the private sector would enable Indian Nations to make up much of the ground lost to state and local governmental units over the past half century. To date, Indian Nations have been hindered from using private-activity tax-exempt financing and thus have been denied the ability to compete for development opportunities on an equal footing with those governments.

---

[256] *See supra* notes 123–34 and accompanying text (describing original 1976 version of the Tribal Tax Status Act that would have permitted Indian tribes to issue tax-exempt IDBs, so long as the proceeds of the bond issue were not "used in any trade or business carried on outside" the Indian Nation's territorial jurisdiction, H.R. 8989, § 5, *supra* note 123).

[257] *See supra* notes 3–5 and accompanying text.

[258] *See supra* note 226 and accompanying text.

These two important measures—providing tribes IDB financing authority equal to that of states and local governments and extending that authority beyond the terminus date that Congress has set for all state and local tax-exempt IDB issues—would enable tribes to take meaningful steps towards economic and social self-sufficiency. Standing alone, however, even these important additions to the Tribal Tax Status Act will not provide all the development and investment capital necessary for creating self-sustaining economic and social environments within Indian Country.

First, even with substantial IDB financed investment subsidies, private entrepreneurs will still be concerned about the other barriers impeding Indian Country development, such as remoteness from markets and labor force deficiencies. In Indian Country, at least, the tax-exempt IDB subsidy will be perhaps one of the few significant incentives that tribes can rely on to induce private sector investment. As tribal leaders and development economists have recounted many times, however, private sector capitalists and their corporations are extremely risk averse and conscious of the "bottom-line".[259] For many entrepreneurs, the benefits of an IDB's three to four percentage point interest subsidy on capital costs may well be outweighed by the perceived risks of locating a business enterprise in Indian Country.

Second, capital market inefficiencies and imperfections, such as high transaction and information gathering costs, often represent significant entry barriers to national money markets for relatively small governmental borrowing entities, such as Indian tribes.[260] Skepticism on the part of the non-Indian investment community concerning Indian Country investment risks and concern over the consequences of a tribal IDB financed business venture default will create further barriers limiting a tribal government's ability to utilize the IDB financing device effectively.[261] Even if a tribe can entice a business to locate within

---

[259] These types of capital access barriers have also plagued minority businesses, small and medium-sized firms, and other enterprises located in underdeveloped and depressed segments of the national economy. On capital market inefficiencies and small business capital access problems, see *Keeping Business in the City: Hearings Before the Joint Economic Comm., Subcomm. on Fiscal and Governmental Policy*, 95th Cong., 2d Sess., 155 (1978); Garvin, *The Small Business Capital Gap: The Special Case of Minority Enterprise* 26 J. FIN. 445 (1971).

[260] Williams, *Reagan's Initiatives Lead to More Questions Than Answers*, INDIAN TRUTH, Apr. 1983, at 17.

[261] *See supra* notes 227–34 and accompanying text.

its borders by the use of IDB financing incentives, it still must overcome these types of entry barriers in order to sell its bonds in the national money markets.

B. *Creation of an American Indian Financial Institution*

One meaningful form of assistance that would greatly increase the ability of tribes to take full advantage of an amendment providing them with the authority to issue IDBs, as well as to address other significant barriers impeding economic and social development, would be the creation by Congress of an American Indian financial intermediary institution. A principal function of a financial intermediary institution for Indian Country would be to promote social and economic development for Indian Nations by providing or facilitating the provision of long-term investment capital and capital credit for the development of tribal economies. Such an institution could serve as a principal purchaser of tribal tax-exempt public and private activity bonds, providing tribes with a ready market for their debt instruments.

The creation of a financial intermediary institution for Indian Country has been advocated by Indian leaders and spokespeople for nearly a decade. In 1976, a task force commissioned by the American Indian Policy Review Commission recommended the creation of an American Indian Development Authority that would provide the "means for the efficient development of viable reservation economies, which will afford Indian people the maximum opportunity for choice of both their style and standard of living, and which will move them towards self-sufficiency as individuals and as tribes."[262] Subsequent proposals for such an institution have adopted the general thrust, if not the details, of the Task Force's original recommendations advocating a federally created intermediary institution with the capacity to provide development and investment capital to Indian Nations.[263]

Numerous precedents exist for the creation of a federal financial intermediary institution designed to correct structural capital allocation and distribution deficiencies in the national economy. The most famous precedent, of course, is the Re-

[262] TASK FORCE SEVEN, *supra* note 3, at 132.

[263] *See, e.g., Oversight of Economic Development on Indian Reservations, supra* note 3, at 8–20 (panel discussion) (proposing the creation of an American Indian Development Finance Corporation).

construction Finance Corporation (RFC), which provided critically needed loans and capital to banks and other financial intermediaries, railroads, commercial and industrial enterprises, agriculture, public agencies, defense contractors, and other segments of the national and even international economy during the Depression.[264] Other well known contemporary examples include the Federal National Mortgage Association (FNMA), the Government National Mortgage Association (GNMA), the Federal Home Loan Bank System, the Federal Home Loan Mortgage Corporation, and the Farmers Home Administration. These federally created institutions play critical, intermediary roles in vital segments of the national economy.[265] An American Indian financial intermediary institution could be modeled after any of these federally created agencies. The activities and organization of the international development institutions of the World Bank group also suggest possible models upon which to structure an American Indian financial intermediary vehicle.[266]

Regardless of the final structure adopted for such a vehicle, an American Indian financial intermediary should be created by Congress as a federally chartered, primarily Indian owned and operated, corporate institution. Its structure and federally granted privileges should be similar to that of other federally created financial intermediaries, such as the Federal National Mortgage Association.[267] The American Indian financial intermediary's charter should clearly indicate its legislatively mandated, nonpolitical purpose of providing and facilitating capital access opportunities for Indian Nations, either through loans, loan guarantees, credits, grants, or other appropriate devices. The charter should also include authority for the provision of technical and managerial assistance to tribal clients, promotional efforts designed to inform the private sector of Indian Country

---

[264] *See* Bickley, *On Evaluation of the Reconstruction Finance Corporation with Implications for Current Capital Needs of the Steel Industry*, in NEW TOOLS FOR ECONOMIC DEVELOPMENT, *supra* note 25, at 144–59.

[265] *See generally*, FED. NAT'L MORTGAGE ASS'N, A GUIDE TO FANNIE MAE (1979) (describing the general structure and role of each of these intermediary institutions).

[266] *See, e.g.*, *Oversight of Economic Development on Indian Reservations*, *supra* note 3, at 20 (statement of Joe Baca, National Congress of American Indians Economic Development Committee member) (discussing the various activities and organization of the World Bank group of institutions, including the International Finance Corporation, as models for appropriate institutional structures and powers of an American Indian Development Corporation).

[267] *See* FED. NAT'L MORTGAGE ASS'N, *supra* note 265, at 2–3 (describing FNMA as privately owned and controlled by its stockholders and directed by its board of directors as an independent financial corporation).

development opportunities, and the identification and coordination of private sector investment opportunities and applicable government financial and technical assistance programs.

An American Indian financial intermediary institution should have the authority and the commitment to serve as a financier, investor, lender, loan coordinator, syndicator, guarantor, and broker for Indian Country economic and social development. It should support and facilitate the widest variety of tribally desired private and joint venture development activities, as well as provide or secure financing of traditional public capital infrastructure. To this end, the institution should have the authority to purchase tribal tax-exempt and taxable debt obligations. This type of authority would provide tribes with a ready and accessible market for their debt instruments. Furthermore, such an institution could achieve economies of scale in terms of investigating the credit worthiness of an issuing tribe or a business seeking tribal IDB financing assistance; it could work with the tribe in structuring an issue that the institution could purchase while adhering to sound investment principles. Such an authority to purchase tribal bonds would significantly minimize the tribes' access barriers to the municipal bond market.[268]

Another method to assure liquidity and to broaden the potential tribal tax-exempt investment base would be to provide the institution with the related authority to pool these municipal bonds and to issue interest bearing securities or bonds to investors upon this pool, backed by the assets represented by the tribal bonds. Such securities or bonds could be modeled upon the modified pass-through type of instruments issued by the Government National Mortgage Association (GNMA). GNMA is a financial intermediary institution within the Department of Housing and Urban Development that supplies mortgage credit in support of federal housing objectives.[269] Since 1970, GNMA has administered a program under which it guarantees payment of principal and interest on securities issued to investors by holders of pools of federally guaranteed or insured mortgages constituting the assets backing the securities. The GNMA guarantee is backed by the full faith and credit of the federal government, and therefore GNMA's securities are regarded as virtually risk free investments by the bond market. Under the modified pass-through type of instrument issued by GNMA,

---

[268] *See supra* note 259 and accompanying text.
[269] *See* 12 U.S.C.A. § 1717 (West Supp. 1984).

each investor in the securities receives a proportionate share of principal and interest due from the pooled mortgages, whether or not collected from the mortgagor. The issuer of the GNMA guaranteed security, usually a mortgage banking company, makes cash advances out of its own funds to cover delinquencies until reimbursed by the mortgagors, or by GNMA as guarantor.[270]

An American Indian financial intermediary vehicle would perform a similar function as GNMA with respect to tribal debt obligations, pooling a large number of tribal bonds and issuing federally guaranteed securities in the money markets; the bonds would represent the assets backing the securities. The principal and interest on these securities should be guaranteed by the federal government in a manner similar to GNMA's securities. This type of authority to pool tribal bonds and to issue securities or bonds backed by these assets, together with a federal guarantee, would permit the institution to attract additional secondary sources of investment capital for tribal bonds and would also permit the institution to expand its own primary purchasing activities.

While serious consideration should be given to issues such as the institution's organization and management, its development and investment philosophy, its initial capitalization, and the degree and scope of federal financial support and federally authorized powers and privileges, the creation of an American Indian financial intermediary vehicle as briefly described here provides an important example of a federal initiative that *would* represent a significant step toward self-sufficiency for Indian nations. Such an initiative would enable tribes to take full advantage of such legislation as the Tribal Tax Status Act and would provide effective means for assisting tribes in addressing the structural barriers to economic and social development.

In short, though the recommendations suggested here deserve serious consideration by Congress, even if enacted as proposed, these remedial measures would still advance Indian Nations only a few small steps on the long road to true self-determination. The success of tribal efforts to promote economic and social self-sufficiency through IDB financing and an American Indian financial intermediary will ultimately depend upon the perceptions of both private sector entrepreneurs and potential

---

[270] *See* FED. NAT'L MORTGAGE ASS'N, *supra* note 265, at 81–83 (description of GNMA and its various functions and securities offerings).

tribal IDB bond investors respecting the risks and barriers associated with Indian Country development investment. Those perceptions will be influenced principally by the federal government's continued commitment to provide tribes with other meaningful forms of assistance in mitigating those risks and barriers.

VI. Conclusion

This government "has charged itself with moral obligations of the highest responsibility and trust" in its dealings with American Indian Nations.[271] It should, therefore, do much more than merely achieve "horizontal equity" for tribes respecting their tax status under the Internal Revenue Code. Sadly, the present version of the Tribal Tax Status Act fails to achieve even this limited goal.

The proposals suggested in this Article are urged as *minimum requirements* of this nation's trust responsibility to assist Indian people in their efforts towards self-sufficiency.[272] Congress should act quickly to cure the defects in the present Tribal Tax Status Act, because the barriers to tribal economic and social self-sufficiency have been created and sustained in large part by the federal government's own failed Indian Country development policies. Legislation granting tribes a federal tax status truly equal to state and local governments, including tax-exempt IDB authority, would indeed represent an important step on the long road to self-sufficiency for Indian Nations. Yet even this step, as small as it may seem, would represent the acceptance of a critical reconstructive challenge for both white and red America.

First, the actions urged here require a rethinking of past approaches to the formulation of federal Indian policy. Focusing

---

[271] Seminole Nation v. United States, 316 U.S. 286, 296–97 (1942).

[272] For instance, housing, education, health care, and a host of other major Indian Country problems associated with decades of chronic underdevelopment will continue to confront federal Indian policymakers long after defects in the Tribal Tax Status Act have been corrected by Congress. *See, e.g.*, NTCA Statement, *supra* note 3 (position papers E, F, and G) (detailing the numerous critical problems in Indian Country housing, education and health care). Thus, the recommendations advanced here do not pretend to represent any ultimate panacea for the poverty and misery inflicted upon Indian Nations by the historical failures of this country's past Indian policies. As the title of this Article indicates, the measures urged here are admitted to be but small steps that could provide important assistance to Indian Nations on their long road toward true self-determination.

on the causes, rather than the symptoms, of underdevelopment, these recommendations ask white America, historically wed to its own collective myths in dictating past Indian policies,[273] to empower Indian Nations with the means to begin constructing their own vision of economic and social self-sufficiency. Thus, for white America, the challenge of the reconstructive project urged here is the first step towards emancipation from the desire for power—which works its own dehumanizing mastery over those whom it possesses. Their choice in accepting this challenge is therefore ultimately a personal and moral one.

For red America, surrounded by an acquisitive modern society that regards tribalism as an inconvenient anachronism,[274] the challenge represented by deciding to participate in such a transformative project transcends issues of mere personal or moral choice. Working with white America in creating policies and programs that can furnish the capacity for self-determination represents a far greater challenge, one directed at the task of assuring tribalism's continued existence amidst the individualistic technocracies of the West. That challenge will require red America to work continually within the structures of an alien consciousness that history and fate have irredeemably imposed on it. Together with white policymakers, Indian Nations must endeavor to discover and to develop those liberative root ideals claimed to exist at the core of Western ideology that might permit the empowerment and liberation of indigenous tribal cultures.

Such discoveries, of course, consistent with the most ancient traditions of Indian gift giving,[275] enrich both the giver and receiver and create an eternal covenant chain[276] among people as well as nations. This covenant, formed to link and thereby to transcend two radically divergent world views, perhaps might then free each race to attain together a far more noble and harmonious vision of our shared humanity and human potential.

---

[273] *See supra* notes 69–89 and accompanying text.

[274] *See generally* U.S. COMM'N ON CIVIL RIGHTS, INDIAN TRIBES: A CONTINUING QUEST FOR SURVIVAL 7–13 (1981). The Commision describes the various "backlash" events and groups organized by non-Indians in the 1970's in response to assertions of tribal powers, and also claims of "special rights" and "preferential treatment" of Indians. *Id.* at 10–11.

[275] *See, e.g.*, K. LLEWELLYN & E. HOEBEL, THE CHEYENNE WAY 169–269 (1941) (describing Cheyenne gift giving customs).

[276] *See* C. COLDEN, HISTORY OF THE FIVE INDIAN NATIONS 90 (1902); VIII DOCUMENTS RELATIVE TO THE COLONIAL HISTORY OF THE STATE OF NEW YORK 619–24 (E. O'Callaghon & E. Baily ed. Albany, 1853–87).

ARTICLE

THE PROPOSED UNIFORM NEW PAYMENTS CODE: ALLOCATION OF LOSSES RESULTING FROM FORGED DRAWERS' SIGNATURES

STEVEN B. DOW*
NAN S. ELLIS**

The 3-4-8 Committee, appointed by the Permanent Editorial Board of the Uniform Commercial Code, has completed a third draft of the proposed Uniform New Payments Code. The new code substantially revises the treatment of funds transfers currently governed by articles 3 and 4 of the Uniform Commercial Code and contains two major changes in the allocation of losses resulting from forged drawers' signatures. First, customer liability for unauthorized items has been revised, and several new strict liability provisions have been included. Second, the long-standing rule of Price v. Neal *has been abolished and replaced with a scheme of transmission liability.*

In this Article, Professors Dow and Ellis review the current system of allocating losses resulting from forged drawers' signatures under articles 3 and 4 of the Uniform Commercial Code and identify changes made by the current draft of the New Payments Code. The authors approve of the New Payments Code's proposed two-tiered approach to the allocation of losses resulting from forgeries. They argue, however, that business customers should receive the same protections as consumer customers under the New Payments Code when forged checks involving small sums are handled. The authors also argue that the abolition of the rule of Price v. Neal *in favor of a system of transmission liability is unjustified and should be rejected. While reviewing the two Codes, Professors Dow and Ellis present alternatives for revision of the present system of payments law.*

The allocation of loss in cases of forgeries and material alterations represents one of the most problematic aspects of the

---

* Assistant Professor of Business Law, Michigan State University. B.L.S., Bowling Green State University, 1973; J.D., Ohio State University, 1978.

** Assistant Professor of Business Law, Loyola College, Baltimore, Md. B.A., Ohio State University, 1974; J.D., Ohio State University, 1977. Professor Ellis is a member of the American Bar Association Ad Hoc Committee on the New Payments Code. The views expressed herein are those of the authors and do not necessarily reflect any view or position taken by the committee.

study of commercial paper. Annual losses from forged checks[1] alone have been estimated at between $60 million and $1 billion annually.[2] Currently, the allocation of such loss is governed primarily by articles 3 and 4 of the Uniform Commercial Code (U.C.C.),[3] supplemented by the common law, especially the law of mistake.[4]

The proposed Uniform New Payments Code (N.P.C.) presents a substantial revision of articles 3 and 4 of the U.C.C.[5] One goal of the N.P.C. is to create a legal framework that will govern

---

[1] In this context, the term "forged checks" includes checks upon which there is a forged drawer's signature and also checks upon which there is a forged indorsement. Typically, however, the term refers only to a check with a forged drawer's signature.

[2] Hudak & MacPherson, *Forged, Altered or Fraudulently Obtained Checks*, 23 PRAC. LAW. 73, 73, Apr. 15, 1977; O'Malley, *The Code and Double Forgeries*, 19 SYRACUSE L. REV. 36, 36 (1967); Comment, *Commercial Paper and Forgery: Broader Liability for Banks?*, 1980 U. ILL. L.F. 813, 813.

[3] U.C.C. arts. 3–4 (1978). For a table listing those states that have adopted the U.C.C. and the respective state codifications, see 1 U.L.A. 1 (Supp. 1985).

[4] *See* U.C.C. § 1-103 (1978); *see also infra* notes 41, 121–23 and accompanying text.

[5] The Permanent Editorial Board (P.E.B.) of the U.C.C. was established by the National Conference of Commissioners on Uniform State Laws (N.C.C.U.S.L.) and the American Law Institute (A.L.I.), joint sponsors of the U.C.C., to make proposals for revisions and amendments to the U.C.C. as needed over time. The 3-4-8 Committee was appointed by the P.E.B. to consider amendments to articles 3, 4, and 8 of the Code. The Committee's amendments to article 8 were approved in a separate draft in 1977.

In 1978 the P.E.B. charged the 3-4-8 Committee with drafting a comprehensive payments code to establish a legal framework for all types of payments other than cash. The Committee has been working on a number of drafts since 1978. The Uniform New Payments Code (N.P.C.), if adopted, would govern both electronic fund transfers and fund transfers by checks, drafts, or similar paper instruments. These payments systems are currently governed by articles 3 and 4 of the U.C.C. A brief history of the development of the N.P.C. can be found in a memorandum to the N.C.C.U.S.L. prepared by Hal S. Scott, Reporter to the 3-4-8 Committee. Memorandum from Hal S. Scott to the Nat'l Conf. of Comm'rs on Uniform State Laws (June 5, 1983) (available from the American Law Institute, 4025 Chestnut St., Philadelphia, PA 19104) [hereinafter cited as Scott Memorandum]. For recent commentary on the N.P.C., see Brandel & Soloway, *Electronic Fund Transfers and the New Payments Code*, 38 BUS. LAW. 1355 (1983); Brandel & Geary, *Electronic Fund Transfers and the New Payments Code*, 37 BUS. LAW. 1065 (1982); Miller, *A Report on the New Payments Code*, 39 BUS. LAW. 1215 (1984); Scott, *Corporate Wire Transfers and the Uniform New Payments Code*, 83 COLUM. L. REV. 1664 (1983).

Once a final draft by the 3-4-8 Committee is approved by the P.E.B., it must be approved by the N.C.C.U.S.L. and the A.L.I. Although the N.C.C.U.S.L. has discussed the proposed N.P.C. at two meetings, it has not completed a first reading. The A.L.I. has neither discussed nor debated the proposed N.P.C.

All citations to the N.P.C. in this Article are to the UNIFORM NEW PAYMENTS CODE (P.E.B. Draft No. 3, 1983) (available from the American Law Institute, 4025 Chestnut St., Philadelphia, PA 19104) [hereinafter cited as N.P.C.]. The provisions of this draft are, of course, subject to change by subsequent drafts.

all types of funds transfers except cash transfers.[6] The N.P.C. also seeks to have the same legal consequences attach to similar kinds of funds transfers, regardless of the method of transfer.[7] These contemplated changes will affect the allocation of losses in cases of both forgeries and material alterations of checks. The most significant impact will occur in cases in which the wrongdoer forges the signature of the drawer on a check.

This Article will analyze the impact of the N.P.C. on the allocation of loss in the specific case of forged drawers' signatures. It will (1) briefly review the current state of the law under the U.C.C. as to the allocation of loss caused by a forged drawer's signature; (2) identify the major changes that will occur if P.E.B. Draft No. 3 of the N.P.C. is enacted as federal or uniform state law; (3) discuss and evaluate the rationales offered by the drafters of the N.P.C. for these changes; (4) evaluate these changes; and (5) offer some alternatives.

These issues will be discussed from the perspective of the potential individual lawsuits that may follow when a wrongdoer forges the signature of a drawer on a check.[8] This pragmatic approach is the easiest to follow, especially for those not already familiar with payments law under the U.C.C.

The N.P.C. contains two major changes with respect to the allocation of losses caused by a forged drawer's signature. First, the rules dealing with the customer's (drawer's) liability for unauthorized items paid out of the customer's account have been modified considerably, especially the rules dealing with negligence.[9] Second, the N.P.C. would abolish the longstanding doctrine of *Price v. Neal*.[10] This change will fundamentally alter

---

[6] *See* Scott Memorandum, *supra* note 5, at 1, 6; N.P.C., *supra* note 5, §§ 2, 10, and accompanying comments.

[7] *See* Scott Memorandum, *supra* note 5, at 1; Miller, *supra* note 5, at 1215. Some of the N.P.C.'s proposed changes in payments law can be attributed to neither of the drafters' goals. The drafters appear to have utilized this opportunity to make changes in the law that they felt were desirable, but that really were not necessary to achieve the goals of the N.P.C. *See* Scott Memorandum, *supra* note 5, at 2.

[8] The N.P.C. has abandoned most of the terminology currently employed in articles 3 and 4 of the U.C.C. In order to facilitate an easy transition for the reader between discussions of the U.C.C. material and the N.P.C. material, we will, where possible, discuss the N.P.C. rules using U.C.C. terminology. The term "item" will be used interchangeably with the term "check."

[9] *See infra* notes 51–105 and accompanying text.

[10] 3 Burr. 1354, 97 Eng. Rep. 871 (K.B. 1762); *see infra* notes 106–84 and accompanying text.

the allocation of forgery losses among the various banks and other parties participating in the check collection process.

I. BACKGROUND: THE CHECK COLLECTION PROCESS AND THE EFFECTS OF A FORGERY

A. *The Check Collection Process*[11]

After the drawer of a check has issued it to the payee,[12] the payee may initiate the check collection process,[13] or he may transfer the check to another person who will initiate the collection process. Collection begins when the payee (or other person to whom the item may have been transferred) deposits the item in a depositary bank.[14] The process of collection is basically the transferring of the item from the depositary bank to the bank on which it was drawn, the payor bank.[15]

---

[11] What follows is a brief overview of the check collection process and is not intended to be a thorough treatment of the topic. For a more thorough treatment of this topic, see generally Leary, *Check Handling Under Article Four of the Uniform Commercial Code*, 49 MARQ. L. REV. 331 (1965); Malcolm, *How Bank Collection Works—Article 4 of the Uniform Commercial Code*, 11 HOW. L.J. 71 (1965); Comment, *supra* note 2.

[12] The U.C.C. defines "issue" as "the first delivery of an instrument to a holder or a remitter." U.C.C. § 3-102(1)(a) (1978). Although not stated in the U.C.C., the "drawer" of a check is the owner of the account on which the check is drawn and out of which it will be paid. The drawer is typically the person who wrote and signed the check. A check can, however, be signed by a drawer's agent. *Id.* § 3-403. *See generally*, R. BRAUCHER & R. RIEGERT, INTRODUCTION TO COMMERCIAL TRANSACTIONS 64 (1977); D. WHALEY, PROBLEMS & MATERIALS ON NEGOTIABLE INSTRUMENTS 3–4, 23 (1981). This Article uses the terms "drawer" and "customer" interchangeably.

The U.C.C. defines a "check" as "a draft drawn on a bank and payable on demand." U.C.C. § 3-104(2)(b) (1978). Like all drafts it is a written order (direction) by a drawer (customer) to the drawee (payor bank) to pay a sum of money to a third party (payee) or other person as instructed by the payee. *Id.* §§ 3-104(2)(a)–(b); R. BRAUCHER & R. RIEGERT, *supra* at 64; D. WHALEY, *supra* at 3–4; *see infra* note 15. A check, whether negotiable or not, is referred to as an "item" in the context of check collections. U.C.C. § 4-104(1)(g) (1978). The "payee" is the person to whom or to whose order the drawer orders the payor bank to pay. R. BRAUCHER & R. RIEGERT, *supra* at 64; D. WHALEY, *supra* at 3–4. Upon receipt of the check the payee becomes a "holder" of the check. U.C.C. § 1-201(20) (1978).

[13] Unlike most participants in the check collection process, the payee is not required to initiate collection within a specific time period. But, under the operation of § 3-501 through § 3-503 of the U.C.C., the payee's recourse against the drawer or indorser, or both, may be lost by a delay in initiating collection. *See, e.g.*, U.C.C. § 3-503(2) (1978).

[14] The U.C.C. defines a "depositary" bank as the "first bank to which an item is transferred for collection even though it is also the payor bank." *Id.* § 4-105(a).

[15] The U.C.C. defines "payor bank" as "a bank by which an item is payable as drawn or accepted." *Id.* § 4-105(b). In article 3, the payor bank is referred to as the drawee bank. *Id.* § 3-102 comment 3. The drawee is the entity on which the check or draft is drawn. *Id.* § 3-503(2). R. BRAUCHER & R. RIEGERT, *supra* note 12, at 64; D. WHALEY *supra* note 12, at 3–4. The check is an order by the customer to the drawee, but it is issued to a third party, the payee. The collection process then is the transmittal of that order to the drawee (or payor bank) so that it may be carried out.

The movement of the item from the depositary bank to the payor bank may follow one of several routes, depending in large part on the location of the depositary bank in relation to the payor bank. Under the U.C.C., the primary responsibility of the depositary bank is to send the item to the payor bank using a reasonably prompt method[16] while exercising ordinary care.[17] The depositary bank may send the item directly to the payor bank.[18] Typically, however, the item passes through one or more intermediary banks[19] before it reaches the payor bank.

Ultimately, the item will be presented to the payor bank for payment.[20] At that point, the payor bank must ascertain whether the item should be paid[21] or dishonored.[22] If the item is paid the proceeds will be remitted to the depositary bank and in turn to the depositary bank's customer (the depositor or payee). If, however, the item is dishonored, it will be returned to the depositary bank, the depositor, and ultimately to the drawer.

A series of credits and debits, or the transfer of money, typically accompanies the physical movement of the item from the depositary bank to the payor bank. When the payee initiates the collection process by depositing the item in the depositary bank, that bank will credit its depositor's account for the amount of the item as if the item will eventually be paid by the payor bank. These credits are usually provisional, however, with the

---

[16] U.C.C. § 4-204(1) (1978). The depositary bank acts "seasonably" if it begins transmission of the check before midnight of the banking day following the banking day of deposit. *See id.* §§ 4-202(2), 4-104(1)(h), 4-107.

[17] *Id.* § 4-202(1).

[18] *Id.* § 4-204(2).

[19] The U.C.C. defines "intermediary bank" as "any bank to which an item is transferred in course of collection except the depositary or payor bank." *Id.* § 4-105(c). Note that the depositary and all intermediary banks are also referred to as collecting banks because they handle the item for purposes of collection. *See id.* § 4-105(d). The payor bank is not a collecting bank, although it may be the depositary bank. *Id.* § 4-105(a)–(b). Collecting banks are subject to the ordinary care and seasonable action provisions of § 4-202 and § 4-204. *Id.* §§ 4-202, 4-204.

[20] "Presentment" in this context means a demand for payment made upon the payor bank on behalf of the payee (holder). *Id.* § 3-504(1). The U.C.C. defines "presenting bank" as "any bank presenting an item except a payor bank." *Id.* § 4-105(e). The presenting bank and any other prior collecting banks act as agents or sub-agents for the payee. *Id.* § 4-201(1).

[21] For a discussion of what constitutes payment, see *infra* notes 46–50 and accompanying text.

[22] For a discussion of what constitutes dishonor, see U.C.C. §§ 3-507, 3-510 and accompanying comments (1978); *see also infra* note 201. Before paying, the payor bank will consider whether the item is "properly payable" under U.C.C. § 4-401 (1978). For a discussion of the concept of properly payable in the context of forgeries see *infra* notes 51–52 and accompanying text.

bank retaining the statutory right to revoke the credit should payment not be forthcoming.[23]

In addition to crediting the depositor's account for the amount of the item, the depositary bank will usually make a corresponding debit entry in the account of the intermediary bank to which the item will be transferred.[24] In the course of processing the item, the intermediary bank will credit the account of the depositary bank and correspondingly debit the account of the bank that will next receive the item.[25] This same procedure will be employed by all subsequent intermediary banks until the item is finally presented for payment.

In determining whether the item should be paid, the payor bank may check the validity of its customer's signature[26] and determine whether the item bears all necessary indorsements.[27] The bank will also determine whether the customer's account has sufficient funds from which to pay the item.[28] If the item is not initially dishonored, the payor bank is required to settle for the item,[29] if only provisionally, by midnight of the banking day[30] on which the item was received.[31] The payor bank may settle

---

[23] The right to revoke credits, referred to as a charge-back, is given to all collecting banks (including the depositary bank). U.C.C. § 4-212 (1978). This right may be exercised by the depositary bank against its customer (the depositor) even if the bank has allowed the customer to draw on the credits. *See id.* § 4-212(4). The right terminates only when the depositary bank has received a final settlement for the item. *Id.* § 4-212(1). On the concept of final settlement see *infra* notes 46–50 and accompanying text.

[24] For example, if the depositary bank sends the check to a Federal Reserve Bank, the depositary bank will debit the account of the Federal Reserve Bank held by the depositary bank. Most checks are collected through the Federal Reserve System. R. BRAUCHER & R. RIEGERT, *supra* note 12, at 105. *See generally* Malcolm, *supra* note 11, at 72–73.

[25] Malcolm, *supra* note 11, at 72; Comment, *supra* note 2, at 815–16.

[26] The check is not properly payable unless the drawer's signature is valid. *See infra* note 51 and accompanying text.

[27] The absence of a necessary indorsement renders the check not properly payable. Perini Corp. v. First Nat'l Bank, 553 F.2d 398, 403 (5th Cir. 1977); Bank of the West v. Wes-Con Dev. Co., 15 Wash. App. 328, 331, 548 P.2d 563, 566 (1976). *See generally* B. CLARK, THE LAW OF BANK DEPOSITS, COLLECTIONS AND CREDIT CARDS § 6.4[2], at 6–36 (rev. ed. 1981).

[28] Insufficient funds in the drawer's account to pay the item will render the item not properly payable. U.C.C. § 4-104(1)(*i*) (1978). The payor bank's refusal to pay a check under such circumstances will result in no liability. *Id.* § 4-402 comment 2. The payor bank, however, may choose to pay the check even though doing so creates an overdraft. *Id.* § 4-401(1). In such a case the bank has an implied right of action against its customer for reimbursement to the extent of the amount of the overdraft. *Id.* § 4-401 comment 1; *see e.g.*, Ashford Bank v. Capital Preservation Fund, Inc., 544 F. Supp. 26, 30–31 (D. Mont. 1982).

[29] U.C.C. § 4-104(1)(j) (1978).

[30] *Id.* § 4-104(1)(c).

[31] *Id.* § 4-301(1).

in a variety of ways,[32] the most common of which is to credit the account of the presenting bank for the amount of the item.[33] If the payor bank decides to dishonor the check, it will return the check and revoke any provisional credits that it may have given.[34]

B. *The Effects of a Forgery on the Check Collection Process*

Under present law, a forgery falls into the general classification of unauthorized signatures, including those made without actual, implied, or apparent authority.[35] Relative to the other parties involved in the transaction, a forger should bear any loss resulting from his wrongdoing. He is an unlikely candidate from whom to recover, however, because he has probably absconded with or spent the fruits of his labor by the time his wrongdoing is detected. Thus, innocent parties are left to dispute the loss.

There are two basic types of check forgeries: forged drawers' signatures and forged indorsements.[36] In both cases, the customer (drawer) can usually require the payor bank to recredit his account.[37] The consequences for the payor bank of the two types of forgery, however, are quite different. In the case of a forged indorsement, the payor bank is usually allowed to pass the loss up the chain[38] to the first party who dealt with the

---

[32] *See id.* § 4-104(1)(j).

[33] In paying an item, the payor bank will debit the account of the drawer and remit the proceeds of the check to the depositary bank, which will then remit to the customer (depositor). *See* Comment, *supra* note 2, at 815; *see also* U.C.C. § 4-213(4) (1978).

[34] If it is unable to return the actual check, the payor bank may send written notice of the dishonor. U.C.C. § 4-301(1) (1978); *see, e.g.*, Blake v. Woodford Bank & Trust Co., 555 S.W.2d 589, 599 (Ky. Ct. App. 1977); Northwestern Nat'l Ins. Co. v. Midland Nat'l Bank, 96 Wis. 2d 155, 163, 292 N.W.2d 591, 596 (1980).

The intermediary banks and the depositary bank would carry out a reversal procedure as well. U.C.C. § 4-212(1) (1978). The payor bank must exercise its right to revoke any provisional credits within applicable time limits set by statute, agreement, or clearing-house rule. *Id.* § 4-213(1)(d) and comment 4; *see, e.g.*, Berman v. United States Nat'l Bank, 197 Neb. 268, 283–85, 249 N.W.2d 187, 195–96 (1976). Delay may result in liability for an item that was not properly payable or for one that the payor bank intended to dishonor. *See* U.C.C. §§ 4-301, 4-302, 4-213(1) (1978); *see e.g.*, Berman v. United States Nat'l Bank, 197 Neb. 268, 283–85, 249 N.W.2d 187, 195–96 (1976).

[35] U.C.C. § 1-201(43) (1978). This section expressly includes forgery under the definition of unauthorized signature. *Id.*

[36] Comment, *supra* note 2, at 820 n.49.

[37] In neither case of forgery would the item be properly payable under U.C.C. § 4-401 (1978).

[38] The phrase "up the chain" refers to the returning of the check through the series of collecting banks to the starting point of its journey in the collection process.

forger.[39] In contrast, in the case of a forged drawer's signature, the allocation of the loss depends to a large extent on whether the payor bank pays the item when it is presented for payment. If the payor bank pays the item, it will most likely bear the loss itself.[40]

The distinction in the treatment of these two types of forgeries has not always existed[41] and can be traced to the case of *Price v. Neal*,[42] as codified in section 3-418 of the U.C.C.[43] *Price v. Neal* stands for the proposition that payment by a drawee over a forged drawer's signature is final in favor of a good faith purchaser of the draft for value.[44] While the supporting rationale for this rule is unclear,[45] it remains an integral factor in allocating loss in the case of forged drawers' signatures.

---

[39] Payroll Check Cashing v. New Palestine Bank, 401 N.E.2d 752, 754 (Ind. Ct. App. 1980); North Carolina Nat'l Bank v. Hammond, 298 N.C. 703, 706, 260 S.E.2d 617, 621 (1979); Bank of the West v. Wes-Con Dev. Co., 15 Wash. App. 328, 331, 548 P.2d 563, 566 (1976); *see* Note, *Allocation of Liability for Checks Bearing Unauthorized Indorsements and Unauthorized Drawer's Signatures*, 24 WAYNE L. REV. 1077, 1079, 1081 (1978) [hereinafter cited as Note, *Allocation*].

[40] *See infra* notes 41–44 and accompanying text; *see also* Payroll Check Cashing v. New Palestine Bank, 401 N.E.2d 752, 754–55 (Ind. Ct. App. 1980); Note, *Allocation*, *supra* note 39, at 1079.

[41] At early common law the drawee who paid either a forged indorsement or a forged drawer's signature could obtain restitution from prior parties, because these parties were said to have been unjustly enriched by receiving money mistakenly paid by the drawee. *See* Note, *Allocation*, *supra* note 39, at 1086; Comment, *supra* note 2, at 820.

[42] 3 Burr. 1354, 97 Eng. Rep. 871 (K.B. 1762). The authority for this rule can be traced to the earlier case of Jenys v. Fawler, 2 Strange 946, 93 Eng. Rep. 959 (K.B. 1715). Note, *The Doctrine of* Price v. Neal *Under Articles Three and Four of the Uniform Commercial Code*, 23 U. PITT. L. REV. 198, 200 (1961) [hereinafter cited as Note, *Doctrine*].

[43] Section 3-418 of the U.C.C. is expressly designed to codify the rule of *Price v. Neal*. *See* U.C.C. § 3-418 comment 1 (1978).

[44] 3 Burr. 1354, 97 Eng. Rep. 871 (K.B. 1762).

[45] At least five conflicting justifications for the *Price v. Neal* rule have been advanced by various commentators.

First, because the drawee is in the best position to detect the forgery and to prevent the loss, failure of the drawee to do so is considered negligence. Ames, *The Doctrine of* Price v. Neal, 4 HARV. L. REV. 297, 298 (1891); O'Malley, *Common Check Frauds and the Uniform Commercial Code*, 23 RUTGERS L. REV. 189, 202 (1969) [hereinafter cited as O'Malley, *Check Frauds*]; Note, *Allocation*, *supra* note 39, at 1090. Although the drafters of the U.C.C. cite this rationale as the traditional justification for the distinction between treatment of the two types of forgeries, they appear to reject it in favor of the finality rationale discussed below. *See* U.C.C. § 3-418 comment 1 (1978). In comment 3 to U.C.C. § 3-417, however, the drafters state that "[t]he justification for the distinction between forgery of the signature of the drawer and forgery of an indorsement is that the drawee is in a position to verify the drawer's signature by comparison with one in his hands, but has ordinarily no opportunity to verify an indorsement." *Id.* § 3-417 comment 3.

To the extent that a negligence theory forms the rationale for *Price v. Neal*, this rationale is deficient. First, the loss for the forged drawer's signature will fall on the payor (drawee) bank even if the bank exercises the utmost care. Second, the defense of contributory negligence is unavailable to the payor (drawee) bank. *Id.* § 3-418 com-

II. Allocation of Loss If the Payor Bank Pays the Check

Under section 4-213 of the U.C.C., final payment[46] of a check can be made in basically any one of the following three

---

ment 4. Even if the presenter is negligent in taking the check from the forger, the payor bank (drawee) will be unsuccessful in shifting the loss to that party. In addition, such a negligence theory is problematic given the immense volume of checks and the computerized processing procedures commonly used by banks today. *See* Hudak & MacPherson, *supra* note 2, at 73; Murray, Price v. Neal *in the Electronic Age: An Empirical Survey*, 87 Banking L.J. 686, 688, 696–97 (1970); Comment, *supra* note 2, at 825.

Related to the negligence rationale is the idea that the rule encourages banks to be cautious in examining signatures. Farnsworth, *Insurance Against Check Forgery*, 60 Colum. L. Rev. 284, 302 (1960). While the doctrine might encourage caution in payment procedures, one commentator suggests that "the expense of litigation required for restitution would provide an equally effective incentive for caution." Comment, *supra* note 2, at 825.

A second suggested rationale is that the payor bank (drawee) has a duty to know the drawer's signature and is estopped from denying its validity after payment. O'Malley, *Check Frauds*, *supra*, at 202.

> It was incumbent upon the plaintiff, to be satisfied "that the bill drawn upon him was the drawer's hand," before he accepted or paid it: but it was not incumbent upon the defendant, to inquire into it . . . . The plaintiff lies by, for a considerable time after he has paid these bills; and then found out "that they were forged:" and the forger comes to be hanged. He made no objection to them, at the time of paying them.

Id. at 202–03 n.83 (quoting Price v. Neal, 3 Burr. at 1357, 97 Eng. Rep. at 872).

A third rationale argues that public policy favors finality. Promoting speed and certainty in commercial transactions requires an end to the process of check collection at some point. This rationale is the most commonly accepted justification for the rule. O'Malley, *Check Frauds*, *supra*, at 203; Note, *Allocation*, *supra* note 39, at 1090. The drafters of the U.C.C. appear to accept this justification as a "less fictional rationalization" than the best position rationale. U.C.C. § 3-418 comment 1 (1978). It should be noted, however, that the payor bank is permitted to reopen the transaction when there is a breach of warranty or when there is a forged indorsement. U.C.C. §§ 3-418, 3-417, 4-207; *see supra* notes 38–39 and accompanying text; *see infra* notes 110–12 and accompanying text. These exceptions appear to be contrary to a policy of finality.

Fourth, some believe that *Price v. Neal* is justified because the equities weigh against the payor bank (drawee). At least one commentator concludes that "there can be little doubt that this is the theory upon which Lord Mansfield relied in denying recovery [in *Price v. Neal*] . . . ." O'Malley, *Check Frauds*, *supra*, at 203 n.85. This argument assumes that between two innocent parties, one of whom must bear the loss, the one with legal title to the draft should prevail. That party is the holder of the item. *See* Ames, *supra*, at 299.

Fifth, some argue that the doctrine allows for wider loss distribution because payor banks are encouraged to obtain forgery insurance, the cost of which is distributed in the form of banking charges to customers. Comment, *supra* note 2, at 826; Farnsworth, *supra*, at 302–03. This result more uniformly spreads the losses from forgery among all potential forgery victims. Moreover, forgery insurance is generally not available to parties other than banks. Farnsworth, *supra*, at 297–301.

None of these justifications seems to offer a completely satisfactory explanation for the doctrine.

[46] The use of the term "final payment" in § 4-213 of the U.C.C. has a somewhat different meaning than the phrase "payment . . . is final" in § 3-418 of the U.C.C. Section 4-213 refers primarily to the timing of final payment whereas § 3-418 refers primarily to the legal consequences of final payment. U.C.C. §§ 3-418, 4-213 (1978);

ways:[47] (1) the payor bank can pay the item in cash; (2) the payor bank can complete the process of posting as described in section 4-109 of the U.C.C.;[48] or (3) the payor bank, having made a provisional settlement for the item, can fail to revoke the settlement[49] within the time permitted by statute, clearing-house rule, or agreement.[50]

Note, *Allocation*, *supra* note 39, at 1088 n.75. One commentator explains: "Section 4-213 prescribes those events which constitute final payment by a payor bank, while section 3-418 sets out those conditions under which a payment may be recovered even though final under Section 4-213." Griffith, *Final Payment and Warranties on Presentment Under the Uniform Commercial Code—Some Aspects*, 23 DRAKE L. REV. 34, 34 (1973).

[47] In addition to the three methods of effecting final payment listed in the text, an item will be deemed finally paid where a payor bank has "settled for the item without reserving a right to revoke the settlement and without having such a right under statute, clearing house rule or agreement." U.C.C. § 4-213(1)(b) & comment 4 (1978). This method was omitted from the textual discussion because it is not a typical method of final payment.

[48] The U.C.C. defines the process of posting as follows:

> The "process of posting" means the usual procedure followed by a payor bank in determining to pay an item and in recording the payment including one or more of the following or other steps as determined by the bank:
> (a) verification of any signature;
> (b) ascertaining that sufficient funds are available;
> (c) affixing a "paid" or other stamp [to the check];
> (d) entering a charge or entry to a customer's account;
> (e) correcting or reversing an entry or erroneous action with respect to the item.

Id. § 4-109.

The most controversial provision of § 4-109 is subsection (e), which arguably provides that final payment does not occur until after the expiration of the time within which the bank can correct or reverse an entry. The Wisconsin Supreme Court, in West Side Bank v. Marine Nat'l Exch. Bank, 37 Wis. 2d 661, 155 N.W.2d 587 (1968), interpreted this subsection to mean that the process of posting was not complete until the midnight deadline (or such later time under a clearing-house rule) had expired, because only then did the opportunity to correct the entry expire. This holding is widely criticized as being inconsistent with the drafters' intent. *See, e.g.*, J. WHITE & R. SUMMERS, HANDBOOK OF THE LAW UNDER THE UNIFORM COMMERCIAL CODE § 16-4, at 623–25 (2d ed. 1980); *see also id.* § 17-7, at 700–03. *See generally* Malcolm, *Reflections on* West Side Bank*: A Draftsman's View*, 18 CATH. U.L. REV. 23 (1968); Note, *Final Payment and the Process of Posting Under the Uniform Commercial Code*, 68 COLUM. L. REV. 349 (1968).

[49] For a discussion of the right to revoke, see *supra* note 23 and accompanying text. In addition, a payor bank becomes "accountable" for an item held beyond midnight of the banking day of receipt if it has not made a provisional settlement for the item. The payor bank is also accountable for any item that it has not paid or returned before the bank's midnight deadline, regardless of whether it has made a provisional settlement for the item. This undue retention doctrine can be found in U.C.C. § 4-302 (1978). For the history of the undue retention doctrine see Blake v. Woodford Bank & Trust Co., 555 S.W.2d 589, 598–99 (Ky. Ct. App. 1977). *See generally* Note, *Retention of a Check: Payor Bank's Liability Under Section 4-302*, 10 B.C. INDUS. & COM. L. REV. 116 (1968).

[50] Provisions that set time limits typically require that the payor bank dishonor the item by its midnight deadline, which is "midnight on its next banking day following the banking day on which it receives the relevant item . . . ." U.C.C. § 4-104(1)(h) (1978). The member banks of a clearing-house can allow themselves additional time. *Id.* § 4-103(1)–(2); *see id.* § 4-213 comment 4.

As check collection basically involves three parties—the drawer, the payor bank, and the prior transferors—three possible suits that can arise as a result of paying a forged check and that involve these three parties will be examined.

A. *Drawer v. Payor Bank*

1. Allocation of Loss Under the U.C.C.

Upon receipt of a bank statement revealing a forgery, a bank customer whose signature has been forged on a paid item will usually demand that the payor bank recredit his account. As a general rule, the payor bank will be required to recredit the customer's account, because the payor bank is only permitted to charge a customer's account when an item is properly payable.[51] It is clear that a check with a forged drawer's signature is not properly payable.[52]

The payor bank may assert several defenses in response to its customer's demand for recredit. The customer may be precluded from asserting the unauthorized nature of his signature if he was negligent and thereby contributed to the making of the forgery or if he was negligent in detecting and reporting forgeries by promptly reviewing his bank statement.[53]

---

[51] *See id.* § 4-401(1). The U.C.C. states: "As against its customer, a bank may charge against his account any item which is otherwise properly payable from that account . . . ." *Id.* By implication, the payor bank may not charge its customer's account if the item is not properly payable.

[52] The U.C.C. definition of properly payable only states that the term "includes the availability of funds for payment at the time of decision to pay or dishonor." *Id.* § 4-104(1)(*i*). Yet clearly a check containing a forged drawer's signature is not properly payable. "Any unauthorized signature is wholly inoperative as that of the person whose name is signed unless he ratifies it or is precluded from denying it . . . ." *Id.* § 3-404(1). Thus, because the customer in the case of a forged drawer's signature has not directed the payor bank to pay the item and to charge his account, courts and commentators agree that the item is not properly payable. *See, e.g.*, Fireman's Fund Ins. Co. v. Security Pac. Nat'l Bank, 85 Cal. App. 3d 797, 831, 149 Cal. Rptr. 883, 905 (1978); B. CLARK, *supra* note 27, at § 16.2; Note, *Uniform Commercial Code—Articles 3 and 4—Bank Required to Disburse Funds After Final Payment*, Northwestern Nat'l Ins. Co. v. Midland Nat'l Bank *96 Wis. 2d 155, 292 N.W.2d 591 (1980)*, 64 MARQ. L. REV. 408, 410 (1980) [hereinafter cited as Note, *Uniform Commercial Code*]; Comment, *supra* note 2, at 819.

[53] Under the U.C.C., the customer "who by his negligence substantially contributes to . . . the making of an unauthorized signature" will be precluded from asserting the forgery. U.C.C. § 3-406 (1978). If the customer is precluded or estopped from denying the validity of his signature, the signature is presumed to be valid because under the U.C.C. each signature on an instrument is admitted unless specifically denied in the pleadings. *Id.* § 3-307(1).

Section 4-406(1) of the U.C.C. creates the so-called bank statement duty whereby

The U.C.C. provides that the payor bank must assert the negligence defenses that it might have against the customer or be precluded from proceeding against prior transferors of the forged item.[54] Partly because of this provision,[55] the payor bank frequently asserts customer negligence as a defense.[56] Even if the payor bank can establish customer negligence, however, the payor bank can still be required to recredit its customer's account if the customer can establish that the payor bank was also negligent.[57]

Additionally, as a defense to the demand for recredit, the payor bank may assert the statute of limitations or the failure of the customer to "discover and report his unauthorized signature" within one year, regardless of negligence by customer or payor bank.[58] Failure to report will result in preclusion from asserting the forgery against the bank.[59]

"the customer must exercise reasonable care and promptness to examine the statement and items to discover his unauthorized signature . . . and must notify the bank promptly after discovery thereof." *Id.* § 4-406(1). Failure to comply with the bank statement duty precludes the customer from asserting the forgery "if the bank also establishes that it suffered a loss by reason of such failure." *Id.* § 4-406(2)(a). This would occur when, for example, the forger becomes insolvent or flees the jurisdiction during the delay period. Thus, the payor bank would lose the opportunity to recover its losses because the customer failed to report the forgery promptly. *Id.* The customer's negligence, however, is excused if the bank was also negligent. *Id.* § 4-406(3). In addition the dilatory customer will be precluded from asserting forgery with respect to subsequent forgeries by the same wrongdoer. *Id.* § 4-406(2)(b).

[54] *Id.* § 4-406(5). On the issue of types of claims that could be based on the forgery see *infra* notes 105–37 and accompanying text.

[55] Another reason for the payor to resist the customer's demand for recredit is the unlikelihood, under the U.C.C., that the payor will be able to pass the loss to the other parties in the check collection process. *See infra* notes 105–37 and accompanying text.

[56] We base this assertion mainly on the paucity of cases in which the payor bank did not assert customer negligence as a defense. Of course, the lack of cases may be attributable to some extent to the willingness of banks to recredit their customer's account voluntarily upon notification of forgery. Murray, *supra* note 45, at 701–05. As the amount of the forgery (and thus the loss to the payor bank) increases, however, the bank undoubtedly becomes less willing to recredit their customer's account voluntarily.

Based on the findings of Prof. Murray, it is also not unusual for the payor bank simply to recredit its customer's account and either to absorb the loss or to collect under its forgery insurance policy. *Id.* at 701–05, 713–15. *See infra* notes 152–55.

[57] The U.C.C. provides that the customer's negligence operates as a preclusion only in favor of a "payor who pays the instrument in good faith and in accordance with the reasonable commercial standards of the drawee's or payor's business." U.C.C. § 3-406 (1978). "The preclusion . . . does not apply if the customer establishes lack of ordinary care on the part of the bank in paying the item(s)." *Id.* § 4-406(3).

[58] *Id.* § 4-406(4).

[59] *Id.* § 4-406(4). While ordinarily the negligence of the bank will excuse the negligence of the customer, the negligence of the bank will not excuse or toll the running of the statute. The customer will most likely be precluded from proceeding against collecting banks and other prior good faith transferors as well. *See infra* notes 185–91 and accompanying text. The customer, however, will still have a valid cause of action against the forger, at least until the corollary state statute of limitations runs out. *Id.* § 4-406(4).

2. Allocation of Loss Under the N.P.C.

Section 101(1) of the N.P.C. provides that the payor bank agrees not to pay any item unauthorized[60] by the customer, thus effectively retaining the properly payable concept of section 4-401 of the U.C.C.[61] As under the U.C.C., the customer has a cause of action against the payor bank under the N.P.C. should the payor bank refuse to recredit his account upon request.[62] The defenses available to a payor bank under the N.P.C. against the customer's action are set out in sections 200, 202, and 203.

Section 200 contains several important rules. First, if an item was paid by the payor bank after receiving notice that the item was unauthorized,[63] or if the item was not an accepted access device,[64] the customer is not liable to any party and his account may not be charged.[65] This provision effectively holds the customer immune from any loss or liability when, for instance, the payor bank pays an item out of the customer's account in a case in which the forger has intercepted a new blank check from the mail prior to its reaching the customer.[66] On the other hand, when a thief or forger steals a pad of blank checks from the customer, the resulting forgeries would qualify as accepted access devices, because the checks were at one time under the control of the customer.[67]

In cases involving the use of an accepted access device and the payment by the payor bank prior to timely notification that

---

[60] The N.P.C. defines unauthorized items to include items on which the signature of the customer (drawer) is forged. N.P.C., *supra* note 5, § 54(6).

[61] *See id.* § 101 comment 1.

[62] *Id.* § 101(4); *see id.* § 101 comment 4. Such a cause of action is implied under § 4-401 of the U.C.C. *See supra* note 51 and accompanying text; D. WHALEY, *supra* note 12, at 190, 193–94.

[63] The section provides that the "notice must be received at such time and in such manner as to afford the [bank] a reasonable opportunity to act [on it]." N.P.C., *supra* note 5, § 200(1). This language is taken from § 4-403 of the U.C.C., which deals with stop-payment orders. *See* U.C.C. § 4-403 (1978).

[64] An "accepted access device" is defined in the N.P.C. to include those devices that the customer requests and receives or over which the customer has control. N.P.C., *supra* note 5, § 201. "Access device" is defined as "a card, check, code, passbook, or any other means of access to an account, or any necessary combination thereof, that may be used to initiate an order." *Id.* § 50(18).

[65] *Id.* § 200(1).

[66] The rule regarding accepted access devices is necessary to avoid making the customer liable for $50 of the loss. *Id.* § 200(2)(a); *see infra* notes 69–73, 93–105 and accompanying text. A customer's own negligence may preclude an assertion that an item is unauthorized. N.P.C., *supra* note 5, § 202. It is unclear whether the drafters of the N.P.C. intended a customer's negligence to preclude an assertion that an access device is not an accepted one.

[67] *See supra* note 64.

the item is unauthorized, the N.P.C. provides a set of loss allocation rules that represent a major departure from the rules of the U.C.C. Under section 200(2) of the N.P.C. the allocation of loss between the customer and the payor bank for a forged drawer's signature depends upon the amount of the check, the type of account involved, and the presence or absence of customer negligence.[68]

If the value of the forged item or series of items paid by the payor bank is less than $500 and the item is drawn on a consumer account,[69] the customer will be liable for a maximum of $50, regardless of whether the customer is negligent.[70] The customer will also be liable for any additional losses caused by a failure to examine the bank statement properly and to report the payment of unauthorized items.[71]

If, on the other hand, the value of the item or series of items on which the consumer drawer's signature is forged is greater than $500, or if the item, regardless of the amount, was not paid out of a consumer account, the amount of the loss that the customer must bear is not limited by the N.P.C. and is essentially based on a determination of whether the customer was negligent.[72] Under section 200(3) of the N.P.C., the payor bank has the burden of proof with respect to showing proper authorization or showing customer liability for an unauthorized order.[73]

---

[68] N.P.C., *supra* note 5, § 200(2).

[69] An account is defined in the N.P.C. as "a liability in money, credit extended or interest in assets on which orders may be drawn or to which orders may be credited." *Id.* § 50(1). A consumer account is "an account established with an account institution [i.e. a bank] in the name of one or more individuals, unless such individuals have represented in writing to the account institution that the account is not to be used primarily for personal, family or household purposes." *Id.* § 50(12). This provision is one of several provisions in which different rules apply when a consumer or a consumer account is involved. *See, e.g., id.* §§ 101(2)(b) (availability of consequential damages), 3(1)(a) (prohibition on variances of N.P.C. rules with consumers). These provisions represent a significant departure from the U.C.C. The drafters of the N.P.C., however, are considering the elimination of all or most of the special rules regarding consumer transactions from the N.P.C. *See* Miller, *supra* note 5, at 1220–22.

[70] This result assumes that the customer received proper disclosures pursuant to § 701 of the N.P.C. Section 701 requires the payor bank to provide the customer with summaries of various terms and conditions of the account. N.P.C., *supra* note 5, § 701.

[71] *See infra* notes 86–90 and accompanying text; *see also supra* note 53 and accompanying text.

[72] N.P.C., *supra* note 5, §§ 200(2)(b), 202–203. Failure to comply with the duties regarding examination of the bank statement under § 203 or other acts and omissions set forth in § 202 constitute negligence. *Id.* §§ 202–203.

[73] For example, the payor bank would have to show that the items were ordered and received by the customer or were otherwise under the control of the customer. *See id.* § 54. The payor bank would also have the burden of proving that the item was drawn on a consumer account. *Id.* § 200(2)(a). Should the payor bank wish to hold the customer

In cases in which the value of the forged item or items paid out of a consumer account amounts to more than $500, or in which the item or items are not paid out of a consumer account, the N.P.C. provides that the customer shall not be liable nor shall his account be charged unless the customer is precluded through negligence from asserting that the item is unauthorized.[74]

Section 202 of the N.P.C., entitled "Strict Liability and Negligence Preclusions Against Drawer," contains several strict liability provisions, none of which are found in its counterpart, section 3-406, of the U.C.C.[75] Under section 202, the customer is precluded from asserting that an item is unauthorized unless the customer shows that certain of his acts or omissions in no way contributed to the loss resulting from the forgery.[76] The negligence of the payor bank is not relevant in the determination of loss allocation when one of these strict liability provisions applies.[77] The strict liability provision relevant to forged drawers' signatures is subsection (1)(d). Under this subsection the customer must promptly report the theft or loss of blank checks upon discovery or he will face preclusion.[78]

In cases where the strict liability provisions do not apply, section 202(2) utilizes a general negligence standard that is very similar to that employed in section 3-406 of the U.C.C.[79] If a customer's negligence substantially contributes to the forgery of his check, the customer is precluded from asserting that an item is unauthorized unless the party seeking to assert the customer's negligence has itself paid or given value for the item in bad faith or has otherwise not acted in accordance with the reasonable commercial standards of its business.[80]

Three points should be made about section 202. First, neither the statute nor the comments define negligence or give any

liable under § 200(2)(a) when several items are forged, the bank would also have the burden of proving that the items were forged by the same person or group of persons acting together. *Id.* §§ 200(2)(a), (3).

[74] *Id.* §§ 200(2)(b), 202–203.

[75] *Compare id.* § 200(1)(a)–(d) *with* U.C.C. § 3-406 (1978); *see infra* notes 77–78 and accompanying text.

[76] N.P.C., *supra* note 5, § 202(1).

[77] This provision represents a major departure from the loss allocation scheme of the U.C.C. The negligence of the payor bank will always offset that of the customer under U.C.C. §§ 3-406, 4-406(3) (1978). *See supra* note 57 and accompanying text.

[78] N.P.C., *supra* note 5, § 202(1)(d).

[79] The N.P.C. provides that a person "is otherwise precluded from asserting as drawer that an order is unauthorized if the person's negligence substantially contributed to the order becoming unauthorized . . . ." *Id.* § 202(2).

[80] *Id.* § 202(2).

content to the phrase "substantially contributed." In failing to define the terms the drafters have missed an excellent opportunity to resolve the substantial confusion that has existed under the U.C.C. with respect to these concepts.[81] Second, unlike the N.P.C. strict liability preclusions, section 202 allows the ordinary negligence of the customer to offset that of the payor bank.[82] Third, the section does not give any content to the concept of reasonable commercial standards, even though a comment to the section suggests that failure of the payor bank to examine its customer's signature on items under a certain amount may meet reasonable commercial standards.[83] Thus, under the N.P.C., a court has the opportunity to place the loss from a forged drawer's signature on the negligent customer, even if the payor bank failed to read and to compare the signature. Such a result is not possible under the U.C.C.[84]

Under the N.P.C., failure to examine and to correct a bank statement also constitutes customer negligence. Section 203 of the N.P.C. requires all customers, both consumer and nonconsumer, to examine statements promptly, with reasonable care, and to report to the payor bank the payment of any unauthorized items.[85] If the customer fails to act accordingly, he will be precluded from asserting that a paid item was unauthorized, unless the payor bank fails to show that it acted in good faith and in accordance with the reasonable commercial standards of its business.[86]

Section 203 of the N.P.C. is written along the lines of section 4-406 of the U.C.C. and should operate in a similar fashion.[87] Nevertheless, a few changes should be noted. First, the customer's duty to detect forgeries exists whether or not the items were actually returned to the customer.[88] This change recognizes the

---

[81] *See* Whaley, *Negligence and Negotiable Instruments*, 53 N.C.L. REV. 1 (1974) [hereinafter cited as Whaley, *Negligence*].

[82] This provision parallels the U.C.C. rule. *See supra* note 77.

[83] *See* N.P.C., *supra* note 5, § 202 purpose comment 4.

[84] Whaley, *Negligence*, *supra* note 81, at 15. Recall that in cases in which an N.P.C. strict liability provision applies, the customer is liable for the resulting loss even though the payor bank is negligent. Under the U.C.C. the payor bank's negligence will always offset that of the customer. *See supra* note 57 and accompanying text.

[85] N.P.C., *supra* note 5, § 203.

[86] The section specifies that "[m]ere payment of an unauthorized 'item' is not failure to act in accordance with reasonable commercial standards." *Id.* § 203(3).

[87] *Compare id.* § 203 *with* U.C.C. § 4-406 (1978).

[88] N.P.C., *supra* note 5, § 203(1)–(2)(b), § 200 purpose comments 1 & 3.

increasing use of check truncation.[89] Second, when the items have not been returned to the customer, the period allowed the customer for reviewing his statement and reporting unauthorized items has been extended from fourteen days under the U.C.C. to sixty days under the N.P.C.[90]

Under section 204 of the N.P.C., the payor bank is required to assert any of these valid defenses against the customer upon request (by any other participant in the check collecting process) or it will be barred from asserting claims against collecting banks and other prior transferors based on the unauthorized item.[91] The U.C.C. contains a similar rule.[92]

Thus, the N.P.C. contains two sets of rules whose application depends upon whether the forged item was drawn on a consumer account and whether the value of the forged item or series of items amounted to more than $500. In contrast, under U.C.C. liability rules, no customer is liable for losses caused by a forged drawer's signature unless that customer is negligent.[93] A practical problem arises, however, when the customer complains about the payment of a forged item under the U.C.C. scheme. Frequently, the payor bank refuses to recredit the account and alleges that the customer was negligent,[94] because the payor bank will be unable to pass the loss on to the collecting banks or other prior transferors if it recredits the customer's account, voluntarily or otherwise.[95] When the payor bank interposes a negligence defense, the customer, even a non-negligent customer, will frequently give up and absorb the loss, because litigating the matter is not economical when the amount at stake is under $500.[96]

Under the N.P.C., when the loss amounts to less than $500 the payor bank is prohibited from arguing that the consumer

---

[89] *See* N.P.C., *supra* note 5, § 203 existing law comment 1. Check truncation refers to the type of check collection in which items are retained by the depositary bank or the payor bank and are not physically returned to the customer. *See* notes 170–74 and accompanying text.

[90] *Compare* U.C.C. § 4-406(2)(b) (1978) *with* N.P.C., *supra* note 5, § 203(2)(b).

[91] N.P.C., *supra* note 5, § 204(6).

[92] U.C.C. § 4-406(5) (1978); *see supra* notes 54–55 and accompanying text.

[93] *See supra* notes 53–55, 59, and accompanying text.

[94] *See supra* note 56 and accompanying text.

[95] *See infra* notes 105–37 and accompanying text. There is considerable authority for the proposition that U.C.C. § 4-406(5) (1978) requires the payor bank to assert any § 3-406 negligence against the customer. *See* Whaley, *Negligence*, *supra* note 81, at 20–21.

[96] *See* N.P.C., *supra* note 5, § 200 purpose comment 2.

was negligent as a way to avoid recrediting the account.[97] Under the N.P.C. scheme, the non-negligent consumer will be worse off than under the U.C.C., because he will absorb $50 of the loss; in another important respect, however, he will be better off, because the N.P.C. rules relieve the consumer of the burden and expense of litigating the issue of negligence with the payor bank when less than $500 is at stake. The negligent consumer, on the other hand, will clearly be better off under the N.P.C. than under the U.C.C. Under the U.C.C., the negligent customer absorbs all of the loss, not to mention the possibility of litigation expenses.[98] Under the N.P.C., the negligent consumer will absorb no more than $50 of the loss when the forgery amounts to less than $500.[99]

The N.P.C. two-tiered approach is sound, partly because it recognizes the economic realities of litigating disputes involving small amounts. This rationale, however, should also be applied to the business customer for whom the costs of litigating the issue of negligence are just as great. For many businesses, especially small ones, these costs are just as burdensome as for consumers. What is not economical for a consumer is, in this case, no more economical for a business customer.[100]

Although the N.P.C. two-tiered approach is a sound one, the $50 liability provision is unjustified and should be removed from the N.P.C. The imposition of the $50 loss on consumers when the forgery or forgeries is under $500 is designed to "provide some incentive against negligent use . . ." of checks.[101] This rationale does not support the $50 liability provision. The $50 liability provision will do little to deter negligent behavior because the $50 loss is imposed whether or not the customer is negligent. Because the $50 liability rule will do little, if anything, to deter negligence, and because its automatic imposition will

---

[97] When the loss is less than $500 the payor must recredit all but $50 of the amount to the consumer's account even though the consumer was negligent. A negligence defense is not available to the payor bank unless the value of the forged item or series of items is greater than $500, or the item, regardless of the amount, was not paid out of a consumer account. *See id.* § 200(2)(a)–(b).

[98] *See supra* note 53 and accompanying text.

[99] *See supra* notes 68–74 and accompanying text.

[100] The economics of the matter should be based not on the assets of the customer, but on the costs of litigation relative to any possible recovery. Even though a business customer may be wealthier than a consumer customer, the absolute costs of litigation relative to the possible recovery will typically be the same for both.

[101] N.P.C., *supra* note 5, § 200 purpose comment 2; *see also id.* § 200(2)(a).

operate as a penalty on the non-negligent consumer, the $50 liability provision should be eliminated in all cases.[102]

Removing the $50 penalty will not encourage negligent behavior by consumers because other portions of section 200 of the N.P.C. provide adequate deterrence. In nearly every act of forgery, there is a considerable chance that the resulting loss will be over $500. When the loss is over $500, the potential liability of the customer is not limited.[103] Because the customer cannot know whether any given forged item will be in an amount greater than $500,[104] the customer has a substantial incentive to be careful with all checks. Removing the $50 liability provision will not remove the incentive to be careful in handling checks.[105]

The two-tiered system of liability under the N.P.C. is sound and represents a needed improvement over the current rules under the U.C.C. The protections given to consumers under section 200(2) of the N.P.C., however, should be extended to business customers, and the $50 liability provision should be

---

[102] It is not clear why the drafters imposed the $50 liability provision only on consumers and not on business customers. Many forgeries occur in a business or organizational setting. Under the N.P.C., business customers will suffer no loss unless the payor bank proves that the businesses were negligent. *Id.* § 200(2)(b). If consumers need a $50 penalty as an incentive against negligent behavior, then the same incentive is needed with business customers. Another inequity of the $50 liability provision is that it may be imposed on the non-negligent consumer even when the payor bank is grossly negligent.

[103] Liability is unlimited in two respects. First, the customer may lose whatever funds are in the account if the amount of the forged item is equal to or greater than the amount of funds in the account. Second, if the amount of the forged item is greater than the amount of funds in the account, § 101(5) of the N.P.C. entitles the payor bank to reimbursement from the customer for the overdraft. N.P.C., *supra* note 5, § 101(5). While this subsection applies only to authorized items, under § 202 of the N.P.C., the customer may be precluded from asserting that the item was unauthorized. *See id.* § 202. Thus, under the operation of these provisions of the N.P.C., the customer has potentially unlimited liability.

[104] In all but the most unusual case, the amount of the forged check is determined by the forger, not the customer.

[105] The $50 liability provision is modeled after a similar provision in the federal Truth-in-Lending Act. N.P.C., *supra* note 5, § 200 purpose comment 2. The Truth-in-Lending Act provision places a $50 limit on the consumer's liability for losses resulting from the fraudulent use of credit cards. 15 U.S.C. § 1643 (1982 & Supp. 1983). In the credit card area, the $50 liability provision does serve as an incentive against negligent behavior because the consumer has no liability for any losses over that amount, regardless of the consumer's care or negligence. (The Truth-in-Lending Act does not utilize different loss allocation rules depending on the amount of the loss. 15 U.S.C. § 1643 (1982 & Supp. 1983). The $50 liability provision in the N.P.C. is unnecessary because the customer is subject to unlimited liability when a forgery or forgeries amount to more than $500. Removing the $50 liability provision from the N.P.C. should not change the appropriateness of the two-tiered loss allocation scheme, which is based on whether the amount of the loss is over or under $500. The rationale behind this two-tiered approach is based on a recognition of the economic realities of litigation when small amounts of money are involved.

removed because it operates as a penalty on the non-negligent customer and is unnecessary to deter careless handling of checks.

B. *Payor Bank v. Collecting Banks and Prior Transferors*

1. Allocation of Loss Under the U.C.C.

If the payor bank is forced to recredit its customer's account, it will search for someone from whom to recoup its loss. The major obstacle to the payor bank's attempt to recover up the chain is the doctrine of *Price v. Neal*, as codified in the final payment provisions of the U.C.C.[106] Section 3-418 of the U.C.C. provides that:

> Except for recovery of bank payments . . . and except for liability for breach of warranty on presentment under the preceding section, payment or acceptance of any instrument is final in favor of a holder in due course, or a person who has in good faith changed his position in reliance on the payment.[107]

In addition, sections 4-213 and 4-302 of the U.C.C. provide that upon noncash final payment,[108] the payor bank becomes "accountable" for the item.[109]

---

[106] 3 Burr. 1354, 97 Eng. Rep. 871 (K.B. 1762); *see supra* notes 42–45 and accompanying text; *see also* U.C.C. §§ 3-417, 4-207 and accompanying comments.

[107] U.C.C. § 3-418 (1978).

[108] Section 4-213(1) provides that "[u]pon final payment under subparagraphs (b), (c) or (d) the payor bank shall be accountable for the amount of the item." *Id.* § 4-213(1). Subparagraphs (b), (c) and (d) all refer to noncash modes of payment. *Id.* § 4-213(1)(b)–(d). The payor bank is not accountable with respect to cash payment under (a) because it has already disbursed the proceeds of the check in cash. *Id.* § 4-213(1)(a); *see id.* § 4-213 comment 7; *see also* Note, *Uniform Commercial Code*, *supra* note 52, at 411. In addition, the U.C.C. provides that where there is no provisional settlement, the payor bank is accountable for an item retained past midnight of the banking day of receipt of the item, and in any event is accountable for an item retained past its midnight deadline. U.C.C. § 4-302 (1978); *see supra* note 49 and accompanying text.

[109] The U.C.C. imposes a duty on the payor bank to account for the item. U.C.C. § 4-213(1) (1978). This duty is met "if and when a settlement for the item satisfactorily clears." *Id.* § 4-213 comment 7. The term "accountable" has been held to be synonymous with the term liable, particularly in the context of § 4-302 of the U.C.C. *See* Rock Island Auction Sales, Inc. v. Empire Packing Co., 32 Ill. 2d 269, 271–72, 204 N.E.2d 721, 723 (1965); Blake v. Woodford Bank & Trust Co., 555 S.W.2d 589, 601 (Ky. Ct. App. 1977); Sun River Cattle Co. v. Miners Bank, 164 Mont. 237, 242, 521 P.2d 679, 684 (1974); Fromer Distrib., Inc. v. Bankers Trust Co., 36 A.D.2d 840, 840–41, 321 N.Y.S.2d 428, 430 (1971); Northwestern Nat'l Ins. Co. v. Midland Nat'l Bank, 96 Wis. 2d 155, 160, 292 N.W.2d 591, 596 (1980); *see also* Note, *Allocation*, *supra* note 39, at 1089.

The relevant final payment provisions of the U.C.C. do, however, contain an exception for situations involving breach of warranty.[110] The warranty provisions are contained in sections 3-417 and 4-207. With minor variation, the warranties contained in both sections are the same.[111] Despite this exception, the

---

[110] The U.C.C. expressly states that § 3-418 and § 4-302 do not apply when a breach of warranty exists. U.C.C. §§ 3-418, 4-302 (1978). Section 4-213 does not contain a similar exemption. *See id.* § 4-213. Perhaps this was an oversight by the drafters. On the other hand, this failure could be explained if the function of § 4-213 is merely one of determining when final payment occurs and not of determining its legal effect. *See supra* note 46. Under this reasoning, "even if final payment has been made under Section 4-213, recovery of that payment is still available under Section 3-418 if there is a breach of warranty on presentment." Griffith, *supra* note 46, at 37. *But cf.* North Carolina Nat'l Bank v. South Carolina Nat'l Bank, 449 F. Supp. 616, 619 (D.S.C. 1976), *aff'd* 573 F.2d 1305 (4th Cir. 1978) (the court held that breach of warranty was irrelevant after final payment because the payor bank was accountable for the item).

[111] Section 3-417 of the U.C.C. provides that:

(1) Any person who obtains payment [for an item] . . . and any prior transferor warrants to a person who in good faith pays [an item] . . . that

(a) he has good title to the instrument or is authorized to obtain payment . . . on behalf of one who has good title; and

(b) he has no knowledge that the signature of the maker or drawer is unauthorized, except that this warranty is not given by a holder in due course acting in good faith [to certain parties]; and

(c) the instrument has not been materially altered, except that this warranty is not given by a holder in due course acting in good faith [to certain parties].

(2) Any person who transfers an instrument and receives consideration warrants to his transferee and if the transfer is by indorsement to any subsequent holder who takes the instrument in good faith that

(a) he has good title to the instrument or is authorized to obtain payment . . . on behalf of one who has good title and the transfer is otherwise rightful; and

(b) all signatures are genuine or authorized; and

(c) the instrument has not been materially altered; and

(d) no defense of any party is good against him; and

(e) he has no knowledge of any insolvency proceeding instituted with respect to [certain parties on the instrument].

U.C.C. § 3-417(1)-(2) (1978).

Section 4-207 of the U.C.C. provides that:

(1) Each customer or collecting bank who obtains payment . . . of an item and each prior customer and collecting bank warrants to the payor bank or other payor who in good faith pays . . . the item that

(a) he has good title to the item or is authorized to obtain payment . . . on behalf of one who has a good title; and

(b) he has no knowledge that the signature of the . . . drawer is unauthorized, except that this warranty is not given by any customer or collecting bank that is a holder in due course and acts in good faith [to certain parties]; and

(c) the item has not been materially altered, except that this warranty is not given by any customer or collecting bank that is a holder in due course and acts in good faith [to certain parties].

(2) Each customer and collecting bank who transfers an item and receives

payor bank will be unable to shift the loss because, for the most part, no warranties will have been breached. Courts have uniformly held that the warranty of good title contained in subsection (1)(a) of each section is breached only by a forged indorsement, not by a forged drawer's signature.[112] Similarly, under subsection (1)(b) of each warranty provision, the party presenting the item warrants to the payor bank only that "he has no knowledge that the signature of the maker or drawer is unauthorized . . . ."[113] Therefore, except in the rare instance where the presenter actually knows that the drawer's signature is forged, no warranty will be breached.[114]

Further, under subsection (2) of each warranty provision, the so-called transfer warranties, the transferor warrants that all signatures are genuine; however, these warranties run only to transferees and thus will provide no protection for the payor

---

a settlement or other consideration for it warrants to his transferee and to any subsequent collecting bank who takes the item in good faith that

(a) he has good title to the item or is authorized to obtain payment . . . on behalf of one who has a good title and the transfer is otherwise rightful; and

(b) all signatures are genuine or authorized; and

(c) the item has not been materially altered; and

(d) no defense of any party is good against him; and

(e) he has no knowledge of any insolvency proceeding instituted with respect to [certain parties on the item].

In addition each customer and collecting bank so transferring an item and receiving a settlement or other consideration engages that upon dishonor and any necessary notice of dishonor and protest he will take up the item.

(3) The warranties and the engagement to honor set forth in the two preceding subsections arise notwithstanding the absence of indorsement . . . .

(4) Unless a claim for breach of warranty under this section is made within a reasonable time after the person claiming learns of the breach, the person liable is discharged to the extent of any loss caused by the delay in making claim.

Id. § 4-207(1)–(4).

[112] *See, e.g.*, Bagby v. Merrill Lynch, Pierce, Fenner & Smith, Inc., 491 F.2d 192, 199 (8th Cir. 1974); Sun N' Sand v. United Cal. Bank, 21 Cal. 3d 671, 684–87, 582 P.2d 920, 929–32, 148 Cal. Rptr. 329, 339–40 (1978); Payroll Check Cashing v. New Palestine Bank, 401 N.E.2d 752, 754, 756 (Ind. Ct. App. 1980); Aetna Life & Casualty Co. v. Hampton State Bank, 497 S.W.2d 80, 84 (Tex. Civ. App. 1973); *see also* North Carolina Nat'l Bank v. Hammond, 298 N.C. 703, 708, 260 S.E.2d 617, 623 (1979) (the warranty of good title refers only to the validity of the chain of indorsements; presumably the court meant that all indorsements necessary to the chain of title must be authentic). *See generally* J. White & R. Summers, *supra* note 48, § 16-2, at 609; Whaley, *Forged Indorsements and the UCC's "Holder"*, 6 Ind. L. Rev. 45 (1972).

[113] U.C.C. §§ 3-417(1)(b), 4-207(1)(b) (1978). Note that under § 3-417 and § 4-207 the warranty on presentment is made by the presenting party and by most prior parties as well. *See id.* §§ 3-417(1), 4-207(1).

[114] *See, e.g.*, Dozier v. First Ala. Bank, 363 So. 2d 781, 783 (Ala. Civ. App. 1979); Manufacturers & Traders Trust Co. v. County Trust Region of the Bank of New York, 59 A.D.2d 645, 646, 398 N.Y.S.2d 298, 298 (1977).

bank because it is not a transferee under the U.C.C.[115] Therefore, a cause of action for breach of warranty is not available, except when the presenter or prior transferor has knowledge of the forgery.

A payor bank, unable to prevail against prior transferors on a breach of warranty claim, may attempt to hold prior indorsers liable under the indorser's contract according to section 3-414(1) of the U.C.C.[116] Under the express terms of section 3-414(1), "every indorser engages that *upon dishonor* and any necessary notice of dishonor and protest he will pay the instrument . . . ."[117] If the payor bank has paid the instrument, however, there has been no dishonor to trigger the indorser's liability. Second, the contract does not run to the payor bank, because the indorser promises only to pay holders of the item; the payor bank is generally not considered a holder of an item.[118] Moreover, if the item has been dishonored the payor bank has suffered no damages. Thus, the payor bank can never enforce the indorser's contract.[119]

The payor bank may also attempt to sue for common law restitution.[120] Section 1-103 incorporates the common law of mistake into the U.C.C., unless contradicted by other provisions.[121] The common law right to sue for restitution has been limited, however, by section 3-418 of the U.C.C., which pro-

---

[115] An item is not transferred to the payor bank. It is presented for payment. As there is no transfer, the payor bank is not a transferee. *See* Dozier v. First Ala. Bank, 363 So. 2d 781, 784 n.3 (Ala. Civ. App. 1978); Payroll Check Cashing v. New Palestine Bank, 401 N.E.2d 752, 756 (Ind. Ct. App. 1980); North Carolina Nat'l Bank v. Hammond, 298 N.C. 703, 706–07, 260 S.E.2d 617, 621 (1979). *See generally* J. WHITE & R. SUMMERS, *supra* note 48, at 608; Comment, *supra* note 2, at 823.

Comment 4 to § 4-207 of the U.C.C. supports the contention that the term "transferees" does not include payor bank: "In this section as in Section 3-417, the (a), (b) and (c) warranties to transferees and collecting banks under subsection (2) are in general similar to the (a), (b) and (c) warranties to payors under subsection (1); but the warranties to payors are less inclusive because of exceptions reflecting the rule of *Price v. Neal* . . . , and related principles." U.C.C. § 4-207 comment 4 (1978).

[116] U.C.C. § 3-414(1) (1978).

[117] *Id.* (emphasis added).

[118] The payor bank is not a holder of a check because the check is not negotiated to it. Instead, it is presented for payment. R. BRAUCHER & R. RIEGERT, *supra* note 12, at 65; Comment, *supra* note 2, at 820 n.49.

[119] The case law consistently supports this conclusion. *See, e.g.*, Dozier v. First Ala. Bank, 363 So. 2d 781, 783 n.1 (Ala. Civ. App. 1978); Kirby v. First & Merchants Nat'l Bank, 210 Va. 88, 92, 168 S.E.2d 273, 276 (1969).

[120] *See supra* note 41.

[121] *See* U.C.C. § 1-103 (1978). "Unless displaced by the particular provisions of the Act, the principles of law and equity, including the law merchant and the law relative to . . . mistake . . . shall supplement its provisions." *Id.*

vides that payment is final in favor of a holder in due course[122] or one who in good faith changes his position in reliance upon the payment.[123] By implication, restitutionary recovery is allowed only when payment was made to one who was not a holder in due course or to one who did not in good faith change his position in reliance upon the payment.[124]

Whether the payor bank's restitutionary rights are also limited by the accountability provisions of sections 4-213 and 4-302 of the U.C.C. is open to question. Neither section expressly limits accountability or liability[125] to holders in due course or to parties who have relied on the payment.[126] Most courts ignore the uncertainty and apply section 3-418 without considering the language of the article 4 provisions,[127] even when the item is clearly in the check collection process.[128] Courts and commentators that have considered the conflict are divided as to the result. Some assert that the limitations contained in section 3-418[129] apply equally to article 4, so that restitution would be possible against a person who was not a holder in due course or who did

---

[122] *Id.* § 3-418. The requisites for acquiring holder in due course status are set forth in *id.* § 3-302.

[123] *Id.* § 3-418. Good faith is defined in the U.C.C. as "honesty in fact in the conduct or transaction concerned." *Id.* § 1-201(19).

[124] For the relevant text of § 3-418 of the U.C.C. see *supra* text accompanying note 107. An illustrative case is First Nat'l City Bank v. Altman, 3 U.C.C. Rep. Serv. (Callaghan) 815 (N.Y. Sup. Ct. 1966), *aff'd* 27 A.D.2d 706, 277 N.Y.S.2d 813 (1967), in which the court allowed restitutionary recovery on the second of two checks. The payee had sold two packets of diamonds, receiving payment in the form of two checks upon which the drawer's signature was forged. The payee (wisely) withheld delivery of the first packet of diamonds until after receiving final payment on the first check. He failed, however, to take this precaution the second time around. The court held that as to this second check, the payee lacked the requisite reliance. The court ignored the second basis under which the payee could have obtained protective status, that is, by being a holder in due course. This case is frequently criticized on this ground. *See, e.g.*, B. CLARK, *supra* note 27, § 6.2[1].

[125] *See supra* note 49 and accompanying text.

[126] In other words, both § 4-213 and § 4-302 state that under certain circumstances the payor bank is accountable for the amount of the item, but neither section limits that accountability to holders in due course or to those who can show good faith reliance.

[127] *See, e.g.*, Perini Corp. v. First Nat'l Bank, 553 F.2d 398, 416, 419 (5th Cir. 1977); Fireman's Fund Ins. Co. v. Security Pac. Nat'l Bank, 85 Cal. App. 3d 797, 825, 831, 149 Cal. Rptr. 883, 902, 905 (1978); Payroll Check Cashing v. New Palestine Bank, 401 N.E.2d 752, 757 (Ind. Ct. App. 1980); Maplewood Bank & Trust Co. v. F.I.B., Inc., 142 N.J. Super. 480, 484, 362 A.2d 44, 47 (N.J. Super. Ct. App. Div. 1976); Richardson Co. v. First Nat'l Bank of Dallas, 504 S.W.2d 812, 816 (Tex. Civ. App. 1974).

[128] Article 4 of the U.C.C. governs bank deposits and collection. When it conflicts with a provision of article 3, article 4 should prevail. *See* U.C.C. § 4-102 (1978).

[129] The limitations referred to are those that limit the finality of payment to holders in due course and to those who in good faith relied on the payment.

not rely in good faith on the payment.[130] Other courts and commentators find a conflict between articles 3 and 4 and reach the result mandated by section 4-102 of the U.C.C.: in case of a conflict, article 4 is to prevail.[131] Thus, they read the accountability provisions of sections 4-213 and 4-302 as cutting off all restitutionary claims by the payor bank.[132] There is fairly uniform consensus that payment to one who obtained payment in bad faith would be recoverable by the payor bank.[133]

---

[130] *See, e.g.*, National Sav. & Trust Co. v. Park Corp., 722 F.2d 1303, 1306 (6th Cir. 1983), *cert. denied*, 104 S. Ct. 1916 (1984); Blake v. Woodford Bank & Trust Co., 555 S.W.2d 589, 601–02 (Ky. Ct. App. 1977); Demos v. Lyons, 151 N.J. Super. 489, 498, 376 A.2d 1352, 1356 (N.J. Super. Ct. Law Div. 1977); *see also* R. BRAUCHER & R. RIEGERT, *supra* note 12, at 128–29; Note, *Uniform Commercial Code*, *supra* note 52, at 415; Comment, *supra* note 2, at 828–37.

An interesting approach was taken by the court in Bank Leumi Trust Co. v. Bally's Park Place, Inc., 528 F. Supp. 349 (S.D.N.Y. 1981). In this case, the payee deposited for collection a check with knowledge of the insolvency of the drawer's estate. Although the drawer's account had been closed for two months, the payor bank failed to dishonor the check by its midnight deadline. The check was subsequently paid and not disputed for almost seven months. The payor bank sued the payee in restitution to recover the money paid "by mistake." The court allowed the payor bank to recover the funds by applying the common law of mistake. The court disregarded § 3-418 and § 4-302 of the U.C.C., saying:

> These sections of the U.C.C. apply to interbank settlement procedures, and not to subsequent actions for restitution. A contrary reading of Sections 3-418, 4-301 and 4-302 would be inconsistent with Section 1-103 which retains the common law governing mistake. It would lead to the unintended result of allowing a payee, unjustly, to retain monies improperly obtained.

Id. at 1549. The court overlooked the fact that under § 1-103 the common law applies only if it is not displaced by the particular provisions of the Code. U.C.C. § 1-103 (1978). Therefore, the common law (through § 1-103) must be read in light of § 3-418, § 4-301, and § 4-302. The court could have reached the same result by allowing restitution against all but holders in due course or parties changing their position in good faith reliance on the payment. Then the court could have quite simply found that the payee could not qualify as a holder in due course because the payee knew of the drawer's insolvency.

[131] *See supra* note 128.

[132] *See, e.g.*, Ashford Bank v. Capital Preservation Fund, Inc., 544 F. Supp. 26, 29 (D. Mont. 1982); Kirby v. First & Merchants Nat'l Bank, 210 Va. 88, 91, 168 S.E.2d 273, 275 n.4 (1969); Northwestern Nat'l Ins. Co. v. Midland Nat'l Bank, 96 Wis. 2d 155, 164, 292 N.W.2d 591, 599 (1980) (this case is discussed in Note, *Uniform Commercial Code*, *supra* note 52); *see also* B. CLARK, *supra* note 27, at 6–7; J. WHITE & R. SUMMERS, *supra* note 48, § 16-2, at 613–18; Edwards, *Recovery of Final Payments Under the Uniform Commercial Code*, 6 OHIO N.U.L. REV. 341, 351 (1979). Some courts have applied § 4-213 of the U.C.C. to block recovery by a payor bank without considering whether the parties obtaining payment were holders in due course or had relied in good faith on the payment. *See, e.g.*, Fromer Distrib., Inc. v. Bankers Trust Co., 36 A.D.2d 840, 842, 321 N.Y.S.2d 428, 430 (1971). It is important to note that not all cases decided under U.C.C. § 4-213 are forged drawer's signature cases. Most often they are cases in which the customer's account had insufficient funds from which to pay the item. The precedential value of these cases must be considered in light of the common law rule that restitutionary relief was generally not available to the payor bank when it paid checks from accounts with insufficient funds by mistake. *See* Comment, *supra* note 2, at 834.

[133] *See, e.g.*, Perini Corp. v. First Nat'l Bank, 553 F.2d 398, 404, 416 (5th Cir. 1977);

Finally, the payor bank may try to avoid liability by asserting negligence on the part of the depositary bank. Even after *Price v. Neal*, many courts gave the payor bank the right to recover if the holder was negligent in taking the instrument from the forger.[134] The U.C.C. drafters, however, clearly rejected this line of cases.[135] Under the U.C.C., the negligence of the holder is only relevant if it amounts to a lack of good faith.[136] Thus, negligence claims against the depositary bank or other prior transferor have been unsuccessful.[137]

2. Allocation of Loss Under the N.P.C.

If the payor bank pays an item that has a forged drawer's signature, the loss allocation among the payor bank, collecting banks, and other prior transferors is governed primarily by section 204 of the N.P.C., entitled "Transmission Liability."[138] Section 204 provides that any party, including any collecting bank, who transfers an unauthorized item[139] is liable to all parties to whom the item is transferred and who pay or give value for the item in good faith.[140] Claims by the payor bank under this section against any prior transferor must be made within a reasonable time[141] and claims may be barred if the payor bank fails to assert

---

Bartlett v. Bank of Carroll, 218 Va. 240, 247, 237 S.E.2d 115, 119 (1977); U.C.C. § 3-418 comment 3 (1978); *see also* J. WHITE & R. SUMMERS, *supra* note 48, § 16-2, at 613–18; Griffith, *supra* note 46, at 39.

[134] Note, *Doctrine*, *supra* note 42, at 210; Note, *Allocation*, *supra* note 39, at 1086.

[135] U.C.C. § 3-418 comment 4 (1978). "This section rejects decisions under the original Act permitting recovery on the basis of mere negligence of the holder in taking the instrument." *Id.*

[136] *Id.* A lack of good faith would preclude holder in due course status as well as preclude good faith reliance. *Id.* §§ 3-302(1)(b), 3-418.

[137] Fireman's Fund Ins. Co. v. Security Pac. Nat'l Bank, 85 Cal. App. 3d 797, 823, 831, 149 Cal. Rptr. 883, 901, 905 (1978); Aetna Life & Casualty Co. v. Hampton State Bank, 497 S.W.2d 80, 85 (Tex. Civ. App. 1973); *see also* Note, *Allocation*, *supra* note 39, at 1086.

[138] N.P.C., *supra* note 5, § 204.

[139] A check bearing a forged drawer's signature is an unauthorized item under § 54 of the N.P.C. *Id.* § 54. *See generally* notes 51–52 and accompanying text.

[140] Like the U.C.C., the N.P.C. provides for some types of final payment. The rules governing payment of items are found primarily in sections 420 through 432. N.P.C., *supra* note 5, §§ 420–432. The N.P.C. follows the U.C.C. in many respects with regard to the various ways in which the payor bank may be held accountable for an item; but in each case the payor bank's right to bring a suit against a prior party based on transmission liability under § 204 of the N.P.C. is reserved. *See, e.g.*, *id.* § 420(3). In this way, with respect to forged drawer's signatures, the N.P.C. avoids but does not necessarily resolve the confusion and uncertainty that currently exists under the U.C.C. with respect to final payment of an item.

[141] *Id.* § 204(5).

any valid defense against the customer if requested to do so (by any other participant in the check collection process).[142] The transferor of an unauthorized item would remain liable even when the payor bank has been negligent.[143]

Section 204 will undoubtedly become controversial because it abolishes a rule of loss allocation that has been in effect for over two hundred years. Under the rule of *Price v. Neal*, as codified in the provisions of the U.C.C., payment of an item with a forged drawer's signature is final in favor of a holder in due course or anyone who in good faith relies on the payment.[144] Because the transaction under these circumstances is final and cannot be undone, the loss will remain on the payor bank unless the drawer (customer) was negligent.[145] Section 204 of the N.P.C. will abolish *Price v. Neal* by enabling the payor bank to shift the loss onto collecting banks and other prior transferors under most circumstances.

The drafters of the N.P.C. justify the abolition of *Price v. Neal* in several ways. First, the drafters assert that the rationale for the rule in *Price v. Neal* is not convincing. They state that the "traditional justification" for the rule is that the payor bank is in a superior position to detect the forgery of its own customer's signature.[146] This proposition, they suggest, is "dubious" today in light of modern banking practices in which a large volume of items are processed by computers.[147] The drafters claim that it has become impractical for the payor bank to check the validity of its customer's signature on all items, and thus the payor bank is no longer in the best position to verify the drawer's signature.[148] Commentators generally point to this best position argument to support the abolition of the rule of *Price v. Neal*.[149]

---

[142] *Id.* § 204(6).

[143] Damages under this section are based on the amount of the item plus interest and expenses (including attorney's fees but typically excluding consequential damages). *Id.* § 204(4).

[144] *See supra* notes 42–45, 106–14 and accompanying text.

[145] *See supra* notes 53, 59, and accompanying text.

[146] N.P.C., *supra* note 5, § 204 comment 2.

[147] *Id.*

[148] *Id.*

[149] *See, e.g.*, Brandel & Soloway, *supra* note 5, at 1359; *see also* Scott Memorandum, *supra* note 5, at 19. Even though the drafters of the U.C.C. do not depend on the best position rationale to justify the final payment rule, it is still popular with the courts. *See, e.g.*, Fireman's Fund Ins. Co. v. Security Pac. Nat'l Bank, 85 Cal. App. 3d 797, 822, 149 Cal. Rptr. 883, 899 (1978).

This first justification is flawed for two reasons. First, the rationale offered by the drafters of the U.C.C. for the final payment rule is not that the payor is in the best position to detect the forgery. Instead they emphasize a policy of commercial finality.[150] Second, while modern check processing has made it more difficult and inconvenient for the payor bank to check the signatures of its customers, the payor bank is still in the best position to detect this type of forgery vis-à-vis other participants in the chain. In spite of substantial efforts by the payee or the payee's transferee (for example, a grocery store or check cashing service) to verify the identity of the drawer, a forgery may well go undetected. These parties normally have no means of verifying the signature of the drawer, because they do not have this signature on file.[151] Even when the grocery store or check cashing service deals face to face with the purported drawer, or when it issues check-cashing identification cards, false identification can still result in failure to detect the forgery. Moreover, a payee who accepts a check through the mail for the purchase of goods is clearly not in the best position to verify the drawer's signature.

As a second justification, the drafters argue that it is impractical for the payor bank to check the validity of all signatures.[152] This impracticality "is reflected by the reality that banks do not check signatures under a certain dollar amount even though they will be liable."[153] Implicit in these statements is the idea

---

[150] U.C.C. § 3-418 comment 1 (1978).

> The traditional justification for the result [of placing the loss for forged drawer's signatures on the drawee (payor bank)] is that the drawee is in a superior position to detect the forgery because he has the maker's [drawer's?] signature and is expected to know and compare it; a less fictional rationalization is that it is highly desirable to end the transaction on an instrument when it is paid rather than reopen and upset a series of commercial transactions at a later date when the forgery is discovered.

Id.

[151] The payor bank is in the best position to detect the forgery of the customer's signature primarily because the payor is the only party in the check collection system that will always have the signature of the drawer on file.

[152] N.P.C., *supra* note 5, § 204 comment 2.

[153] *Id.* In addition, purpose comment 4 to § 202 of the N.P.C. also mentions that "[a]t the present time, many banks do not read the signature on checks under a certain amount." *Id.* § 202 purpose comment 4. In comment 2 to § 204, the drafters of the N.P.C. again assert that "banks do not check signatures under a certain amount," yet they cite no support for their assertion. *See id.* § 204 comment 2. The most recent empirical research on this issue indicates that most payor banks do check the drawer's signature on all or a substantial number of checks. Murray, *supra* note 45, at 698–701. Prof. Murray reports that "only one bank out of 91 large American banks has 'adopted the practice of not checking signatures on checks below a certain amount'." *Id.* at 701 (quoting Farnsworth & Leary, *U.C.C. Brief Number 10: Forgery and Alterations of Checks*, 14 PRAC. LAW. 75, 76 (1968)). The authors of this Article are not aware of any

that at some dollar amount, it would become economical for the payor bank to verify its customer's signature given the costs of verification and the certainty and extent of liability. Yet, section 204 does not provide for different loss allocation rules depending on the amount of the item. The drafters of the N.P.C. utilize an economic argument to justify abolishing the current rule, but readily ignore the same economic argument in the course of fashioning the new rule.

Even if the economic realities of modern banking justify relieving the payor bank of the legal duty to verify all signatures, they do not justify relieving the payor bank of the duty to verify at least some signatures. If a greater amount of the loss caused by forged drawers' signatures is to be placed on prior parties, section 204 should be modified to impose liability only on the transfer of an unauthorized item under a specific amount, perhaps $500.[154] Such a provision would allow most items to be processed quickly, without signature verification. Economic considerations would justify verification of the relatively small number of items over this threshold amount.[155] Under the modification proposed here, section 204 transmission liability would apply to unauthorized items under the threshold amount, while the current U.C.C. rules would apply to items over that amount.[156]

---

further empirical research on this question since the publication of Prof. Murray's work in 1970; nor do the authors have any evidence, other than anecdotal evidence, that contradicts the basic findings of Prof. Murray's empirical research.

[154] A recent study showed that, as of 1980 the amount of the average check is $570. ARTHUR D. LITTLE, INC., ISSUES AND NEEDS IN THE NATION'S PAYMENT SYSTEM 12 (1982), *cited in* Scott, *supra* note 5, at 1664 n.1. The threshold figure for liability, like other figures in the N.P.C., would be adjusted to account for inflation according to the formula provided for this purpose. N.P.C., *supra* note 5, § 5.

[155] Recall that the most recent empirical evidence shows that most bankers do verify the signatures of their customers. *See* Murray, *supra* note 45, at 698–701. Financial institutions that process their customers' checks could easily comply with our proposed rule. Computer systems, which bankers claim make it difficult for them to check signatures, are well suited to selecting items over a certain amount which can then be inspected visually. An operator at the depositary bank encodes the amount of the check on its face using magnetic ink character recognition (MICR) symbols. A computer at the payor bank reads these symbols and can be programmed to pick out all items over a certain amount. *See* B. CLARK, *supra* note 27, at § 10.5; N. PENNY & D. BAKER, THE LAW OF ELECTRONIC FUND TRANSFER SYSTEMS §§ 1.01[2], 1.02[1]–[2] (1980). This practice is currently utilized by at least some banks. Murray, *supra* note 45, at 698–701.

[156] It should be emphasized that the risk of loss from forged drawers' signatures should not be shifted away from the payor bank. If the economic argument used by the drafters of the N.P.C., however, justifies such a shift, then that same argument should be used to determine the extent or degree of the shift. The shift, if there must be one, should take place only to the extent that signature verification by the payor bank is not economical. Based on Prof. Murray's findings it is economical to verify signatures on items much smaller than our proposed figure of $500. Murray, *supra* note 45, at 705–13.

The drafters of the N.P.C., in comment 2 to section 204, also discuss the rationale behind the final payment rule of the U.C.C., which is a policy of promoting finality in commercial transactions.[157] Instead of fostering the reopening of a whole series of transactions the U.C.C. simply leaves the loss on the payor bank.[158]

The drafters of the N.P.C. correctly point out a problem with utilizing this rationale for the final payment rule.[159] If public policy favors finality, then the rule of *Price v. Neal* is appropriate because it typically prohibits the payor bank from recovering the payment through a suit in restitution or warranty.[160] This same rationale, it seems, should also operate in the context of forged indorsements, but it does not. In instances of forged indorsements, the payor bank is allowed to undo the whole series of transactions by bringing a warranty action under section 4-207 or section 3-417 of the U.C.C. against the party who obtained payment as well as all other prior transferors.[161] If the U.C.C. favors finality in instances of forged drawers' signatures it should also favor finality in the case of a forged indorsement.[162]

The drafters of the N.P.C. recognize this problem. They have chosen to treat both types of forgeries alike and to shift the loss away from the payor bank.[163] The position taken by the N.P.C. drafters is problematic in several respects.

The drafters of the N.P.C. apparently reject the equally logical alternative position, which is to reaffirm the finality rationale and to apply it to both types of forgeries. This would place the loss in both types of cases on the payor bank. Accordingly, payment over a forged drawer's signature or over a forged indorsement would be final in favor of certain types of transferors.[164] The policy of commercial finality has merit. It prevents

---

[157] N.P.C., *supra* note 5, § 204 comment 2.

[158] U.C.C. § 3-418 comment 1 (1978); Note, *Allocation*, *supra* note 39, at 1090.

[159] N.P.C. *supra* note 5, § 204, comment 2.

[160] *See supra* notes 105–37 and accompanying text.

[161] *See* U.C.C. §§ 3-417(1)(a), 4-207(2)(a)–(b); *supra* notes 38–39 and accompanying text.

[162] Profs. White and Summers write: "We can discern no adequate rationale to explain the difference between the liability of the drawee bank on checks bearing forged indorsements and its liability on those bearing forged drawer's signatures." J. WHITE & R. SUMMERS, *supra* note 48, at 610.

[163] Under the N.P.C. an unauthorized item would include an item which lacks the required valid indorsements. N.P.C., *supra* note 5, § 54. Liability for transmitting such an item would also be governed by § 204. *Id.* § 54 comment 3.

[164] A requirement of holder in due course status or of good faith reliance on the payment could be imposed as it is currently imposed under the U.C.C. U.C.C. § 3-418 (1978).

transactions from becoming so contingent as to disrupt commerce, and it provides a measure of assurance to those who accept checks. The drafters have failed to justify their rejection of this alternative solution to such a significant issue.[165]

The drafters of the N.P.C. might reject this alternative with a best position argument in cases of forged indorsements. If, however, the best position argument is used implicitly by the drafters to shield payor banks from bearing the losses caused by forged indorsements, then the drafters should not discard the same argument in considering the issue of the allocation of losses caused by forged drawers' signatures. Yet the drafters do discard that argument in that situation.

Much of the confusion and dispute over determining the appropriate rationale for the rule of *Price v. Neal* is caused by viewing the alternative rationales as mutually exclusive. In deferring to one policy the other need not be disregarded. Commercial finality has positive benefits for business and commerce, yet the reasoning behind the best position argument has considerable merit as well.[166] The loss allocation scheme should begin with a policy of finality by which the payor bank will bear the loss caused by forged indorsements as well as forged drawers' signatures. When one of the parties is in a significantly better position to detect the forgery, however, the best position argument should prevail over the goal of finality. For example, because an indorsee's transferee is in a significantly better position to detect a forged indorsement than the payor bank, the loss should be shifted away from the payor bank and onto the transferee in cases of forged indorsements.

Because the payee's transferee is not in a better position to detect a forged drawer's signature than the payor bank,[167] the finality rationale should prevail in cases of a forged drawer's

---

[165] The drafters of the N.P.C. state, somewhat poetically, that N.P.C. § 204(1) "marks the death knell" of *Price v. Neal*. N.P.C., *supra* note 5, § 204 comment 2. This change is no small matter. Prof. Murray suggests that without *Price v. Neal* and the protection it affords payees and their transferees, the check would not have become the dominant method of funds transfer. Murray, *supra* note 45, at 688. The death of *Price v. Neal* may, in turn, mark the death knell of checking.

[166] The loss should be placed on the party who is in the best position to avoid the loss by detecting the forgery. When liability is imposed on the party in the best position to detect a forgery, such parties are encouraged to detect forgeries in order to minimize their exposure to liability. This allocation of loss will most effectively reduce the aggregate amount of economic loss caused by forgeries. Placing the loss liability on any other party would be less than optimally efficient, because such persons would need to expend more resources to detect forgeries than the party in the best position.

[167] *See supra* notes 149–52 and accompanying text.

signature. This result is the most efficient and effective way to minimize loss. Removing any liability for paying an item with a forged drawer's signature will effectively remove any incentive for the payor bank to verify the customer's signature.[168] The loss allocation scheme proposed by the drafters of the N.P.C. should be rejected and replaced with the preferable scheme of the U.C.C., which is justified on both finality and best position grounds.[169]

The position taken by the drafters of the N.P.C. on the effects of check truncation is also of concern.[170] In comments to section 204, the drafters argue that the rule in *Price v. Neal* "makes no sense in cases of check truncation."[171] In some respects this statement is true. The development of check truncation, however, does not necessarily call for the abolition of the rule in *Price v. Neal*, because check truncation has already been developing under the rule.[172] From this fact it can be inferred that even though *Price v. Neal* may be retarding the growth of check truncation, at least some bankers are willing, because of the

---

[168] On the matter of payor incentive to verify the customer's signature, see *infra* notes 179–81 and accompanying text.

[169] *See supra* notes 106–37 and accompanying text. A review of the official commentary to sections 3-417, 3-418, and 4-207 of the U.C.C. leads the authors to believe that the drafters embraced both the finality policy and the best position arguments. Together these arguments support the distinction between the treatment of forged drawer's signatures and forged indorsements. U.C.C. § 3-417 comment 3, § 3-418 comment 1, § 4-207 comments 1 & 4 (1978).

[170] *See supra* note 89. With the type of check truncation in which the depositary bank retains the items, payor banks do not receive the items in any physical form. Therefore, the payor bank typically has no opportunity to verify its customer's signature. Such a system would appear to be incompatible with a loss allocation scheme like that under the U.C.C., which ultimately depends on signature verification. The type of check truncation in which the payor bank retains the items is completely compatible with such a scheme, because with this system, the items are returned to the payor bank but not to the customer.

[171] N.P.C., *supra* note 5, § 204 comment 2. Although the drafters fail to specify whether they are referring to the payor bank retention or the depositary bank retention type of check truncation, one can assume that they are referring to the latter.

[172] A new study by Trans Data Corporation found that approximately 25% of all financial institutions truncate checking services. This proportion would be greater today, suggests Trans Data, but for the decline in publicity about truncation within the banking establishment and the channeling of resources into developing and promoting newly authorized depository instruments and services. Garsson, *Truncation Losing Ground as Banks Concentrate on Other Matters*, Am. Banker, Apr. 18, 1984, at 15, col. 1. While the majority of this truncation has been of the payor bank retention type, the depositary bank retention type of check truncation has been successfully implemented on an experimental basis. *See European American Will Soon Test Fast Check Processing System: 'Digitized Image' Reduces Size of Paper by Three-Fourths, Allowing High-Speed Handling*, Am. Banker, June 3, 1984, at 3, col. 1 [hereinafter cited as *Check Processing Systems*]; *see also* N. PENNY & D. BAKER, *supra* note 155, §§ 2.01–2.03 (1980).

cost savings check truncation offers, to absorb the forgery losses that will result from this system. In addition, the retention of the *Price v. Neal* rule while check truncation is developing will spur the development of technology capable of verifying or transmitting signatures electronically.[173]

The drafters of the N.P.C. are also critical of the rule in *Price v. Neal* for "not giving adequate incentives to payees to check on the bona fides of people drawing checks to them."[174] They claim that "[u]nder existing law, a merchant cashing a check need not be concerned with whether a person paying by check is actually the owner of the account on which the check is drawn."[175] In reality, though, current rules provide adequate incentive for merchants to check the bona fides of their customers who pay by check. In order to obtain protection under the final payment rule, the merchant or other payee must establish that he is a holder in due course or has relied in good faith on payment of the item.[176] In addition, all risks, including checks drawn on accounts with insufficient funds and on closed accounts, are borne by the payee who accepts the item or by the transferee of a third party check until final payment of the item.[177] Even after final payment, the payee's transferee contin-

---

[173] The drafters of the N.P.C. expressly assume that such technology will not be available at a reasonable cost. We believe that such an assumption is shortsighted and incompatible with the effort to construct a legal framework for governing payments in the future. *See Check Processing Systems*, *supra* note 172.

If the rule of *Price v. Neal* is truly incompatible with the development of check truncation, perhaps check truncation should be discouraged. After all, check truncation is a cost saving measure and should not form the sole basis of allocating losses which result from forgeries. Furthermore, if the cost savings with check truncation come at the expense of proportionally larger forgery costs, check truncation hardly makes economic sense. Bankers should devise cost saving measures that are compatible with a loss allocation scheme based on sound policy.

[174] N.P.C., *supra* note 5, § 204 comment 2.

[175] *Id.*

[176] *See, e.g.*, National Sav. & Trust Co. v. Park Corp., 722 F.2d 1303, 1305–07 (6th Cir. 1983), *cert. denied*, 104 S. Ct. 1916 (1984); Perini Corp. v. First Nat'l Bank, 553 F.2d 398, 404 (5th Cir. 1977). This incentive depends upon whether the court will utilize § 3-418 or § 4-213 of the U.C.C. in determining the legal effect of final payment, and whether the restrictions under § 3-418 will be read into § 4-213. *See supra* notes 124–33 and accompanying text.

[177] Until final payment, the payor bank has the absolute right to dishonor the item and to revoke any provisional credits so long as it acts within the applicable time limits. *See supra* note 22. Following dishonor, the collecting bank may exercise its right to charge-back against prior collecting banks and ultimately the payee. *See supra* note 23 and accompanying text; *see, e.g.*, Whalen & Sons Grain Co. v. Missouri Delta Bank, 496 F. Supp. 211, 214 (E.D. Mo. 1980) (payor bank not liable to payee for checks drawn on insufficient funds since items were returned before midnight deadline).

ues to bear the risks of altered items, as well as forged or missing indorsements.[178]

Moreover, under current law, a merchant has every reason to be concerned with whether a person paying by check is actually the owner of the account on which the check is drawn. It is quite possible, especially in the case of stolen blank checks, that after receiving notice from the customer about a theft, the payor bank will dishonor or return a forged check.[179] In such cases, the loss will fall squarely on the merchant. The loss falls on the payor bank only if it fails to dishonor the item.

Although the drafters of the N.P.C. are concerned that under current law the payee has insufficient incentive to check on the bona fides of people drawing checks on them, under the N.P.C. the payor bank would not have sufficient incentive to handle its customers' money carefully. Under section 204, the negligence of the payor bank is not relevant to the allocation of loss. In comments to section 204, the drafters state that the N.P.C. "rejects the possibility of allowing the collector of a check to avoid liability by showing that the payor . . . was negligent, e.g. failed to detect an obvious forgery on a $1 million check."[180] Because the negligence of the payor bank will not preclude shifting the loss caused by unauthorized items onto prior parties in the check collection process, section 204 of the N.P.C. does not provide the payor bank with any incentive to be careful in disbursing funds from its customer's account. No other provision of the N.P.C. provides the payor with such an incentive.[181]

---

[178] The transfer or presentment of altered items, as well as items bearing forged or missing indorsements, will generate warranty liability for the payee. U.C.C. §§ 3-417(1)(a), (c), 3-417(2)(a)–(c), 4-207(1)(a), (c), 4-207(2)(a)–(c) (1978); *see supra* notes 110–14 and accompanying text; *see, e.g.*, Sun N' Sand v. United Cal. Bank, 21 Cal. 3d 671, 688–92, 582 P.2d 920, 932–34, 148 Cal. Rptr. 329, 341–44 (1978) (altered items); Atlantic Bank v. Israel Discount Bank, 108 Misc. 2d 342, 346–47, 441 N.Y.S.2d 315, 318 (N.Y. App. Term 1981) (forged indorsement); First Nat'l Bank v. Nunn, 628 P.2d 1110, 1115 (Mont. 1981) (missing indorsement).

[179] *See supra* notes 22–23 and accompanying text.

[180] N.P.C., *supra* note 5, § 204 comment 2.

[181] It is true that under § 101 of the N.P.C., the payor bank agrees with its customer to pay only authorized items out of its customer's account; but in the event that the payor bank pays an unauthorized item over $500, even though grossly negligent, it can simply refuse to recredit the account and thereby put the burden of bringing suit on the customer. Moreover, in such a suit, the payor bank can bring a third party claim against all prior parties under § 204 of the N.P.C. All of this litigation could be prevented through greater care on the part of the payor bank. Under § 204, the payor bank may be precluded from holding prior parties liable only if the payor bank was not acting in good faith. *Id.* § 204(1). Additionally, the payor bank's payment of unauthorized checks may often deplete the customer's account, causing wrongful dishonors. Although the customer will have an action against the payor bank under § 101 for wrongful dishonor,

The drafters of the N.P.C. wish to replace a loss allocation scheme which they mistakenly claim does not provide adequate incentive for payees and depositary banks to be careful in accepting checks with one that provides inadequate incentive for the payor bank to be careful.

Because of the wide-reaching scope of transmission liability under section 204, the payor bank can recover from any collecting bank or other prior transferor.[182] Transmission liability extends not only to the payor bank but to any subsequent party who pays, accepts, or gives value for the item in good faith.[183] Therefore, the presenting bank that is liable to the payor bank may, in turn, hold any prior collecting bank or other prior transferor liable under section 204. This process could continue until the loss is eventually passed to the person who dealt with the forger, which usually will be the payee or the party who took the item from the payee.[184]

C. *Drawer v. Collecting Banks and Other Prior Transferors*

1. Allocation of Loss Under the U.C.C.

If the payor bank refuses to recredit its customer's account,[185] becomes insolvent, or can successfully defend on the grounds of customer negligence,[186] the customer could initiate an action

---

a wrongful dishonor may result in substantial injuries and expenses, not to mention inconvenience, which may not be compensable under § 101. Under § 101(2) of the N.P.C., consequential damages are only available in limited circumstances. *Id.* § 101(2). Moreover, the customer's action against the payor bank for wrongful dishonor under N.P.C. § 101 suffers from a number of other weaknesses, the discussion of which is beyond the scope of this Article. *See generally* Dow, *Wrongful Dishonor of a Check under Uniform Commercial Code § 4-402 and Proposed New Payments Code § 101: A Comparison*, 1984 SELECTED PAPERS OF THE AM. BUS. LAW ASS'N REGIONAL PROCEEDINGS [TRI-STATE AND MIDWEST BUSINESS LAW ASS'NS] 32.

[182] N.P.C., *supra* note 5, § 204(1).

[183] *Id.* § 204(1). "Each customer, transmitting account institution or transferor of an unauthorized draw order is liable to all parties to whom the draw order is subsequently transmitted . . . ." *Id.*

[184] The forger is typically insolvent or unavailable. The payor bank may bring a suit directly against the payee or depositary bank under § 204 of the N.P.C. The depositary bank or payee will be liable to the payor bank. *Id.* This action will substantially shorten the sequence, while bringing about the same result as the series of suits suggested in the text.

[185] *See, e.g.*, Perini Corp. v. First Nat'l Bank, 553 F.2d 398 (5th Cir. 1977) (because the customer utilized a facsimile signature machine and had authorized the payor bank to pay any items bearing the facsimile signature regardless of who had used the machine, the payor bank was not required to recredit the customer's account).

[186] *See supra* notes 53-57 and accompanying text.

against one or more collecting banks or other prior transferors. Nothing in the U.C.C. expressly provides the customer with a cause of action against these parties. If the customer brings an action nevertheless, the obstacles to recovery would be essentially the same as those that the payor bank would face in an action against these same parties.

If the customer asserts breach of warranty against a collecting bank or other prior transferor, he will most likely be unsuccessful. Although it appears that the drawer can properly bring such an action,[187] he will stand in the payor bank's position and thus will typically be unable to establish liability based on breach of warranty, because no warranties have been breached in the typical forged drawer's signature situation.[188]

If the customer asserts either common law negligence or restitution as grounds for recovery, the action will also be unsuccessful, because courts have held that the final payment rule will block the customer to the same extent that it would block the payor bank's action.[189] Allowing the drawer to prevail against the collecting bank in a case in which the drawer would not prevail against the payor bank (and in which the payor bank would not prevail against the collecting bank), would conflict with the policy of final payment under the U.C.C.[190] In shifting the loss from the drawer to a collecting bank, the transaction would be reopened, and the loss would be shifted to a party who would not have been in the best position to prevent the loss.[191] The drawer whose signature was forged and who failed to obtain a recredit from the payor bank will probably not prevail against collecting banks (or transferors subsequent to the forgery), except in cases in which the party taking the instrument did so in bad faith.[192]

---

[187] Sun N' Sand v. United Cal. Bank, 21 Cal. 3d 671, 680–83, 582 P.2d 920, 927–28, 148 Cal. Rptr. 329, 336–37 (1978).

[188] *See supra* notes 110–15 and accompanying text.

[189] Perini Corp. v. First Nat'l Bank, 553 F.2d 398, 416 (5th Cir. 1977); Fireman's Fund Ins. Co. v. Security Pac. Nat'l Bank, 85 Cal. App. 3d 797, 831, 149 Cal. Rptr. 883, 905 (1978).

[190] Perini Corp. v. First Nat'l Bank, 553 F.2d 398, 417 (5th Cir. 1977); Fireman's Fund Ins. Co. v. Security Pac. Nat'l Bank, 85 Cal. App. 3d 797, 831, 149 Cal. Rptr. 883, 905 (1978). *See generally* Note, *Direct Drawer Suits Against Collecting Parties for Loss Arising From Unauthorized Checks*, 17 WAKE FOREST L. REV. 844, 857 (1981).

[191] See *supra* note 45 for a discussion of the rationales behind the final payment rule.

[192] *See* Perini Corp. v. First Nat'l Bank, 553 F.2d 398, 417 (5th Cir. 1977). In this case, the question on remand was whether the depositary bank acted in bad faith when it paid checks, questionably indorsed and worth over $1 million, to a man with a taped

2. Allocation of Loss Under the N.P.C.

The problems that the drawer would encounter under the N.P.C. in bringing an action against collecting banks and other prior transferors are similar to those encountered under the U.C.C. The courts would very likely interpret section 204 of the N.P.C. as allowing the customer whose bank paid an unauthorized item to sue these prior parties for transmission liability.[193] Although neither the language of the section nor the comments expressly allow such an action, the courts will probably interpret parties who pay under section 204(1) to include a customer whose account has been charged for an unauthorized item.[194] Support for this conclusion can be found in cases interpreting the term "payor" under the U.C.C. to include a customer.[195] The customer's cause of action, however, will have limited significance because, like the payor bank, collecting banks could assert the customer's negligence as a defense. Under sections 200 and 202, the customer who is negligent or who has engaged in any of the acts for which the strict liability rule applies will be precluded from asserting that the item was unauthorized.[196] This provision will bar the customer from bringing an action under section 204 of the N.P.C. for transmitting an unauthorized order. To assert this negligence preclusion, the collecting bank or other prior transferor must, of course, have acted in good faith and in accordance with the reasonable commercial standards of its business.[197]

The comments to section 204 do suggest, though, that the collecting banks may not be able to assert the negligence of the customer as a defense.[198] Recall that when the payor bank brings an action against the collecting banks under section 204, the collecting banks may *not* assert the negligence of the payor bank

---

on Fu-Manchu mustache.

It may also be possible for the payor bank, and thus the drawer, to prevail if the presenter was neither a holder in due course nor changed his position in good faith reliance upon the payment. *See supra* rotes 117–34 and accompanying text.

[193] *See supra* notes 139–43 and accompanying text.

[194] N.P.C., *supra* note 5, § 204(1). Under this subsection, transmission liability extends to all parties who pay in good faith. *Id.*

[195] *See, e.g.*, Sun N' Sand v. United Cal. Bank, 21 Cal. 3d 671, 680–83, 582 P.2d 920, 927–28, 148 Cal. Rptr. 329, 336–37 (1978).

[196] N.P.C., *supra* note 5, §§ 200(2)(b), 202(1)–(2); *see supra* notes 60–80 and accompanying text.

[197] N.P.C., *supra* note 5, § 202(2).

[198] *Id.* § 204 comment 2.

to avoid transmission liability.[199] By analogy, the collecting banks may not assert the negligence of the customer to avoid transmission liability to the customer. On the other hand, section 204(6) makes the negligence of the customer relevant in an action by a payor bank against collecting banks; it requires the payor bank to assert customer negligence as a defense to a customer's action to recredit the account.[200] The intended significance of section 204(6) when the customer, instead of the payor bank, brings the action is unclear. It seems, though, that if negligence of the payor bank would not preclude section 204 transmission liability on the part of collecting banks and other prior parties, the negligence of the customer should likewise not preclude liability when the customer brings the action.

III. Allocation of Loss When the Payor Dishonors the Check

A. *Allocation of Loss Under the U.C.C.*

If the payor bank discovers the forged drawer's signature and returns the item in a timely manner, the loss allocation scheme differs substantially from the scheme when the item is paid.[201]

---

[199] *See supra* text accompanying notes 139–43.

[200] N.P.C., *supra* note 5, § 204(6).

[201] There is some question whether, from a technical standpoint, returning a check because of a forged drawer's signature constitutes a dishonor or merely a refusal to pay followed by a return. Under the U.C.C., an instrument is dishonored when, "in the case of bank collections the instrument is seasonably returned by the midnight deadline (Section 4-301)." U.C.C. § 3-507 (1978); *see supra* notes 21–22 and accompanying text. The U.C.C. uses the term "dishonor" as well as the term "return" in discussing the payor bank's actions in revoking provisional settlement and in recovering any payments already made. *Id.* § 4-301. From the perspective of these two sections, dishonor appears to be used in the broad sense of returning the check unpaid or of returning the check and revoking any provisional settlement. On the other hand, comment 2 to § 3-510 of the U.C.C. utilizes the term "dishonor" in a much narrower sense. *Id.* § 3-510 comment 2. According to this comment, returning an item that bears a "forgery" (presumably this includes a forged drawer's signature) is not evidence of a dishonor, but is evidence of a justifiable refusal to pay. *Id.*

From one perspective this issue has little practical importance. Whether the failure of the bank to pay a check because of a forged drawer's signature is classified as a dishonor, or merely a justifiable refusal to pay, the check clearly has not been paid, so the final payment rule and its various manifestations will not come into play. From another perspective, however, the issue is a critical one. Liability under the various article 3 contracts often depends, at least in part, on whether there has been a dishonor. *See, e.g.*, *id.* §§ 3-413(2), 3-414(1). Some components of article 4 liability also depend on whether there has been a dishonor. *See, e.g.*, *id.* § 4-207(2) (final sentence). If the payor bank's refusal to pay a check because of a forged drawer's signature does not

Although there are virtually no reported cases where the payor bank discovered a check with a forged drawer's signature and dishonored it in a timely manner,[202] the U.C.C. clearly allocates the loss in such cases. In general, if the payor bank discovers the forged drawer's signature and dishonors the item, it will revoke any provisional credits it may have given to the presenting bank.[203]

The presenting bank can then bring one of two different actions against prior collecting banks.[204] The presenting bank can bring an action for breach of the transfer warranty[205] that all signatures are genuine.[206] Any collecting bank found to be in breach of this warranty may, in turn, bring an action against prior parties on the same warranty.[207] Alternatively, the presenting bank can bring an action against any prior indorser on the indorser's contract.[208]

Any prior indorser found liable may, in turn, bring a contract action against the indorser prior to it.[209] Through either course

---

consititute a dishonor, the liability of some parties may not be triggered.

Any resolution of this issue is somewhat speculative because there are virtually no reported cases in which the payor bank discovered a forged drawer's signature before final payment and returned the check before its midnight deadline. For purposes of allocating the loss among the parties involved in a check collection, a court would be likely to consider the payor bank's refusal to pay such a check a dishonor and apply the liability provisions under articles 3 and 4 accordingly. Part III of this Article is written from this perspective.

Note that the payor bank can discover the forgery and can dishonor the item at any time before final payment occurs. Comment, *supra* note 2, at 817. At that time, the payor bank can also revoke any provisional credits it may have given. U.C.C. § 4-301(1) (1978). Collecting banks may do the same even if the customer has been allowed to draw on the credits. *Id.* § 4-212.

[202] It is unclear whether the lack of cases exists because payor banks rarely discover a forged drawer's signature within the time allowed to dishonor the check, or because banks allocate any losses in such situations without resort to a lawsuit, by simply utilizing the right of charge-back. *See* U.C.C. §§ 4-212, 4-301 (1978).

[203] *Id.* § 4-301. Again, this conclusion assumes that final payment has not occurred.

[204] J. WHITE & R. SUMMERS, *supra* note 48, § 16-2 at 608. For an overview of the check collection process and the participants in that process, see *supra* notes 11–34 and accompanying text.

[205] On the narrow issue of who makes such a warranty and to whom is it made, see U.C.C. §§ 3-417(2), 4-207(2) (1978).

[206] *Id.* §§ 3-417(2)(b), 4-207(2)(b).

[207] *Id.* § 4-207(2)(b). *See also id.* § 3-417(2)(b). Under § 4-207(2)(b), each customer and collecting bank who transfers the item and receives settlement or consideration makes this warranty to the transferee and subsequent collecting banks who take the item in good faith. *See supra* notes 110–15 and accompanying text.

[208] The indorser's contract is found in U.C.C. § 3-414(1) (1978). Liability under this contract is triggered by the payor's dishonor. *See supra* notes 116–19 and accompanying text. Liability may also depend on notice of dishonor, U.C.C. §§ 3-501(2), 3-502, 3-508, 3-511 (1978), and on protest, *id.* §§ 3-501(3), 3-502(2), 3-509, 3-511. "Protest" is governed by U.C.C. § 3-509 (1978). Nearly identical contract language is found in the subsection establishing the article 4 transfer warranty. *Id.* § 4-207(2).

[209] *See supra* note 208. If any indorser is found liable under contract, he might then

of action, the loss will eventually be shifted back up the chain[210] and rest on the party who first dealt with the forger, who will then be left with a cause of action against the forger.

Therefore, in most forgery cases the loss will be shifted back up the chain through the collecting banks and will eventually rest with the forger or the party who first dealt with the forger. This result will occur when the check bears a forged indorsement, or when a check bears a forged drawer's signature and is not paid. Only in the case in which a forged drawer's signature is paid will the loss remain with the payor bank.

B. *Allocation of Loss Under the N.P.C.*

Because the overriding rule of section 204 transmission liability will remove nearly all incentive from the payor bank to verify its customer's signature,[211] cases in which the payor bank detects the forgery and dishonors or returns the item will certainly be rare under the N.P.C. There will be little incentive to be careful as long as the payor bank can establish good faith, a relatively easy task. Under the N.P.C., the payor bank will usually dishonor an unauthorized item only accidently or for some reason not related to the forgery, such as insufficient funds.

Section 204 also governs liability and loss allocation in the case of dishonored, unauthorized items, and the section's application, again, passes the loss up the chain. As a practical matter, this liability would ultimately rest with the party who

---

bring a warranty action against prior parties. Any "transferee . . . who takes the item in good faith" can be sued for breach of warranty. U.C.C. § 4-207(2) (1978); *see also id.* § 3-417(2). Similarly, a party found liable for breach of warranty might bring a contract as well as a warranty action against prior parties. The promise in § 3-414(1) of the U.C.C. is made to "the holder or any subsequent indorser who takes it up . . . ." *Id.* § 3-414(1).

[210] It should be emphasized that not every party who participates in the check collection process will be liable for breach of warranty in all cases, nor will every party be liable under contract in all cases. Liability under either theory depends upon a number of conditions. For example, transfer warranty liability under both article 3 and article 4 depends on whether the potential defendant in a warranty action received some consideration or settlement for the check. Liability for indorsements, at least under article 3, depends on whether the potential defendant indorsed the check. *See supra* note 208. Nevertheless, in the majority of cases, the loss can be shifted back up the chain either through contract liability, or warranty liability, or both.

[211] The only incentive, other than the good faith requirement, is indirectly supplied by the payor bank's liability for wrongful dishonor under § 101 of the N.P.C. *See supra* note 180 and accompanying text. *But see supra* note 181.

first dealt with the forger, a result identical to that under the U.C.C.[212]

Whether grocery stores, depositary banks, or other parties who deal with forgers can pass the loss on to the negligent customer is questionable. Section 204(6) requires the payor bank or other collecting bank to assert any valid defense against the customer.[213] In the case of a dishonored, unauthorized item, however, the customer would not have brought an action against the payor bank, so that there would be no opportunity for any party to assert any defense, negligence-based or otherwise, against that customer. Loss allocation in this type of case can be simply addressed through section 100 of the N.P.C.[214] This section sets forth the contract of the drawer and is essentially the same as the drawer's contract under section 3-413(2) of the U.C.C.[215] Section 100 of the N.P.C. provides that this contract liability exists only with authorized orders; but recall that under section 202 of the N.P.C., the negligent or strictly liable customer will be precluded from asserting that an item is unauthorized.[216] Unless the collecting bank or other prior transferor is negligent or fails to act in accordance with reasonable commercial standards, it should be able to pass the loss on to the negligent customer. If the customer is not negligent or if the party who dealt with the forger failed to act in accordance with reasonable commercial standards, the loss will remain with that party.[217]

IV. Conclusion

With respect to the allocation of loss between the customer (drawer) and the payor bank in cases of a forged drawer's

[212] *See supra* notes 201–10 and accompanying text.

[213] N.P.C., *supra* note 5, § 204(6).

[214] Section 100 provides that: "The drawer of an authorized [item] agrees with the [holder] that upon dishonor of a [draft or check] by the drawee or any subsequent [collecting bank], and any necessary notice of protest or dishonor, the drawer will pay the amount of the authorized [item] to the [holder]." *Id.* § 100. Comment 2 to this section states that: "The obligation of the drawer is limited to authorized [items]. If the [item] is altered, forged, or transferred without authorization, it becomes an unauthorized [item] and the drawer is under no obligation to pay it." *Id.* § 100 comment 2.

[215] The U.C.C. provides that: "The drawer engages that upon dishonor of the draft and any necessary notice of dishonor or protest he will pay the amount of the draft to the holder or to any indorser who takes it up . . . ." U.C.C. § 3-413(2) (1978). *Cf.* N.P.C., *supra* note 5, § 100; for the text of N.P.C. § 100, see *supra* note 214.

[216] *See supra* notes 74–86 and accompanying text.

[217] The collecting bank or other prior transferor could pass the loss to the customer by bringing an action, based on the contract under § 100 of the N.P.C., against the customer whose signature was forged. In other words, it would bring suit to enforce the drawer's contract.

signature, the two-tiered approach of N.P.C. is a needed improvement over the current rules under the U.C.C. The N.P.C. approach is soundly based on a recognition of the economic realities of litigating disputes involving small amounts. Consequently, the protections given to consumers under section 200(2) of the N.P.C. should be extended to business customers. The $50 liability provision, however, is inappropriate and unnecessary because it will be imposed in cases in which customers are not negligent as well as in cases in which customers are negligent, and because section 200 of the N.P.C. contains other provisions that will deter negligent behavior.

The abolition of the rule of *Price v. Neal* is unjustified. In its place the drafters have proposed a rule of transmission liability, embodied in section 204, that would allow the payor bank in nearly every instance of a forged drawer's signature to shift any loss onto collecting banks and other prior transferors. This shifting would, in effect, insulate the payor bank from liability in virtually every type of forgery. In our view, this change is a substantial and unwarranted departure from the current loss allocation scheme under the U.C.C. While modern banking practices, such as check truncation, may require some changes in the rules governing the allocation of loss in cases of forged drawers' signatures, such changes do not require the significant changes contained in the current draft of the N.P.C. Thus, this draft of the N.P.C. should not be adopted unless the rules for allocating loss in cases of forged drawers' signatures are reconsidered and substantially revised.

NOTE

OF THINGS TO COME—THE ACTUAL IMPACT OF *HERBERT v. LANDO* AND A PROPOSED NATIONAL CORRECTION STATUTE

DALE M. CENDALI*

The United States Supreme Court, in New York Times v. Sullivan, *established a special substantive standard requiring "actual malice" in public figure libel cases.* Herbert v. Lando *is the forerunner of subsequent cases in which the Supreme Court declined to grant special procedural protections in libel litigation. While initially* Herbert *created a fear of excessive intrusion into the editorial process, an analysis of the case reveals that the increased costs of libel litigation have in fact had a greater chilling effect, both on the media's prepublication behavior and its subsequent litigation tactics.*

In this Note, Ms. Cendali argues that currently proposed solutions to the problems of libel litigation are inadequate to combat successfully the concerns of intrusiveness and costliness in modern libel law. She instead proposes a national "Correction Statute," which encourages settlement, discourages costly and intrusive litigation, and seeks to promote the ascertainment of truth.

When the Supreme Court decided *New York Times v. Sullivan*[1] in 1964 and established the rule that a public figure plaintiff in a libel action must prove that a media defendant acted with "actual malice,"[2] the case was heralded as "unquestionably the greatest victory won by defendants in the modern history of the law of torts"[3] and "an occasion for dancing in the

---

* Associate, Fried, Frank, Harris, Shriver & Jacobson, New York, N.Y. B.A., Yale University, 1981; J.D., Harvard University, 1984. The author thanks Professor Arthur Miller of the Harvard Law School for his advice during the development of this manuscript.

[1] 376 U.S. 254 (1964).

[2] *Id.* "Actual malice" is a term of art defined as publishing a misstatement of fact "with knowledge that it was false or with reckless disregard of whether it was false or not." *Id.* at 279–80. *See also* Cantrell v. Forest City Publishing Co., 419 U.S. 245, 251–52 (1974) ("actual malice" is distinct from common law "malice" or ill will); St. Amant v. Thompson, 390 U.S. 727, 731 (1968) ("reckless disregard" requires showing that defendant had serious doubts as to the truth of publication).

[3] W. PROSSER, HANDBOOK OF THE LAW OF TORTS § 118, at 819 (4th ed. 1971); *see generally* Kalven, *The* New York Times *Case: A Note on The Central Meaning of the First Amendment*, 1964 SUP. CT. REV. 191; Berney, *Libel and the First Amendment—A New Constitutional Privilege*, 51 VA. L. REV. 1 (1965); Pedrick, *Freedom of the Press and the Law of Libel: The Modern Revised Translation*, 49 CORNELL L.Q. 581 (1964); *The Supreme Court, 1963 Term*, 78 HARV. L. REV. 143, 201–05 (1964).

streets."[4] Fifteen years later, when the Court held in *Herbert v. Lando*[5] that to establish actual malice, libel plaintiffs could conduct discovery into the thought processes of defendants, press advocates called the decision an "Orwellian invasion of the mind,"[6] and "judicial Agnewism."[7] The press feared that *Herbert* would "chill" the media in its exercise of its First Amendment rights either by allowing intrusive discovery into the thought processes of editors and reporters[8] or by increasing the number of libel suits designed to harass the media.[9]

Less publicized, but in fact more significant, was *Herbert*'s potential for increasing litigation costs in libel suits. By holding that the First Amendment did not afford media defendants any special procedural protection from discovery, the Supreme Court ensured that defending a libel suit would be an expensive and time-consuming task. Moreover, *Herbert* was the first of several Supreme Court cases that have guaranteed large costs in libel litigation. In *Keeton v. Hustler Magazine*[10] and *Calder v. Jones*,[11] the Court interpreted the jurisdictional minimum contacts test to permit a state to assert jurisdiction over a foreign publisher that circulated its publication within the state, thereby exposing publishers to litigation in distant and unfamiliar jurisdictions. To the extent that the dicta in *Hutchinson v. Proxmire*[12] encouraged courts to deny summary judgment for defendants, *Hutchinson* may also have prolonged litigation and increased its expense. Finally, while *Bose Corp. v. Consumers Union*[13] preserved the independent appellate review of the *Sullivan* standard's application by lower courts, it did so at the expense of

---

[4] Alexander Meiklejohn, *quoted* in Kaufman, *The Media and Juries*, L.A. Daily J., Nov. 9, 1982, at 4, col. 3.

[5] 441 U.S. 153 (1979).

[6] L.A. Times, Apr. 19, 1979, § 1, at 16, col. 1 (quoting Ralph Otwell, editor of the Chicago Sun Times); *see also* Miami Herald, Apr. 20, 1979, at 6-A ("Orwellian domain").

[7] Wash. Star, Apr. 20, 1979, at A-10.

[8] *See* Dale & Dale, *Full Court Press: The Imperial Judiciary vs. The Paranoid Press*, 7 PEPPERDINE L. REV. 241, 273 (1980), *citing* Keynote Address by Allen Neuharth, Am. Newspaper Publishers Ass'n, 93d Annual Convention of the Am. Newspaper Publishers Ass'n (Apr. 23, 1979) (*Herbert* decision "ruled that lawyers can rummage through reporters' and editors' minds"); *see also* L.A. Times, Apr. 19, 1979, at 16, col. 1.

[9] *See, e.g.*, EDITOR & PUBLISHER, May 5, 1979, at 24 ("public officials will use [the *Herbert*] decision to try to harass newspapers with expensive law suits").

[10] 104 S. Ct. 1473 (1984); *see infra* notes 65–72 and accompanying text.

[11] 104 S. Ct. 1482 (1984); *see infra* notes 73–79 and accompanying text.

[12] 443 U.S. 111 (1979); *see infra* notes 33–34 & 99–100 and accompanying text.

[13] 104 S. Ct. 1949 (1984); *see infra* notes 80–87 and accompanying text.

the increased costs associated with independent appellate review, which occurs after most litigation expenses have already been incurred.

This Note examines the actual impact of *Herbert* in light of these more recent decisions and concludes that the potential chilling effect of intrusive discovery was greatly exaggerated and that *Herbert*'s contribution to increased litigation costs is actually the more dangerous result for both plaintiffs and defendants. *Herbert* was the forerunner of a series of cases that refused to acknowledge unique procedural protections for media defendants. As this Note will discuss, this failure to grant procedural protections has had a substantive, collective impact on both the media's prepublication behavior and its subsequent litigation tactics. The Note suggests that the press should shift its criticism of current libel law to advocate practical legislative solutions to the special problems associated with libel litigation and to emphasize the media's shared interest with both courts and plaintiffs in promoting a system for the "just, speedy, and inexpensive determination of every action."[14] The Note's analysis of the potential solutions to the problems of intrusiveness and costliness shows that the solutions proposed to date are inadequate. The Note concludes by proposing a national "Correction Statute," which would limit the press's liability while encouraging the correction of published factual inaccuracies and the avoidance of protracted litigation.

I. Constitutional Context of *Herbert v. Lando*

New York Times v. Sullivan marked the first time in its 170-year history that the Supreme Court reviewed a state judgment in a civil libel suit under the First Amendment.[15] The police commissioner of Montgomery, Alabama brought an action alleging that he had been libeled by an advertisement in the *New*

---

[14] Fed. R. Civ. P. 1.

[15] 376 U.S. at 300 n.3. (Goldberg, J., concurring). The First Amendment free speech and press clauses apply to the states through the Fourteenth Amendment. *See, e.g.*, Gitlow v. New York, 268 U.S. 652, 666 (1925) (upheld New York's statute which punished those who advocated overthrowing organized government by unlawful means); Stronberg v. California, 283 U.S. 359, 368 (1931) (declared unconstitutional that part of California's statute that prohibited the display of a red flag as a sign of government opposition).

York Times entitled "Heed Their Rising Voices."[16] The advertisement described conditions in the South and sought contributions for the civil rights movement. Although it did not mention Sullivan by name, Sullivan felt that the advertisement defamed him.[17] He sued the *New York Times* for $500,000 in an Alabama state court and received a jury award for that amount. The Alabama Supreme Court affirmed.[18]

At the time that *Sullivan* was tried, there were libel suits for claims totaling more than $300 million pending against news organizations nationwide.[19] Other Montgomery officials and the governor of Alabama had sued the *New York Times* over the same advertisement, and a jury had awarded another $500,000 in damages.[20] The *Times* had a circulation of roughly 650,000; yet on the day that the advertisement appeared, approximately 394 copies reached Alabama, and of those only thirty-five were distributed in Montgomery County.[21]

The United States Supreme Court overturned the lower court's verdict in favor of Sullivan. Justice Brennan, writing for the Court, concluded that the First Amendment shielded even defamatory falsehoods about a public official relating to his official capacity unless they were published with "actual malice."[22] Proof of "actual malice" requires a showing that the statement was made "with knowledge that it was false or with reckless disregard of whether it was false or not."[23]

---

[16] *See* N.Y. Times, Mar. 29, 1960, at 25 (reproduced as Appendix to the opinion of the Court in *Sullivan*, 376 U.S. at 292).

[17] The fact that the advertisement did not mention Sullivan by name, together with the relatively minor inaccuracies in the piece that Sullivan cited, raises serious concerns that, without special protections, such publications as the *Times* could be limitlessly exposed to liability for minor discrepancies. For example, the Court in *Sullivan* pointed out that nine students were expelled by the State Board of Education, "not for leading [a] demonstration [as the *Times* had indicated] . . . but for demanding service at a lunch counter." *Sullivan*, 376 U.S. at 259. Furthermore, it noted, "[n]ot the entire student body, but most of it, had protested the expulsion, not by refusing to register, but by boycotting classes on a single day." *Id.*

[18] New York Times Co. v. Sullivan, 273 Ala. 656, 144 So. 2d 25 (1962). At the time, this was the largest judgment in Alabama history. C. LAWHORNE, DEFAMATION AND PUBLIC OFFICIALS 216 (1971).

[19] Garbus, *The New Challenge to Press Freedom*, N.Y. Times, Jan. 29, 1984, § 6 (Magazine), at 38. Five suits of claims totaling $1.7 million were filed concerning CBS's coverage of civil rights protests. *Id.*

[20] *Sullivan*, 376 U.S. at 278 n.18.

[21] *Id.* at 260 n.3. The *Times* had waived any objections to personal jurisdiction by making what was deemed by the Alabama trial court and the Alabama Supreme Court to be a general appearance. *Id.* at 264 n.4.

[22] *Id.* at 279–80.

[23] *Id.* at 280.

Sullivan remains the well-settled constitutional standard.[24] Subsequent decisions extended *Sullivan* to criminal libel,[25] to unelected public officials,[26] and to public figures who, while not government employees, had thrust themselves into public controversy.[27]

In the 1970's, however, the Court began to indicate that in balancing the law of defamation with the protections of the First Amendment, the proper accommodation may sometimes favor the plaintiff's interests. The first sign was in 1971, when only a plurality of the Court wanted to extend the *Sullivan* rule to actions brought by public figures about matters of "public or general interest."[28] Then, in 1974, *Gertz v. Robert Welch, Inc.*[29] established that nonpublic figure plaintiffs bringing defamation actions about nonpublic issues did not have to meet the actual malice standard, but only a lesser showing of fault.[30]

The Court then began to limit its definition of a public figure. For example, the Court held in *Time, Inc. v. Firestone*[31] that a sensational divorce may have been a "cause celebre" but it was not a "public controversy," and that the estranged wife was not a public figure because she did not voluntarily thrust herself to the forefront of a public issue to influence its outcome.[32] Simi-

---

[24] *See, e.g.*, Herbert v. Lando, 441 U.S. 153, 169 (1979) (*Sullivan* doctrine "represented a major development . . . [that] has been repeatedly affirmed as the appropriate First Amendment standard applicable" in public figure libel suits); Bose Corp. v. Consumers Union, 104 S. Ct. 1949, 1959 (1984) (*Sullivan* doctrine affects standard for appellate review).

[25] Garrison v. Louisiana, 379 U.S. 64, 77–78 (1964) (Court struck down state criminal libel statute because it did not adhere to the *Sullivan* interpretation of "actual malice").

[26] Rosenblatt v. Baer, 383 U.S. 75, 84–86 (1966) (unelected county recreation supervisor is a public official under the *Sullivan* rule).

[27] Curtis Publishing Co. v. Butts, 388 U.S. 130, 155 (1967) (well-known university athletic director is a public figure who must meet "actual malice" test).

[28] Rosenbloom v. Metromedia, Inc., 403 U.S. 29, 43–44 (1971) (*Sullivan* rule applied where report of plaintiff's arrest for distribution of obscene publication concerned a matter of public or general interest).

[29] 418 U.S. 323 (1974). Plaintiff was an attorney representing a policeman convicted of murder in civil rights suits. The attorney was reported in defendant's newspaper to have been part of a Communist plot against the local police force, to have participated in a "frame-up" of the policeman, and to have had a criminal record. *Id.* at 325–26. The jury found for the plaintiff. The district court entered a judgment n.o.v. for the defendant on the ground that the *Sullivan* public figure standard applied. The court of appeals affirmed. *Id.* at 329–32. The Supreme Court held that Gertz was not a public figure because his role in the controversy was not one of a voluntary exposure to the risk of defamatory falsehoods. *Id.* at 351–52. *See generally* Smith, *The Rising Tide of Libel Litigation*, 44 MONT. L. REV. 71 (1983).

[30] *Id.* at 347.

[31] 424 U.S. 448 (1976).

[32] *Id.* at 453–54. Mrs. Firestone did, however, hold periodic press conferences, but because these could not have influenced the legal outcome of the divorce, holding the

larly, in *Hutchinson v. Proxmire*,[33] the Court held that a scientist did not thrust himself into a public controversy merely by applying for and accepting a federal research grant.[34] Finally, the plaintiff in *Wolston v. Reader's Digest Association*[35] did not have to meet the *Sullivan* test because he was "dragged unwillingly" into a controversy through a subpoena to testify before a grand jury in an espionage investigation.[36]

After *Sullivan* was decided, the Court concentrated most of its efforts on defining the substantive constitutional standard in defamation cases. In 1979, however, *Herbert v. Lando* became the first in a series of cases to address the relationship between substantive protections and procedural rules.

Lieutenant Colonel Anthony B. Herbert, an army officer stationed in Vietnam during the late 1960's, contended that he had observed many war crimes and atrocities committed by American troops. He reported these acts to his superior officers, who took no action. Ultimately, Herbert brought formal charges against these officers, which led to a full army investigation. He later wrote a book about his experiences.[37]

On February 4, 1973, CBS televised a *60 Minutes* program that Herbert believed portrayed him as a liar and an opportunist who had committed brutalities and who had tried to use the alleged conduct of his superior officers as an excuse for his own relief from command. The producer of the segment, Barry Lando, subsequently wrote an article about Herbert in *Atlantic Monthly* magazine that reiterated the broadcast's theme. Herbert brought a libel suit in federal district court.[38] He alleged $44,725,000 in damages for injury to his reputation and impairment of his book, *Soldier*, as literary property.[39]

Herbert conducted extensive pretrial discovery. Lando alone was deposed in twenty-six separate sessions over one year,

---

news conferences did not constitute the requisite attempt to affect a result. *Id.* at 454–55 n.3.

[33] 443 U.S. 111 (1979).

[34] *Id.* at 135. *Hutchinson* is also important because Chief Justice Burger implied that summary judgment is an inappropriate device to decide such state of mind issues as actual malice and should thus be avoided in libel suits. *Id.* at 120 n.9. *See infra* notes 160–169 and accompanying text.

[35] 443 U.S. 157 (1979).

[36] *Id.* at 166–67.

[37] Herbert v. Lando, 73 F.R.D. 387, 391 (S.D.N.Y.), *remanded* 568 F.2d 974 (2d Cir. 1977), *rev'd* 441 U.S. 153 (1979).

[38] *Id.*

[39] Herbert v. Lando, 568 F.2d 974, 982 (2d Cir. 1977), *rev'd* 441 U.S. 153 (1979).

generating 2,903 pages of transcript and 240 exhibits.[40] He answered "innumerable questions about what he knew, or had seen; whom he interviewed; . . . and the form and frequency of his communications with sources."[41] Lando "painstakingly deciphered and explained" his often "cryptic" interview notes.[42] Numerous disputes arose, which the parties managed to resolve; however, there was one area of questioning upon which the parties could not agree. Herbert wanted Lando to answer questions concerning the producer's thought processes as he put the broadcast together.[43]

Judge Haight recognized that defining the proper boundaries for pretrial discovery in a public figure defamation action was a question of first impression.[44] He concluded that, like any plaintiff, a plaintiff in a libel action is entitled to a liberal interpretation of the discovery rules, and that although Herbert's inquiry might not lead to admissible evidence, under the normal standard of Federal Rule 26(b)(1), admissibility is not a prerequisite to production.[45] Because actual malice requires proof that a publisher "in fact entertained serious doubts as to the truth of his publication,"[46] Judge Haight was concerned that if he narrowed the scope of discovery "into areas which in the nature of the case lie solely with the defendant, then the law in effect [would provide] an arras behind which malicious publication

---

[40] *Id.*
[41] *Id.*
[42] *Id.* at n.18.
[43] The district court summarized these questions into seven categories:

1. Lando's conclusions during his research and investigation regarding people or leads to be pursued, or not to be pursued, in connection with the "60 Minutes" segment and the *Atlantic Monthly* article;
2. Lando's conclusions about facts imparted by interviewees and his state of mind with respect to veracity of persons interviewed;
3. The basis for conclusions where Lando testified that he did reach a conclusion with respect to persons, information or events;
4. Conversations between Lando and Wallace about matter to be included or excluded from the broadcast publication;
5. Lando's intentions as manifested by the decision to include or exclude material;
6. Conversations between Lando and source persons subsequent to the inception of this action;
7. Lando's activities as well as conversations between Lando, Wallace and/or other CBS employees concerning Herbert or the "60 Minutes" segment between broadcast and publication of the *Atlantic Monthly* article.

Herbert, 73 F.R.D. at 392.
[44] *Id.* at 390–91, 393.
[45] *Id.* at 394–95; *see* FED. R. CIV. P. 26(b)(1).
[46] *Herbert*, 73 F.R.D. at 393, *citing* St. Amant v. Thompson, 390 U.S. 727, 731 (1968).

could go undetected and unpunished."[47] He stated that "[n]othing in the First Amendment requires such a result."[48]

A sharply divided panel of the Court of Appeals for the Second Circuit reversed.[49] Chief Judge Kaufman wrote the opinion for the court in sweeping terms. He discussed *Miami Herald Publishing Co. v. Tornillo*,[50] in which the Supreme Court held that a newspaper could not be compelled by a state statute to accept editorial replies, and *Columbia Broadcasting System v. Democratic National Committee*,[51] which held that broadcasters were not required by the First Amendment to accept paid political advertisements.[52] Kaufman concluded that these cases argued that full and candid discussion within the newsroom had to be encouraged and that discovery into the editorial process was constitutionally protected.[53] He suggested that "we must permit only those procedures in libel actions which least conflict with the principle that debate on public issues should be robust and uninhibited."[54]

In a decision that produced five different opinions, the Supreme Court reversed six to three.[55] Justice White, writing for the majority, rejected the appellate court's analysis, finding it "incredible to believe" that the Supreme Court in *CBS* and

---

[47] *Id.* at 394.

[48] *Id.*

[49] Herbert v. Lando, 568 F.2d 974 (2d Cir. 1977), *rev'd* 441 U. S. 153 (1979).

[50] 418 U.S. 241 (1974).

[51] 412 U.S. 94 (1973).

[52] *Herbert*, 568 F.2d. at 978–79.

[53] *Id.* at 979.

[54] *Id.* at 980. Judge Oakes based his concurrence on his belief that the First Amendment provided special procedural protections for the press and that therefore such decisions as *Tornillo* and *CBS* "require[d] a constriction of the normal discovery rules" for libel litigation. *Id.* at 991 (Oakes, J., concurring). Judge Meskill, in his dissent, rejected the idea that discovery into an editor's state of mind would chill First Amendment activity "to any greater extent than it is already being chilled as a result of the very review permitted by *New York Times v. Sullivan.*" *Id.* at 995–96 (Meskill, J., dissenting).

[55] Herbert v. Lando, 441 U.S. 153 (1979). In addition to the opinion of the Court, four justices wrote separately. In his concurring opinion, Justice Powell added that he believed that a district court in applying the discovery rules should consider First Amendment interests and protect against discovery abuse. *Id.* at 178–79 (Powell, J., concurring). Justice Brennan, in dissent, went further and suggested a formal bifurcated discovery procedure whereby a plaintiff would first have to make a prima facie demonstration of falsehood before the trial judge would allow discovery into the editorial process. *Id.* at 181 (Brennan, J., dissenting in part). Justice Stewart's dissent stressed relevance, claiming that because a libel suit concerns what "was in fact published . . . [w]hat was *not* published has nothing to do with the case" and can have no bearing on the publisher's knowledge of falsity. *Id.* at 200 (Stewart, J., dissenting) (emphasis in original). Finally, Justice Marshall also advocated a special "strict standard of relevance" in libel actions. *Id.* at 206 (Marshall, J., dissenting).

Tornillo "silently effected a substantial contraction of the rights preserved to defamation plaintiffs."[56] He cited eighteen cases in which inquiry by plaintiffs into the defendants' states of mind was permitted without any constitutional objection,[57] and he cited twenty-three cases in which the defendants introduced evidence to either establish their own good faith or lack of malice.[58] Justice White stressed that the actual malice test was the firmly established constitutional doctrine that had been "repeatedly affirmed as the appropriate First Amendment standard in libel actions";[59] and although this standard provided protection, he concluded that "[i]nevitably, unless liability is to be

---

[56] *Id.* at 168.

[57] *Id.* at 165–66 n.15, *citing* Johnson Publishing Co. v. Davis, 271 Ala. 474, 124 So. 2d 441 (1960); Freeman v. Mills, 97 Cal. App. 2d 161, 217 P.2d 687 (1950); Scott v. Times-Mirror Co., 181 Cal. 345, 184 P. 672 (1919); Sandora v. Times Co., 113 Conn. 574, 155 A. 819 (1931); Rice v. Simmons, 2 Del. 309 (1837); Western Union Telegraph Co. v. Vickers, 71 Ga. App. 204, 30 S.E.2d 440 (1944); Cook v. East Shore Newspapers, 327 Ill. App. 559, 64 N.E.2d 751 (1945); Berger v. Freeman Tribune Publishing Co., 132 Iowa 290, 109 N.W. 784 (1906); Thompson v. Globe Newspaper Co., 279 Mass. 176, 181 N.E. 249 (1932); Conroy v. Fall River Herald News Co., 306 Mass. 488, 28 N.E.2d 729 (1940); Cyrowski v. Polish-American Pub. Co., 196 Mich. 648, 163 N.W. 58 (1917); Friedell v. Blakely Printing Co., 163 Minn. 226, 203 N.W. 974 (1925); Cook v. Globe Printing Co., 227 Mo. 471, 127 S.W. 332 (1910); Butler v. Gazette Co., 119 A.D. 767, 104 N.Y.S. 637 (1907); Briggs v. Byrd, 34 N.C. 377 (1851); McBurney v. Times Publishing Co., 93 R.I. 331, 175 A.2d 170 (1961); Lancour v. Herald & Globe Ass'n, 112 Vt. 471, 28 A.2d 396 (1942); Farrar v. Tribune Publishing Co., 57 Wash. 2d 549, 358 P.2d 792 (1961).

[58] *Id.* at 165–67 n.15, *citing* Bohan v. Record Pub. Co., 1 Cal. App. 429, 82 P. 634 (1905); Hearne v. De Young, 119 Cal. 670, 52 P. 150 (1898); Ballinger v. Democrat Co., 203 Iowa 1095, 212 N.W. 557 (1927); Snyder v. Tribune Co., 161 Iowa 671, 143 N.W. 519 (1913); Courier-Journal Co. v. Phillips, 142 Ky. 372, 134 S.W. 446 (1911); Conner v. Standard Pub. Co., 183 Mass. 474, 67 N.E. 596 (1903); Davis v. Marxhausen, 103 Mich. 315, 61 N.W. 504 (1894); Julian v. Kansas City Star Co., 209 Mo. 35, 107 S.W. 496 (1908); Paxton v. Woodward, 31 Mont. 195, 78 P. 215 (1904); Las Vegas Sun, Inc. v. Franklin, 74 Nev. 282, 329 P.2d 867 (1958); Lindsey v. Evening Journal Ass'n, 10 N.J. Misc. 1275, 163 A. 245 (1932); Kohn v. P & D Publishing Co., 169 A.D. 580, 155 N.Y.S. 455 (1915); Hains v. New York Evening Journal, 240 N.Y.S. 734 (Sup. Ct. 1930); Goodrow v. Malone Telegram, Inc., 235 A.D. 3, 255 N.Y.S. 812 (1932); Goodrow v. Press Co., 233 A.D. 41, 251 N.Y.S. 364 (1931); Kehoe v. New York Tribune, 229 A.D. 220, 241 N.Y.S. 676 (1930); Varvaro v. American Agriculturist, Inc., 222 A.D. 213, 225 N.Y.S. 564 (1927); Van Arsdale v. Time, Inc., 35 N.Y.S.2d 951 (Sup Ct. 1942), *aff'd*, 265 A.D. 919, 39 N.Y.S.2d 413 (1942); Weichbrodt v. New York Evening Journal, 11 N.Y.S.2d 112 (Sup. Ct. 1939); Cleveland Leader Printing Co. v. Nethersole, 84 Ohio St. 118, 95 N.E. 735 (1911); Cobb v. Oklahoma Publishing Co., 42 Okla. 314, 140 P. 1079 (1914); Times Pub. Co. v. Ray, 1 S.W.2d 471 (Tex. Civ. App. 1927), *aff'd*, 12 S.W.2d 165 (1929); Pfister v. Milwaukee Free Press Co., 139 Wis. 627, 121 N.W. 938 (1909).

[59] *Id.* at 169, *citing* Time, Inc. v. Firestone, 424 U.S. 448 (1976) (where plaintiff did not occupy a role of societal prominence, public figure standard of reckless disregard does not apply); Gertz v. Robert Welch, Inc., 418 U.S. 323, 342 (1974) (*Sullivan* "actual malice" standard does not apply to private plaintiff); St. Amant v. Thompson, 390 U.S. 727 (1968) (public figure plaintiff required to prove "reckless disregard"); Curtis Publishing Co. v. Butts, 388 U.S. 130 (1967) (*Sullivan* standard applies to defamation actions by "public figures" as well as "public officials").

completely foreclosed, the thoughts and editorial processes of the alleged defamer would be open to examination."[60]

While much of the initial commentary on *Herbert* focused on the need to have liberal discovery rules in order to prove actual malice,[61] the potentially most significant part of the decision received little attention.[62] The Court recognized that "it would not be surprising" if plaintiffs after *Herbert* conducted more discovery and if "the costs and other burdens of this kind of litigation would escalate and become much more troublesome for both plaintiffs and defendants."[63] It declined, however, to allow the press any special procedural protections. "Mushrooming litigation costs" due in part to pretrial discovery, the Court noted, "are not peculiar to the libel and slander area."[64]

In recent cases, the majority's failure to recognize the need for unique procedural protection has magnified the importance of the *Herbert* discovery rule. Two such cases are *Keeton v. Hustler Magazine*[65] and *Calder v. Jones*.[66] The plaintiff in *Keeton* was a New York resident bringing a diversity action in New Hampshire against Hustler, Inc., an Ohio corporation with its principal place of business in California. The defendant's only connection with New Hampshire was that 10,000 to 15,000 copies of *Hustler Magazine* were circulated there each month. In most jurisdictions Keeton's claim would have been barred by the statute of limitations; however, New Hampshire had an unusually lengthy six-year limitations period.[67] Under the doc-

---

[60] *Id.* at 160. The Court also seemed concerned about the difficulty in defining a privilege for the editorial process and whether it would apply only to media defendants. *See id.* at 170. The Court was also unsure why Lando was willing to testify as to what he "knew" and what he had "learned" but not what he "believed," *see id.*; *see also* 4 U.S. Supreme Court, Oral Arguments, 42–44 (Oct. Term 1978) (It was questioned whether the privilege would apply to both a media reporter and a private citizen who made a bad statement about someone).

[61] *See, e.g.*, Dale & Dale, *supra* note 8; Oakes, *Proof of Actual Malice in Defamation Actions: An Unsettled Dilemma*, 7 HOFSTRA L. REV. 655 (1979); Los Angeles Times, Apr. 19, 1979, § 1, at 16, col. 1; Miami Herald, Apr. 20, 1979, at 6-A.

[62] *See, e.g.*, Friedenthal, Herbert v. Lando: *A Note on Discovery*, 31 STAN. L. REV. 1059 (1979) (*Herbert* decision used discovery rules as a scapegoat to achieve the desired substantive result—denial of a media privilege in defamation cases).

[63] *Herbert*, 441 U.S. at 176.

[64] *Id.*

[65] 104 S. Ct. 1473 (1984); *see infra* notes 67–72 and accompanying text.

[66] 104 S. Ct. 1482 (1984); *see infra* notes 73–79 and accompanying text.

[67] N.H. REV. STAT. ANN. § 508:4 (1983). Effective August 1981, the limitation period was shortened to 3 years. *Id.* Because Keeton's injury occured prior to Aug. 1981, the old six-year limit was applicable. *Id.* at 1477.

trine of the single publication rule,[68] Keeton could recover damages incurred in all jurisdictions, regardless of their respective state statutes of limitations, if she could bring suit in New Hampshire. The only question was whether a New Hampshire court could exercise personal jurisdiction over Hustler.[69]

Justice Rehnquist, writing for a unanimous Court, held that it could. The Court applied the minimum contacts test of *World-Wide Volkswagen Corp. v. Woodson*[70] and *International Shoe Corp. v. Washington*.[71] More important, however, was what the Court did not consider. Justice Rehnquist wrote, "[W]e reject categorically the suggestion that invisible radiations from the First Amendment may defeat jurisdiction otherwise proper under the Due Process Clause."[72]

Justice Rehnquist developed this view in *Calder v. Jones*,[73] a case decided the same day as *Keeton*. Once again, he wrote for a unanimous Court. The plaintiff, Shirley Jones, was a California resident allegedly defamed by an article in the *National Enquirer*, written by defendant South and edited by defendant Calder. She brought suit in the California Superior Court and served defendants with process in Florida, their place of residence. South had made frequent trips and phone calls to California, but Calder had only been to California twice—once on a pleasure trip and once to testify in an unrelated trial. The *National Enquirer* did not contest personal jurisdiction, but Calder and South did.[74]

The Court held that the exercise of jurisdiction was proper under *World-Wide Volkswagen Corp. v. Woodson*, because defendants knew that the article "would have a potentially devastating impact," the brunt of which would be felt in California, the state in which Jones lived and where the *Enquirer* had its

---

[68] The single publication rule states that in libel actions "(a) only one action for damages can be maintained; [and] (b) all damages suffered in all jurisdictions can be recovered in the one action." RESTATEMENT (SECOND) OF TORTS § 577A(4)(1974).

[69] Keeton v. Hustler Magazine, Inc., 104 S. Ct. 1473 (1984).

[70] 444 U.S. 286, 297–98 (1980).

[71] 326 U.S. 310, 317 (1945).

[72] *Keeton*, 104 S. Ct. at 148 n.12. Even Justice Brennan accepted Justice Rehnquist's principle of not allowing special procedural rules in defamation cases. Justice Brennan devoted his concurrence to the point that only the liberty interests of the defendant and not state sovereignty concerns, should be considered in assessing personal jurisdiction. *Id.* at 1482 (Brennan, J., concurring) (discussing Insurance Corp v. Compagnie des Bauxites, 456 U.S. 694, 702–03 n.10 (1982)).

[73] 104 S. Ct. 1482 (1984).

[74] *Id.*

largest circulation.[75] Calder and South could "reasonably anticipate being haled into court there."[76]

The Court once again rejected the suggestion that "First Amendment concerns enter into the jurisdictional analysis."[77] Justice Rehnquist noted that "the potential chill on protected First Amendment activity stemming from libel and defamation actions is already taken into account in the constitutional limitations on the substantive law."[78] The Court cited *Herbert v. Lando* and reiterated, "We have already declined in other contexts to grant special procedural protections to defendants in libel and defamation actions in addition to the constitutional protections embodied in the substantive laws."[79]

Bose Corp. v. Consumers Union[80] was another decision highlighting the Supreme Court's unwillingness to uphold special First Amendment protections unless those protections were already granted in *Sullivan*. The split decision, moreover, reveals that Justices Rehnquist and O'Connor were actually willing to cut back on *Sullivan* rather than to reaffirm its protections.[81]

Bose was a product disparagement action between a manufacturer, Bose Corporation, and Consumers Union, whose magazine *Consumer Reports* evaluated the plaintiff's stereo loudspeaker system. Bose objected to statements in the article about its system, including one that said that the sound of individual musical instruments tended to wander "about the room," instead of accurately characterizing its tendency to wander "along the wall" between the speakers.[82] Bose sued after Consumers Union refused to print a retraction. The District Court of Massachusetts ruled that the *Sullivan* "actual malice" standard applied

---

[75] *Id.* at 1487.

[76] *Id.*, *citing* World-Wide Volkswagen, 444 U.S. at 297; Kulko v. Superior Court, 436 U.S. 84, 97–98 (1978); Shaffer v. Heitner, 433 U.S. 186, 216 (1977).

[77] *Calder*, 104 S. Ct. at 1487.

[78] *Id.* at 1488.

[79] *Id.*

[80] 104 S. Ct. 1949 (1984).

[81] In a dissent joined by Justice O'Connor, Justice Rehnquist argued that "reckless disregard" was simply a finding on defendant's mens rea. *Id.* at 1968 (Rehnquist, J., dissenting). Justice Rehnquist questioned *Sullivan*'s reviewability standard on the ground that appellate courts were "ill-prepared to make in any context, including the First Amendment context" such findings. *Id.* Justice White did not agree that the entire issue of "reckless disregard" was merely a question of fact. He argued that only the "actual knowledge" component of the *Sullivan* malice standard was a question determined by historical fact and that only it was subject to the kind of review contemplated in Rule 52(a). *Id.* at 1967 (White, J., dissenting).

[82] *Id.* at 1953–54.

and held that plaintiff proved its claim of disparagement.[83] The First Circuit reversed, holding that the defendant was entitled to a de novo review of the record to see if the district court properly applied the constitutional standard under *Sullivan*.[84] After conducting such a review, the court held that Bose had failed to sustain its burden of proof.[85]

Justice Stevens, writing for the majority of the Supreme Court, held that the "clearly erroneous" standard of review in Federal Rule 52(a) did not prescribe the applicable standard of review in a case governed by *Sullivan*'s "actual malice" standard.[86] The basis for the decision was the established tradition of the *Sullivan* rule requiring independent appellate review. Justice Stevens noted that "[t]he requirement of independent appellate review reiterated in *New York Times v. Sullivan* is a rule of federal constitutional law" and that the Court had repeatedly held that an appellate court has an obligation to make such an independent examination.[87] The majority was unwilling to alter *Sullivan*'s protections in order to foster procedural uniformity. *Bose* can thus be distinguished from *Calder* and *Keeton* in that those cases involved new procedural protections rather than existing ones. Justices Rehnquist's and O'Connor's dissent in *Bose*, however, properly characterizes the case as a precarious affirmation of the status quo.

This line of cases highlights the Supreme Court's refusal to extend new unique procedural protections to libel defendants. As will be discussed, the immediate effect of these decisions has led not only to more intrusive discovery in cases involving public figures, but also to increased costs for media defendants who must finance protracted litigation and who are now vulnerable to suit in distant jurisdictions.[88] To evaluate the potential

---

[83] Bose Corp. v. Consumers Union, 508 F. Supp. 1249, 1277 (D. Mass. 1981), *rev'd* 692 F.2d 189 (1st Cir. 1982), *aff'd* 104 S. Ct. 1949 (1984).

[84] Bose Corp. v. Consumers Union, 692 F.2d 189, 197 (1st Cir. 1982), *aff'd* 104 S. Ct. 1949 (1984).

[85] *Bose*, 692 F.2d at 197.

[86] Bose Corp. v. Consumers Union, 104 S. Ct. 1949, 1959 (1984) (Justice Stevens's opinion was joined by Justices Brennan, Marshall, Blackmun and Powell; Chief Justice Burger concurred separately without opinion); *see* FED. R. CIV. P. 52(a).

[87] *Bose*, 104 S. Ct. at 1965.

[88] *Bose*, for example, does little to help lower the costs of litigation because most costs are incurred in the pretrial phase. The more significant chilling effect of *Herbert*—higher costs—remains unresolved despite this procedural protection. For an analysis of *Calder*'s and *Keeton*'s impact on procedural protections for defendants sued in distant and inconvenient jurisdictions, see Levine, *Preliminary Procedural Protection for the Press from Jurisdiction in Distant Forums After Calder and Keeton,* 1984 ARIZ. ST. L.

ramifications of these recent developments, it is particularly appropriate to re-examine *Herbert v. Lando*, the first case to draw the procedural/substantive distinction, to assess its actual impact on libel litigation.

II. INITIAL REACTION TO *Herbert v. Lando*

The media considers virtually all Supreme Court decisions involving the press to be important—"usually the most important decision of the day."[89] It is no surprise, therefore, that on the day it was decided, *Herbert* was reported on the network news.[90] The *New York Times* ran the story on the front page with an additional feature and a detailed summary of the Court's opinion on the inside.[91] Justice Brennan characterized the general reaction as "a virtually unprecedented outpouring of scathing criticism."[92] The president of the American Newspaper Publishers Association suggested that the "state of mind of the majority of the members of the Burger Court needs examination."[93] Commentators in trade publications immediately began formulating techniques to combat *Herbert*, such as answering all state of mind questions with "I can't remember."[94] Although there were those in the press who felt that *Herbert* would not have much of an impact,[95] they were in the decided minority.[96]

J. 459 (concluding that lower courts should continue to give defamation defendants the same protection given other defendants under the Due Process Clause and associated jurisdictional doctrines). The vulnerability of media defendants to suits in distant jurisdictions will have its greatest chilling effect on small publishers with marginal circulation in distant forums. *Id.* at 464–65.

[89] Franklin, *Reflections on* Herbert v. Lando, 31 STAN. L. REV. 1035, 1050 (1979) [hereinafter cited as Franklin, *Reflections*].

[90] *Id.*

[91] *Newsmen Dealt Blow on Defense of Suits for Libel*, N.Y. Times, Apr. 19, 1979, at A1, col. 2; *Editors Concerned by Court's "State of Mind" Decision*, N.Y. Times, Apr. 19, 1979, at B11, col. 4; *Excerpts From Opinions of the Supreme Court*, N.Y. Times, Apr. 19, 1979, at B10, col. 1.

[92] Address by Justice William J. Brennan, Jr., dedication of the S.I. Newhouse Center for Law and Justice (Oct. 17, 1979), *reprinted* in 32 RUTGERS L. REV. 173, 179 (1979) [hereinafter cited as Brennan, *Address*].

[93] EDITOR & PUBLISHER, Apr. 28, 1979, at 12 (statement of Allen Neuharth).

[94] *See, e.g.*, Brown, *Three Little Words*, EDITOR & PUBLISHER, Apr. 28, 1979, at 96; *see also* Franklin, *Reflections*, *supra* note 89, at 1048 n.77 (discussing the "I can't remember" tactic and its likelihood of avoiding a contempt charge but encouraging an adverse jury verdict).

[95] *See, e.g.*, Shaw, *Journalists Fear Impact of Court Rulings*, L.A. Times, Jan. 1, 1979, at 28, col. 1; EDITOR & PUBLISHER, May 5, 1979, at 24.

[96] *See generally* Dale & Dale, *supra* note 8; Franklin, *Reflections*, *supra* note 89.

The press directed much of its criticism at the Burger court.[97] Part of the venom can be explained by the Chief Justice's view that the institutional press is only entitled to the same constitutional rights under the free speech clause as any other person,[98] a view completely antagonistic to the press's call for a special constitutional privilege. Burger's statement in *Hutchinson v. Proxmire*, indicating that proof of actual malice does not "readily lend itself" to disposition by summary judgment,[99] has been popularly perceived as sounding the "death knell" for this efficient procedure in libel suits.[100] In addition, Chief Justice Burger is depicted in *The Brethren* as being personally responsible for the Court's general shift toward conservatism.[101] This climate of personal criticism does little to improve understanding between the press and the Court.

It is not uncommon for the press to react to opinions it dislikes.[102] It is unusual, however, for one member of the Supreme Court to respond to the press's attacks. At the opening of a law center dedicated to a prominent publisher, Justice Brennan, the author of the opinion in *Sullivan*, devoted his address to the relationship between the media and the judiciary. He quoted various newspapers and their vituperative indictments of the Court after *Herbert*.[103] He pointed out that several newspaper accounts of important court decisions, including *Herbert*, were incorrect.[104] He did not believe that *Herbert*'s injury to the press

[97] *See, e.g.*, EDITOR & PUBLISHER, July 7, 1979, at 9 ("[T]he Burger court is determined to unmake the Constitution."); Editorial, *Last Resort*, BROADCASTING, Apr. 23, 1979, at 90 ("[T]he Burger court has left little of the First Amendment to repudiate."); *see also* Dale & Dale, *supra* note 8, at 273.

[98] First Nat'l Bank of Boston v. Bellotti, 435 U.S. 765 (1978) (Burger, C.J., concurring). Chief Justice Burger argued that the press should not have special constitutional protection because, (1) the Framers did not intend to distinguish between the free speech rights of individuals and the free press rights of the media, *id.* at 798–99, and (2) deciding which organizations merit institutional press protection would be reminiscent of the "abhorred licensing system of Tudor and Stuart England." *Id.* at 801. For a general discussion of the licensing system, see L. LEVY, EMERGENCE OF A FREE PRESS 6–7 (1985).

[99] 443 U.S. at 120 n.9.

[100] Garbus, *supra* note 19, at 40; *see* MEDIA INSURANCE 28 (J. Lankenau ed. 1983) (citing *Hutchinson* as support for the proposition that "[t]he Burger Court has created a legal environment less favorable to the media").

[101] B. WOODWARD & S. ARMSTRONG, THE BRETHREN 11–13 (1979).

[102] *See, e.g.*, Dale & Dale, *supra* note 8, for a discussion of press reaction to Gannett v. DePasquale, 443 U.S. 368 (1979) ("Supreme Court saying that the judiciary is a private supreme club," EDITOR & PUBLISHER, July 7, 1979, at 9).

[103] Brennan, *Address*, *supra* note 92, at 179.

[104] *Id.* at 179–80. Two newspapers erroneously characterized *Herbert* as holding that truth would no longer be an absolute defense in libel suits. *Id.*, *citing* Times-Picayune (New Orleans), Apr. 20, 1979, § 1, at 18; Birmingham News, Apr. 19, 1979, at 12. Others

was of the magnitude to justify the firestorm of criticism that resulted.[105] Most of all, Justice Brennan noted a new mood of "acrimony" and a "complete absence of an enterprise" between the press and the Supreme Court.[106]

Justice Brennan attributed this schism to the press's apparent unwillingness to accept the distinction between the familiar "speech" model of the First Amendment, where "the press requires and is accorded" absolute protection,[107] and the "structural" model, which recognizes and balances the press's interests with other societal interests.[108] Justice Brennan observed that by insisting on the pure speech model, the press overlooks their structural protections, such as access rights.[109] He suggested that the press "tailor" its absolutist rhetoric to recognize competing societal concerns.[110]

In trying to explain the difference between the structural and speech models of the First Amendment, Justice Brennan revealed how polarized the positions of the press and the judiciary are with regard to First Amendment perspectives. To much of the judiciary, the press seems unable to understand subtleties in the law; reporters look to the result of the opinion without understanding its context.[111] In the rush for "hot news," stories

---

"read the opinion as reverting to the common law definition of malice." Brennan, *Address*, *supra* note 92, at 179–80, *citing* Wicker, *A Chilling Court*, N.Y. Times, Apr. 20, 1979, at A31, col. 1; Star-Ledger (Newark, N.J.), Apr. 20, 1979, at 26, col. 1; Atlanta Constitution, Apr. 19, 1979, at A-4; The Oregonian, Apr. 19, 1979, at B6.

[105] Brennan, *Address*, *supra* note 92, at 179.

[106] *Id.* at 174.

[107] *Id.* at 176. Brennan argued that this model commands the widest consensus, in that it abides to the "commitment we all feel to the right of self-expression." *Id.*

[108] *See generally* L. TRIBE, AMERICAN CONSTITUTIONAL LAW §§ 12-2 to -22, at 582–700 (1978) (analyzing several Supreme Court decisions in light of its "balancing of interests" and absolutist paradigms).

[109] Brennan, *Address*, supra note 92, at 176.

[110] *Id.* at 181.

[111] *See* Dale & Dale, *supra* note 8, at 305–06. The authors analyze the problem as follows:

> 1. The press generally, excluding the specialists who cover the Supreme Court, does not make a distinction between the "holding" of the Court and the opinions of the Justices. Frequently the press is confused, and misinterprets the opinions and fails to separate the specific holding from elaboration.
>
> 2. The press generally does not appreciate the utility and necessity of the Court's confining itself to the narrow question presented. It seems that the press is unable to understand why the Court does not go beyond its holdings to anticipate potential disputes that might later arise and give specific guidelines to the lower courts on how to follow the ruling.
>
> 3. The press generally does not understand the function of *obiter dictum* in an opinion.
>
> 4. The press is, for the most part, composed of generalists who are impatient

may be written that do not accurately reflect the opinion.[112] The press is self-absorbed with its First Amendment rights and neglects other important legal opinions.[113] Journalism schools appear to perpetuate the absolutist "win" or "lose" perspective.[114] *New York Times* columnist and law lecturer Anthony Lewis suggested that conferences on the media and the law "bring out the machismo in journalists"[115] to the point that one judge normally sympathetic to the press remarked at a conference:

> Where, ladies and gentlemen, do you think these great constitutional rights that you are so vehemently asserting, and in which you were so conspicuously wallowing yesterday, where do you think they came from? The stork didn't bring them. These came from the judges of this country, from these villains here sitting at the table.[116]

On the other hand, the president of the American Newspaper Publishers Association sees the judiciary as behaving like a "private supreme club."[117] Judges are said to be jealous of the press's constitutional protections and are reluctant to recognize

---

with legal niceties, technicalities, and legal jargon. The press tends to "see the forests" and resists the strictures that come from "looking at the trees."

5. The press is impatient with the form of Court opinions, which are written in dry and symbolic language with constant reference to legal precedents.

6. The press fears the impact of Court decisions, believing that the rulings are carved in stone and are only rarely modified, distinguished, or overturned. *Id.* The former practice of the Supreme Court delivering its opinions only on Mondays and their present practice of saving important cases until the end of the term have contributed to the problems the media has had in reporting these decisions. The rush often led to some significant opinions going unreported or being misunderstood. Franklin, *Reflections*, *supra* note 89, at 1052–53 n.93.

[112] Franklin, *Reflections*, *supra* note 89, at 1052. The author cites examples of the press telephoning law professors for opinions of recently decided cases that the professors had yet to read. Franklin suggests that the press's behavior results from a lack of understanding of court decisions rather than from bias. *Id.* at 1053 n.96.

[113] In Dalia v. United States, 441 U.S. 238 (1979), the Court held that it was constitutional for police officers with a warrant to enter premises secretly and to plant surveillance bugs. *Dalia* was decided the same day as *Herbert*, but it received little coverage. Franklin, *Reflections*, *supra* note 89, at 1050. One study found metropolitan newspapers regularly provide more coverage of press clause cases than speech clause cases. *Id.* at 1050 n.85. Nevertheless, individual media organizations are known to be reluctant to publish accounts of adverse libel rulings against them. *Id.* at 1058 n.107.

[114] Franklin, *Reflections*, *supra* note 89, at 1054.

[115] Lewis, *A Preferred Position for Journalism?*, 7 HOFSTRA L. REV. 595, 619 (1979).

[116] *Id.*, *citing* THE MEDIA AND THE LAW 36–37 (H. Simons & J. Califano eds. 1976). Under the rules of the conference, the book does not give the name of the judge it quotes.

[117] EDITOR & PUBLISHER, July 7, 1979, at 9.

the press's role in guarding constitutional rights.[118] This mutual distrust led to a furor over *Herbert v. Lando* that far exceeded its actual effect.

III. ACTUAL IMPACT OF *Herbert*: INTRUSIVENESS

A. *Discovery into the Editorial Process*

As of September 1, 1983, a majority of the states had yet to have a judicial opinion that applied *Herbert v. Lando*.[119] Even though numerous other jurisdictions have followed *Herbert* to compel discovery of the editorial process,[120] there has been no drastic change in libel litigation. For example, courts have denied discovery into the editorial processes of nonparty media,[121] have prohibited discovery on the grounds of relevance,[122] and have stayed discovery until plaintiffs made a showing that there was a genuine issue of fact regarding falsity.[123] *Herbert* has even been cited for the proposition that discovery should not be granted merely to satisfy curiosity or to serve a general public interest.[124] Contrary to the Supreme Court's disinclination to extend special constitutional protections to the press,[125] many

---

[118] *See* Dale & Dale, *supra* note 8, at 307.

[119] LIBEL DEFENSE RESOURCE CENTER 50-STATE SURVEY (H. Kaufman ed. 1983) [hereinafter cited as LRDC SURVEY].

[120] *See, e.g.*, Gadsden County Times v. Horne, 426 So. 2d 1234, 1240 (Fla. 1983); National Ass'n of Gov't Employees, Inc. v. Central Broadcasting Corp., 379 Mass. 220, 232–233, 396 N.E.2d 996, 1004 (1979), *cert. denied* 446 U.S. 935 (1980); Downing v. Monitor Publishing Co., 120 N.H. 383, 386, 415 A.2d 683, 685–86 (1980); *see generally* LRDC SURVEY, *supra* note 119 (review of state laws on discovery of editorial processes). In Boyles v. Mid-Florida Television Corp., 431 So. 2d 627, 639 (1983), the court ordered disclosure of the "reasons for the decisions that were made" when those discussions were between a media client and his lawyer. The court felt that the lawyer was giving business, not legal advice. *Id.*

[121] *See, e.g.*, *In re* Consumers Union, 495 F. Supp. 582, 588 (S.D.N.Y. 1980); *In re* Forbes Magazine, 494 F. Supp. 780, 781 (S.D.N.Y. 1980).

[122] Lancaster v. Daily Banner-News, 274 Ark. 145, 154, 622 S.W.2d 671, 676 (1981).

[123] Bruno & Stillman, Inc. v. Globe Newspaper, 633 F.2d 583, 597 (lst Cir. 1980); Miller v. Transamerica Press, Inc., 621 F.2d 721, 727 (5th Cir. 1980). *Cf.* Downing v. Monitor Publishing Co., 120 N.H. 383, 415 A.2d 683 (1980) (required genuine issue of fact regarding falsity before ordering disclosure of sources).

[124] Australia/Eastern U.S.A. Shipping Conference v. United States, 537 F. Supp. 807, 810 (D.D.C. 1982).

[125] Dale & Dale, *supra* note 8, at 301.

courts continue to recognize First Amendment interests and to weigh them in their calculus.[126]

Individual interpretations of *Herbert* vary, but it is clear that the type of questions *Herbert* allows plaintiffs to ask, such as "Did you doubt *X*'s veracity?,"[127] were recognized as "easy" questions.[128] A well-prepared witness would take an open-ended question like that as an opportunity to tell his or her side of the story.

Herbert's unanticipated but greater impact has been its use by skillful plaintiffs' lawyers to learn about the existence of concrete evidence relevant to the case and then to obtain that evidence through discovery. *Herbert*, for example, has been cited to justify the production of broadcast "outtakes."[129] The rationale is that the outtakes are essential to prove malice by showing the material from which the editor selected the defamatory product. The plaintiff in *Westmoreland v. CBS*, for example, obtained an internal investigation report through discovery because it was relevant to the issue of malice.[130] In a subsequent motion, the court also compelled production of the notes and memoranda used to write the report because it, too, would bear on the defendant's state of mind.[131] The court held that the report "was not newsgathering. It was the kind of inquiry which might be conducted by any entity—a railroad, a

---

[126] *Id.*; Bruno & Stillman Inc. v. Globe Newspaper, 633 F.2d 583, 596–98 (1st Cir. 1980); Solargen Elec. Motor Car Corp. v. American Motors Corp., 506 F. Supp. 546, 553 (N.D.N.Y. 1981).

[127] LIBEL LITIGATION 1981, 186 (R. Winfield ed. 1981). Other examples include: "Did contradictory statements by *Y* lead you to doubt *X*'s veracity?"; "Why or why not?" *Id.*

[128] *Herbert*, 441 U.S. at 172 n.20. The Court noted that these direct questions with respect to the ultimate issue of actual malice would, by defendant's own admission, be "easy" questions to answer. These kinds of questions "are set-up questions for our side. . . . These are not difficult questions to answer." *Id.*

[129] Williams v. ABC, 96 F.R.D. 658, 669 (W.D. Ark. 1983) (court, following *Herbert*, found outtakes discoverable when they involve specific claim of injury arising from a publication or investigation that is alleged to have been knowingly or recklessly false). *But cf.* Steaks Unlimited, Inc. v. Deaner, 623 F.2d 264 (3d Cir. 1980) (state not forced to follow *Herbert* where Pennsylvannia shield law gives media privilege to refuse to surrender outtakes.); United States v. Cuthbertson, 630 F.2d 139 (3d Cir. 1980), *cert. denied* 449 U.S. 1126 (1981) (state with a statutory journalists' privilege can prevent media from being forced to surrender outtakes).

[130] 97 F.R.D. 703 (S.D.N.Y. 1983). The investigation was not privileged because CBS did not treat the report as confidential; it relied on the report at a press conference. *Id.* at 705. The court noted that the report was not only relevant, but that it might also be important evidence. *Id.*

[131] Westmoreland v. CBS, 9 MEDIA L. REP. (BNA) 2316, 2317 (S.D.N.Y. 1983).

hospital, a school or a steel company—in response to a charge that something went wrong."[132]

In sum, although *Herbert* has expanded the scope of discovery material to which plaintiffs now have access, its practical impact on libel litigation has not been as devastating as originally anticipated.[133]

B. *Disclosure of Confidential Sources*

Herbert has affected the disclosure of a media defendant's sources. Although *Herbert* itself never said that confidential sources in libel suits should be given less First Amendment protection than previously afforded, some courts have interpreted the decision as compelling disclosure of such sources.[134] Plaintiffs have argued that one way to prove actual malice is to show that "there are obvious reasons to doubt the veracity of the informant or the accuracy of his reports."[135] Another way is to show that there was no informant and that the publication was "baseless."[136] Failure to disclose the identity of a confidential source could preclude the defendant from relying on the source,[137] lead to a presumption that there was no source,[138] or even result in criminal contempt.[139]

---

[132] *Id.* at 2318.

[133] Perhaps the most creative, yet unsuccessful, use of *Herbert* was in Uhl v. CBS, 476 F. Supp. 1134 (W.D. Pa. 1979). *Uhl* was brought by a hunter who alleged that he was defamed as unsportsmanlike by a film edited to imply that he killed a walking goose. Counsel for the defendant argued that plaintiff's case should be dismissed because he failed to depose CBS employees regarding their states of mind. This prompted the court to remind counsel that such depositions are not mandatory and litigation is not limited to those who can afford "the services of a large double-breasted law firm with platoons of young credit card-carrying associates who can fan out all over the country on a search-and-depose mission." *Id.* at 1141. The jury awarded Uhl one dollar in damages. *Id.*

[134] *See, e.g.*, Rancho La Costa v. Penthouse Int., Ltd., 106 Cal. App. 3d 646, 667, 165 Cal. Rptr. 347, 360 (1980); Downing v. Monitor Publishing, 120 N.H. 383, 386, 415 A.2d 683, 685–86 (1980); McNabb v. Oregonian Publishing Co., 69 Or. App. 136, 685 P.2d 458 (1984), *cert. denied*, 53 U.S.L.W. 3598 (1985); Steaks Unlimited, Inc. v. Deaner, 623 F.2d 264, 279 (3d Cir. 1980); *see generally* Note, *Source Protection in Libel Suits After* Herbert v. Lando, 81 COLUM. L. REV. 338 (1981) (discussing the increase of source disclosure orders in libel suits and arguing that *Herbert* should not have this effect); Note, *Reporters and Their Sources: The Constitutional Right to a Confidential Relationship*, 80 YALE L.J. 317 (1970) (arguing for exception to reporter's privilege in libel cases).

[135] St. Amant v. Thompson, 390 U.S. 727, 727 (1968).

[136] Downing v. Monitor Publishing, 120 N.H. 383, 386, 415 A.2d 683, 685 (1980).

[137] Greenberg v. CBS, 69 A.D.2d 693, 708–709, 419 N.Y.S.2d 988, 997 (1979).

[138] DeRoburt v. Gannett Co. Inc., 507 F. Supp. 880, 886–87 (D. Hawaii 1981).

[139] Garland v. Torre, 259 F.2d 545, 550–51 (2d. Cir.), *cert. denied*, 358 U.S. 910 (1958). The potential importance of the problem is evidenced by one estimate that 33%–50% of all major stories depend on confidential sources. *See*, Note, *Reporters and Their Sources*, *supra* note 134, at 330.

Federal courts have developed a number of tests to decide whether to order disclosure of confidential information. The court in *Bruno & Stillman v. Globe Newspaper*[140] suggested a procedure that first requires the defendants to show a need for confidentiality. If the need for confidentiality is shown, the court next weighs relevance,[141] possibly through an *in camera* examination.[142] Finally, if the information is relevant, the court may require the plaintiff to exhaust alternative sources before it compels production.[143] In one case, the plaintiff moved for disclosure four times before the court believed that the exhaustion requirement was met.[144] Other courts either require plaintiffs first to establish that their claim is meritorious[145] or to forego discovery until after hearing a summary judgment motion on the issue of malice.[146] A trial court, in short, has wide discretion to control the scope of discovery—a discretion that in the ideal situation is "informed by an awareness of First Amendment values."[147]

Herbert's effect on the disclosure of confidential sources has given increased importance to state shield statutes that are designed to protect reporters from forced disclosure of their confidential sources. Twenty-six states currently have shield statutes of varying scope.[148] The Supreme Court in *Herbert* only

---

[140] 633 F.2d 583, 598 (1st Cir. 1980).

[141] Sierra Life v. Magic Valley Newspapers, 101 Idaho 795, 801, 623 P.2d 103, 109 (1980).

[142] *Bruno & Stillman*, 633 F.2d 583, 598 (1st Cir. 1980).

[143] DeRoburt v. Gannett Co. Inc., 507 F. Supp. 880, 886 (D. Hawaii 1981); Mize v. McGraw-Hill, Inc., 82 F.R.D. 475, 478 (S.D. Tex. 1979).

[144] Miller v. Transamerica Press Inc., 621 F.2d 721, 726, *modified on other grounds*, 628 F.2d 932 (5th Cir. 1980).

[145] Senear v. Daily Journal Am., 97 Wash. 2d 148, 155, 641 P.2d 1180, 1183 (1980).

[146] *Bruno & Stillman*, 633 F.2d 583, 598 (1st Cir. 1980); *see also* Cervantes v. Time, Inc., 464 F.2d 986, 994 (8th Cir. 1972), *cert. denied* 409 U.S. 1125 (1973) (reporters relied on nonconfidential sources in successful summary judgment motion).

[147] *Bruno & Stillman*, 633 F.2d 583, 598 (1st Cir. 1980).

[148] *See* ALA. CODE § 12-21-142 (1977); ALASKA STAT. §§ 09.25.150–.25.200 (1983); ARIZ. REV. STAT. ANN. § 12-2237 (1982); ARK. STAT. ANN. § 43-917 (1977); CAL. EVID. CODE § 1070 (West Supp. 1985); DEL. CODE ANN. tit. 10 §§ 4320–26 (1975); Act of Sept. 23, 1983, § 1, ILL. REV. STAT. ch. 110, §§ 8-901 to -909 (1984); IND. CODE ANN. § 34-3-5-1 (West 1983); KY. REV. STAT. § 421.100 (1982); LA. REV. STAT. ANN. §§ 45:1451–:1454 (West Supp. 1982); MD. CTS. & JUD. PROC. CODE ANN. § 9-112 (1984); MICH. COMP. LAWS § 767.5a (1982); MINN. STAT. ANN. §§ 595.021–.025 (West Supp. 1984); MONT. CODE. ANN. §§ 26-1-901 to -903 (1983); NEB. REV. STAT. §§ 20-144 to -147 (1977); NEV. REV. STAT. § 49.275 (1981); N.J. STAT. ANN. §§ 2A:84A-21, -21a, -21.1 to -21.9, -29 (West Supp. 1984); N.M. STAT. ANN. § 38-6-7 (Supp. 1984); N.Y. CIV. RIGHTS LAW §§ 79, 79-h (McKinney 1976 & Supp. 1984); N.D. CENT. CODE § 31-01-06.2 (1976); OHIO REV. CODE ANN. §§ 2739.04, .11, .12 (Page 1981); OKLA. STAT. ANN. tit. 12, § 2056 (West 1980); OR. REV. STAT. §§ 44.510–.540 (1983); 42 PA. CONS. STAT. ANN. § 5942 (Purdon 1982); R.I. GEN. LAWS §§ 9-19.1-1 to -3 (Supp. 1984); TENN. CODE ANN. § 24-1-208 (1980).

specified the minimum constitutional protections for media defendants in libel suits; states are free to legislate beyond this point.[149] For example, the court in *Steaks Unlimited v. Deaner*[150] interpreted the Pennsylvania shield law as foreclosing inquiry into confidential sources and the editorial process. Because state law creates a cause of action for defamation, a state has "substantial latitude" in enforcing the remedy.[151]

Oregon followed a similar analysis in *McNabb v. Oregonian Publishing Co.*[152] In that case McNabb, a police officer, brought a suit, claiming that he was libeled by an article published by the *Oregonian* that called him a racist. During the litigation, McNabb sought discovery of the reporter's notes made while writing the article, the names of sources used, and the substance of interviews not included in the article. The trial court ruled that Oregon's shield law precluded McNabb from discovering these materials. It accordingly granted the defendant's motion for summary judgment.[153] The Oregon Court of Appeals upheld the decision to disallow discovery on the ground that "except for constitutional limitations, to the extent a state authorizes a claim for defamation, it may also limit a party's ability to prove the claim in order to promote other social purposes."[154]

Many courts, however, narrowly apply shield laws to defamation suits.[155] The New York courts, for example, do not permit the state's broad shield law to be used as a "sword" and thus consider the shield protection waived if a libel defendant uses privileged information at trial.[156] Other states limit protection to immunity from contempt when the reporter is a defen-

---

[149] *See generally* L. TRIBE, *supra* note 108, §§ 5-1 to -22, at 224–318.

[150] 623 F.2d 264 (3d Cir. 1980); *see also* Maressa v. New Jersey Monthly, 89 N.J. 176, 188, 445 A.2d 376, 382–83 (1982) (New Jersey shield law prohibits inquiry into the editorial process).

[151] Mazzella v. Philadelphia Newspapers, 479 F. Supp. 523 (E.D.N.Y. 1979), *citing* Gertz v. Robert Welch, Inc., 418 U.S. 323, 346–47 (1974).

[152] 69 Or. App. 136, 685 P.2d 458 (1984).

[153] *Id.* at 462–63.

[154] *Id.* at 463, *citing* Steaks Unlimited, Inc. v. Deaner, 623 F.2d 264, 279 (3d Cir. 1980).

[155] *See, e.g.*, Rosato v. Superior Court, 51 Cal. App. 3d 190, 124 Cal. Rptr. 427 (1975) (shield law overridden by court's inherent power of contempt), *cert. denied*, 427 U.S. 912 (1976); *In re* Farber, 78 N.J. 259, 394 A.2d 330, *cert. denied*, 439 U.S. 997 (1978) (shield law contrary to fair trial guarantee in state constitution); *see generally* Sack, *Reflections on the Wrong Question: Special Constitutional Privilege for the Institutional Press*, 7 HOFSTRA L. REV. 629, 651–54 (1979) (press can obtain protection from sources outside the Supreme Court).

[156] Greenberg v. CBS, 69 A.D.2d 693, 708–09, 419 N.Y.S.2d 988, 997 (1979).

dant.[157] Moreover, in light of *Keeton*[158] and *Calder*,[159] and the relative freedom plaintiffs have in choosing jurisdictions in which to sue, the utility of these shield laws appears doubtful.

Although *Herbert* has increased the likelihood that confidential sources may be disclosed, these recent developments in court practices indicate that a combination of judicial discretion and legislative enactments have made *Herbert*'s actual impact on confidentiality less severe than expected.

C. *Relation to Summary Judgment*

Discovery is only one part of libel litigation; to understand fully the impact of *Herbert*, it is important to see how it relates to other litigation procedures. Because *Herbert* prolongs the discovery process, it similarly affects the timing of summary judgment. Chief Justice Burger's dictum in *Hutchinson*, suggesting that summary judgment is inappropriate when the defendant's state of mind is in question,[160] combines with *Herbert* to exacerbate this delay in the ultimate resolution of the case.

Before *Hutchinson*, summary judgment was considered the rule in disposing of libel cases rather than the exception.[161] Since then, some courts have followed *Hutchinson* and have expressed unwillingness to sanction the use of summary judgment in defamation cases.[162] Other courts, however, apparently have ignored the Chief Justice's suggestion and continue to favor the procedure.[163] The most significant change is that many courts, influenced by the *Hutchinson* debate, have adopted a "neutral

---

[157] CAL. EVID. CODE § 1070(a) (Supp. 1979).

[158] 104 S. Ct. 1473 (1984).

[159] 104 S. Ct. 1482 (1984).

[160] 443 U.S. 111, 120 n.9 (1979).

[161] *See, e.g.*, Southard v. Forbes, Inc., 588 F.2d 140, 145 (5th Cir.), *cert. denied*, 444 U.S. 832 (1979); Martin Marietta Corp. v. Evening Star Newspaper Co., 417 F. Supp. 947 (D.D.C. 1976); Guitar v. Westinghouse Elec. Corp., 396 F. Supp. 1042 (S.D.N.Y. 1975), *aff'd*, 538 F.2d 309 (2d Cir. 1976); Bon Air Hotel, Inc. v. Time, Inc., 426 F.2d 858 (5th Cir. 1970).

[162] *See* Rodriguez v. Nishiki, 65 Hawaii 430, 439, 653 P.2d 1145, 1151 (1982); Hall v. Piedmont Publishing Co., 33 N.C. App. 637, 235 S.E.2d 800, *cert. denied*, 293 N.C. 360, 238 S.E.2d 149 (1977); Burkey v. Delia, 287 Md. 302, 413 A.2d 170 (1980).

[163] *See* Stuart v. Gambling Times, Inc., 534 F. Supp. 170, 172 (D. N.J. 1982); Kotlikoff v. Community News, 89 N.J. 62, 444 A.2d 1086 (1982); Dupler v. Mansfield Journal, 64 Ohio St. 2d 116, 413 N.E.2d 1187, *cert. denied*, 452 U.S. 962 (1981); Mark v. Seattle Times, 96 Wash. 2d 473, 635 P.2d 1081 (1981), *cert. denied*, 457 U.S. 1224 (1982).

approach."[164] The court in *Yiamouyiannis v. Consumers Union*[165] held that "[d]efamation actions are, for procedural purposes, such as discovery, or for summary judgment, to be treated no differently from other actions; any 'chilling effect' caused by the defense of a lawsuit itself is simply to be disregarded."[166]

To the extent that courts have become less willing to grant summary judgment, there has been an increase in discovery. One court noted that "[s]ummary judgment will now be safely granted in libel cases only after it is shown that a sufficient opportunity for discovery has been permitted, and only after the trial judge has made detailed evaluation of the evidence relating to wilfulness or reckless disregard."[167] Summary judgment may still be granted, but its timing is now likely to be after discovery into the state of mind of the defendant has been exhausted.[168] Courts routinely deny summary judgment motions pending additional discovery.[169] *Hutchinson* and *Herbert*, in conjunction, thus have led to both more intrusive discovery and more expensive litigation.

D. *Newsroom Chill*

Immediately after *Herbert v. Lando* was decided, reporters prepared to throw out their notes and first drafts in readiness of the coming onslaught.[170] Since then, however, press fear of intrusive discovery has subsided, and business proceeds as usual. The media has begun to recognize that *Herbert* was mainly a psychological blow.

---

[164] The "neutral approach" is merely the view that no rule exists that favors either granting or denying motions for summary judgment in defamation cases. Schultz v. Newsweek, Inc., 668 F.2d 911, 917 (6th Cir. 1982); Church of Scientology v. Siegelman, 475 F. Supp. 950, 955 (S.D.N.Y. 1979); Gray v. WALA-TV, 384 So. 2d 1062 (Ala. 1980); Thomson v. Cash, 119 N.H. 371, 402 A.2d 651 (1979); Brophy v. Philadephia Newspapers, Inc., 281 Pa. Super. 588, 422 A.2d 625 (1980).

[165] 619 F.2d 932 (2d Cir.), *cert. denied*, 449 U.S. 839 (1980).

[166] *Id.* at 940, *citing* Herbert v. Lando, 441 U.S. 153 (1979).

[167] National Nutritional Foods Ass'n v. Whelan, 492 F. Supp. 374, 379 (S.D.N.Y. 1980).

[168] Fitzgerald v. Penthouse Int'l, Ltd., 525 F. Supp. 585, 597 (D. Md. 1981), *cert. denied*, 460 U.S. 1024 (1983); Loeb v. New Times, 497 F. Supp. 85, 93 (S.D.N.Y. 1980); McIntire v. Westinghouse Broadcasting Co., 479 F. Supp. 808 (D. Mass 1979).

[169] Church of Scientology v. Siegelman, 475 F. Supp. 950 (S.D.N.Y. 1979); Marchiondo v. Brown, 98 N.M. 394, 649 P.2d 462 (1982); Rinaldi v. Viking Penguin, Inc., 52 N.Y.2d 422, 420 N.E.2d 377 (1981).

[170] Garbus, *supra* note 19, at 49.

A recent study assessed the extent of newsroom intrusion. Sixty-seven news men and women from forty media organizations in five states were interviewed extensively.[171] Although the sample was small, the results reveal interesting trends. Eighty-four percent of those surveyed said that their work was not affected in any way by the threat of discovery into the editorial process.[172] Those that did feel that their work was affected characterized *Herbert*'s influence as making them take "greater care" with facts or limiting "wisecracking" in the city room.[173] Nevertheless, of those who did not feel affected, fully eighty percent strongly disapproved of *Herbert*'s holding.[174] Although many thought *Herbert* might affect them in the future, the survey found that much concern was now directed toward fear of newsroom searches or closed courtrooms.[175] Those who most strongly continue to disapprove of *Herbert* see the case as part of a larger struggle between the press and the courts.[176]

The survey concluded that, in fact, the "set-up questions" authorized by *Herbert* would only serve to "boomerang" and to hurt the plaintiffs.[177] The study was limited to an analysis of intrusiveness and was not designed to consider *Herbert*'s influence in prolonging litigation or as a legal tool to acquire confidential sources. It is clear, however, that even if *Herbert* has led to increased disclosure, individual reporters have not changed their work habits. In sum, while the quantity of discoverable material has expanded, *Herbert* has apparently had little impact on the day to day activites of the media.[178] The chill does not seem to have manifested itself in the quality of news reporting.

IV. ACTUAL IMPACT OF *Herbert*: EXPENSE

Herbert may not be the Orwellian terror the press feared; however, its role in adding to the expense of defending a libel

---

[171] Briod, Herbert v. Lando: *Threat To The Press, Or Boomerang for Public Officials?*, 2 COM. & L. 59 (1980).

[172] *Id.* at 74.

[173] *Id.* Several of those interviewed did not want their names used for fear that something they said could later be used against them in a libel litigation. *Id.* at 64.

[174] *Id.* at 75.

[175] *Id.*

[176] *Id.*

[177] *Id.* at 87.

[178] *Id.* at 65. The study concluded that the Supreme Court "handed . . . libel plaintiffs a feather duster with which to go out and fight the press. The press responded to this dubious threat with a barrage from its heaviest cannon. The barrage has abated, but the troops are still at fixed bayonets." *Id.* at 65.

suit may cause a more pervasive chill than editorial intrusiveness. While it is difficult to isolate discovery costs from general litigation expenses, most litigation costs are usually incurred before trial. It is impossible to tell what *Herbert*'s incremental cost effect has been, but common sense suggests that new areas of discovery into the editorial process, increased motions for disclosure of confidential sources, and delayed summary judgment must increase costs. This section looks into the actual cost burden of defending libel suits and concludes that it is this expense that actually has the most deleterious effect on the media's reporting behavior.

A. *Rising Costs*

A combined effect of *Herbert* and *Hutchinson* has been the increase in the likelihood that a libel suit will reach a jury.[179] Although most cases are disposed of favorably to defendants before trial.in a summary judgment motion or a motion to dismiss,[180] if an action survives summary judgment and goes to a jury, plaintiff is likely to win over eighty percent of the time.[181] Bench trials in the past were more favorable to defendants,[182] but this is changing.[183] In the years before *Sullivan*, it was unusual for a defamation award to exceed $10,000;[184] now fifty-three percent of all verdicts are over $100,000.[185] "[T]he rise in jury damage awards in libel cases does not simply parallel what has been happening in other tort areas."[186] The awards are tied to political cases or to cases with strong regional or emotional aspects, such as civil rights cases.[187] More recently, jury awards

---

[179] Smolla, *Let the Author Beware: The Rejuvenation of the American Law of Libel*, 132 U. PA. L. REV. 1, 14 (1983).

[180] Franklin, *Suing Media for Libel: A Litigation Study*, 1981 AM. B. FOUND. RESEARCH J. 795, 802–03 [hereinafter cited as Franklin, *Suing Media*].

[181] *Id.* at 804; *see also Libel Defense Resource Center Study 1982*, 4 LRDC BULL. (Aug. 15, 1982) [hereinafter cited as *LRDC Study 1982*]; *Libel Defense Resource Center Study 1983*, 6 LRDC BULL. (Mar. 15, 1983)[hereinafter cited as *LRDC Study 1983*]; Smolla, *supra* note 179, at 5 (notes study indicating that juries are overwhelmingly more favorable to plaintiffs than are judges).

[182] Franklin, *Suing Media*, *supra* note 180, at 804.

[183] *LRDC Study 1983*, *supra* note 181.

[184] L. ELDREDGE, THE LAW OF DEFAMATION 269 (1978).

[185] MEDIA INSURANCE, *supra* note 100, at 19.

[186] Franklin, *Good Names and Bad Law: A Critique of Libel Law and A Proposal*, 18 U.S.F.L. REV. 1, 10 (1983) [hereinafter cited as Franklin, *Good Names and Bad Law*].

[187] *Id.*

in the millions have become common.[188] There has also been a dramatic increase in juries awarding punitive damages.[189] In *Pring v. Penthouse International Ltd.*, the jury awarded $1,500,000 in compensatory and $25,000,000 in punitive damages[190] before the Tenth Circuit ultimately set aside the verdict and dismissed the action.[191] To the extent that *Hutchinson* and *Herbert* have combined to increase the likelihood that a suit will reach the jury, the defendant's risk of liability has multiplied, as has his litigation costs.[192] Given judicial review of jury verdicts, however, the former cost may not accrue, while the latter appears certain to do so.

Extremely high awards are still statistically rare[193] and are frequently reduced.[194] As exemplified by *Bose*, the greater danger to media defendants is still the cost of defense.[195] Libel attorneys claim that the operative question when deciding whether to publish is no longer "Will we win?" but has now become "Will he sue?"[196] Although the minimum cost of each suit varies, the cost of a full defense of a defamation action in 1975 was estimated at $20,000.[197] Today, the minimum cost to defend a case disposed of through summary judgment with minimal discovery is probably $25,000 to $35,000.[198] *Herbert v. Lando* itself cost between $3 and $4 million.[199] The more recent *Westmoreland v. CBS* case is believed to have cost both parties between $7 and $9 million,[200] with CBS's cost estimated at a sum between $3.75 and $5.75 million.[201] Perhaps because of the expense of litigation, there has been an increase in large settle-

---

[188] *Id.*

[189] MEDIA INSURANCE, *supra* note 100, at 27.

[190] 8 Media L. Rep (BNA) 2409, 2409 (10th Cir. 1982). The trial judge, however, reduced the punitive damages on remittitur to $12.5 million. *Id.*

[191] 695 F.2d 438, 443 (10th Cir. 1982), *cert. denied*, 103 S. Ct. 3112 (1983).

[192] A full trial is four times more expensive than a case disposed of through summary judgment. Smolla, *supra* note 179, at 14 n.81.

[193] Franklin, *Good Names and Bad Law, supra* note 186, at 5.

[194] *Id.* at 5 n.23. Defendants almost always appeal adverse judgments by the trial court, while plaintiffs are likely to cease litigation if they lose at the trial level. Franklin, *Winners and Losers and Why: A Study of Defamation Litigation*, 1980 AM. B. FOUND. RESEARCH J. 455, 462–63.

[195] Bose Corp. v. Consumers Credit Union, 104 S. Ct. 1949 (1984).

[196] Garbus, *supra* note 19, at 49.

[197] Anderson, *Libel and Press Self-Censorship*, 53 TEX. L. REV. 422, 435–36 (1975).

[198] Hentoff, *Free Speech: The Price is Going Up*, THE PROGRESSIVE, May 1983, at 36.

[199] Lewis, New York Times v. Sullivan *Reconsidered: Time to Reform to "The Central Meaning of the First Amendment,"* 63 COLUM. L. REV. 603, 612 (1983).

[200] N.Y. Times, Feb. 18, 1985, § 1, at 1, col. 1.

[201] *See id.*

ments.[202] The incidence of libel suits is increasing,[203] and defense and counseling costs are demanding increasing amounts of a media defendant's income.[204]

The increase in costs has led to an increase in media insurance premiums.[205] Some companies, most notably the New York Times, are self-insured,[206] but most publishers cannot afford this luxury: Insurance funds have been set up to try to protect small publishers.[207] Normally, insurance policies cover only judgments and do not indemnify litigation costs.[208] Recently, however, media clients have been willing to pay higher premiums to protect themselves against onerous expenses.[209] Similarly, punitive damages are often not specifically included in many policies,[210] and some courts have interpreted general "pay all sums" provisions to exclude punitive damages.[211] Moreover, in several states, including California and New York, insurance against punitive damages, payment for an intentional wrong, is barred

---

[202] For example, ABC settled with the drug company, Synanon, for $1.25 million, yet its legal fees were already $6 million. Smolla, *supra* note 179, at 13 n.77. The San Francisco Examiner settled with Synanon for $600,000 and published a front-page apology. N.Y. Times, July 3, 1976, at 24, col. 1. *See also* Friendly, *CBS Settles Large Suits By Public Officials in 2 Cities*, N.Y. Times, Oct. 21, 1982, at A18, col. 1.

[203] *Aggressive TV News Reporting Brings Flood of Lawsuits*, L.A. Daily J., Feb. 4, 1983, at 5, col. 3 [hereinafter cited as *TV News Reporting*].

[204] Franklin, *Suing Media*, *supra* note 180, at 795, 800 n.13. For example, a daily newspaper with a circulation between 100,000 to 300,000 reported the following increase in litigation and counseling costs:

	Litigation	*Counseling*	*Total*
1976	$ 1,400	$ 6,000	$ 7,400
1977	3,100	6,000	9,100
1978	39,200	7,500	46,700
1979	19,500	10,000	29,500
1980	56,900	17,500	74,400

Id.

[205] *Id.*

[206] Franklin, *Good Names and Bad Law*, *supra* note 186, at 18. The New York Times also refuses monetary settlements in an effort to deter suits. Drew Pearson, prominent columnist and self insurer, was sued 275 times, losing only one case for $40,000 but spending hundreds of thousands of dollars in legal fees. Anderson, *supra* note 197, at 432 n.51, *citing* O. PILAT, DREW PEARSON: AN UNAUTHORIZED BIOGRAPHY 10 (1973).

[207] Franklin, *Reflections*, *supra* note 89, at 1049 n.83. Most defamation insurance is underwritten by CNA, Employers Reinsurance, Mutual of Bermuda, Fireman's Fund, and Chubb. MEDIA INSURANCE, *supra* note 100, at 169.

[208] MEDIA INSURANCE, *supra* note 100, at 354.

[209] *Id.* Seventy-five to eighty percent of all insurance payments are for defense costs. *Id.* at 373.

[210] *Id.* at 175–77.

[211] *Id.* at 175, *citing* Tedesco v. Maryland Cas. Co., 127 Conn. 533, 18 A.2d 357 (1941); Braley v. Berkshire Mutual Ins. Co., 440 A.2d 359, 361 (Me. 1982); Universal Indem. Ins. Co. v. Tenery, 96 Colo. 10, 39 P.2d 776 (1934).

as contrary to public policy.[212] Now that *Calder* and *Keeton* have allowed plaintiffs more freedom in choosing a forum, defendants run the risk of suit in a jurisdiction not covered by their insurance.[213]

Litigation costs are similarly high for plaintiffs. Because of the expense and small likelihood of eventual success, libel is rapidly becoming a rich person's action.[214] Moreover, as Justice Marshall noted, many defamation plaintiffs are not animated by a "rational calculus of their chances of recovery."[215] Frequently, plaintiffs want revenge.[216] Some commentators have claimed that politicians may use libel actions as political weapons to stifle criticism.[217]

There also has been an increase in suits brought by corporations and their executives.[218] One telling sign of the times is the decision by Mobil Oil Corporation to offer insurance policies to its one hundred top executives that will indemnify each of them for up to ten million dollars in legal fees each.[219] Media representatives have called the policies "hunting licenses" and have characterized them as an attempt to discourage reporters from writing about Mobil.[220] Previously, libel litigation was not prof-

---

[212] *Id.* at 27. New York requires insurance policies be filed with the state for approval. *Id.* at 270. *See* Public Service Mut. Ins. Co. v. Goldfarb, 53 N.Y.2d 392, 398, 425 N.E.2d 810, 815, 442 N.Y.S.2d 422, 425 (1981) (public policy barred coverage of punitive damages); Employers Reinsurance Corp. v. National Enquirer, Inc., Cal. Sup. Ct., Los Angeles County, Civ. No. C-387419 (unreported decision of Sept. 3, 1982, order entered Sept. 20, 1982) (National Enquirer not entitled to indemnity for punitive damage award in suit by Carol Burnett). Seventeen jurisdictions have yet to address the propriety of punitive damages, but of the thirty-nine that have, twenty-five have upheld insurability. MEDIA INSURANCE, *supra* note 100, at 196–97.

[213] There are already attempts to circumvent state law prohibiting indemnification of punitive damages through the use of "off-shore" insurance companies that do not "do business" in the United States. MEDIA INSURANCE, *supra* note 100, at 275–76.

[214] Franklin, *Good Names and Bad Law*, *supra* note 186, at 29–30.

[215] Herbert v. Lando, 441 U.S. 153, 204 (Marshall, J., dissenting).

[216] Franklin, *Good Names and Bad Law*, *supra* note 186, at 5 n.26.

[217] *See generally* Smolla, *supra* note 179 (noting an ambassador, two governors, and a mayor recently have sued); Note, *Libel and the Reporting of Rumor*, 92 YALE L.J. 85, 85 (1982) (discussing President Jimmy Carter's threat to sue over a story that the White House was bugged).

[218] *See generally* Drechsel & Moon, *Corporate Libel Plaintiffs and the New Media*, 21/2 AM. BUS. L.J. 27 (1983).

[219] *See Tension on the Frontiers of Libel*, N.Y. Times, Dec. 18, 1983, § 3 at 4, col. 3.

[220] *Id.* (statement of Floyd Abrams). For another indication of the antagonism between Mobil and the press, see Mobil's advertisement on the Op. Ed. page of the *New York Times* regarding the Washington Post's publication of the rumor that President Carter had bugged the quarters of President-elect Reagan. Mobil offered a "modest proposal" suggesting that the *Post* label its stories "CST" (contains some truth), "BS" (basically silly), "BR" (believable rumor), and "AFST" (absolutely for sure true) and that it publish a section called "Totally Untrue But Interesting," printed on yellow paper. N.Y. Times, Oct. 29, 1981, at A27, col. 4.

itable enough to support a plaintiff's bar; however, this has changed.[221] The lure of high verdicts is of course one factor, but the advent of conservative public interest law firms[222] and the repeated willingness of certain cult religious groups to sue[223] have started to upset the status quo.

The increasing willingness of various groups to sue is part of a general dissatisfaction with the press. In concentrating on the relationship between the press and the judiciary, it is easy to forget an important third party—the public. The press normally characterizes its absolutist appeals for First Amendment protection as being in the public interest;[224] however, it is unlikely that the public shares this view. After Watergate, the press rushed to expand its investigative journalism.[225] The result has been an increase in libel suits and a dislike for "ambush" techniques.[226] Some perceive the press as arrogant and aloof,[227] imbued with a "salvationist ethic,"[228] and unwilling to admit mistakes.[229] Strident press reaction to decisions like *Herbert* only exacerbates the problem while the extensive publicity afforded actions such as Carol Burnett's suit against the National Enquirer[230] only encourages more suits. "The jury room is an opportunity" for the public to "get even" with the press.[231] There is a fundamental difference in outlook: the press sees its archetypal image as *All the President's Men*, while the public more likely sees it as *Absence of Malice*.

---

[221] Franklin, *Reflection*, *supra* note 89, at 1057 n.16. *But see* N.Y. Times, Feb. 18, 1985, at B-4, col. 3 (noting Capital Legal Foundation's $500,000 debt after handling the *Westmoreland* case).

[222] General Westmoreland, for example, was represented in his suit against CBS by the Capital Legal Foundation. Westmoreland v. CBS, 97 F.R.D. 703 (S.D.N.Y. 1983).

[223] One court has referred to the Church of Scientology as "litigious." Church of Scientology v. Siegelman, 475 F. Supp. 950 (S.D.N.Y. 1979). It has sued one woman alone over seventeen times. Garbus, *supra* note 19, at 48.

[224] Dale & Dale, *supra* note 8, at 289; Franklin, *Reflections*, *supra* note 89, at 1051.

[225] *TV News Reporting*, *supra* note 203; L.A. Daily J., Feb. 4, 1983, at 5, col. 3.

[226] *See* Franklin, *Good Names and Bad Law*, *supra* note 186, at 9 n.42.

[227] *Of Reputations and Reporters*, TIME, Mar. 19, 1984, at 64 (quoting U.S. Appeals Court Judge Irving Kaufman).

[228] Chesire, *The Imperial Press*, NAT'L REV., Aug. 17, 1979, at 1021. Too many reporters feel "what's good for the press is good for the country." Dale & Dale, *supra* note 8, at 304, *citing* Grunwald, *Don't Love the Press, But Understand It*, TIME, July 8, 1974, at 74–75.

[229] MEDIA INSURANCE, *supra* note 100, at 300.

[230] Burnett v. National Enquirer, 144 Cal. App. 3d 991, 193 Cal. Rptr. 206 (1983), *app. dismissed*, 104 S. Ct. 1260 (1984).

[231] Goodale, *Is the Public "Getting Even" with the Press in Libel Cases?*, N.Y.L.J., Aug. 11, 1982, at 1, col. 2.

B. *Investigative Chill*

The increasing costs associated with defending a libel suit probably create the greatest media reporting chill. There is widespread recognition that the fear of litigation is as great as the fear of its outcome.[232] It is difficult to learn, however, which stories were "killed" for being too controversial.[233] Also, many institutions, such as the Washington Post, insist that they have not changed their editorial practices because of potential suits.[234] There has been a decrease in investigative journalism, but this might be attributed to the natural end of a cycle.[235]

It is clear, however, that there is now less economic incentive to publish potentially libelous stories. Few towns have more than one newspaper. In towns where there is competition, newspapers are more likely to compete by expanding their comics section or by offering contests than by attempting a one-time publication of a daring exposé.[236] Professional pride, however, still provides incentives for newspapers to take risks.

While it is difficult to assess individual behavior, it is clear that the state of libel litigation has caused media organizations to take prepublication review precautions.[237] For some news people, however, "[a] lawyer in the newsroom is more dangerous to freedom of the press than a reporter in jail."[238] Similarly, the factors an insurance company considers in evaluating the cost of premiums encourage self-censorship. They impose the heaviest burdens on controversial publishers who are most in need of protection.[239]

Perhaps the most troubling aspect of the chill from litigation expense is the disparate impact on the small publisher. Few organizations have the resources of CBS or the New York

---

[232] *See, e.g.*, New York Times v. Sullivan, 376 U.S. 254, 282–83 (1964); Washington Post v. Keogh, 365 F.2d 965, 968 (D.C. Cir. 1966), *cert. denied*, 385 U.S. 1011; National Nutritional Foods Ass'n v. Whelan, 492 F. Supp. 374 (1980).

[233] ABC's decision not to broadcast a docu-drama of Elizabeth Taylor's life is a rare exception. Smolla, *supra* note 179, at 3.

[234] *See* Smith, *The Rising Tide of Libel Litigation*, 44 MONT. L. REV. 71, 87 (1983).

[235] O'Neill, *The Ebbing of the Great Investigative Wave*, BULL. AM. SOC. OF NEWSPAPER EDITORS, Sept. 1983, at 26.

[236] Anderson, *supra* note 197, at 433.

[237] *See generally* LIBEL LITIGATION 1981, *supra* note 127, at 22–35. For example, some organizations now consult lawyers regularly before publication. *Id.*

[238] EDITOR & PUBLISHER, Aug. 14, 1983, at 8 (statement of columnist Lyle Dennisten).

[239] MEDIA INSURANCE, *supra* note 100, at 360–62. Factors include loss experience, editorial content, education of staff, prepublication review, circulation, use of in-house counsel, and punitive damage coverage. *Id.*

Times. The Alton Telegraph, circulation 38,000, was recently served with a $9.2 million libel judgment, which forced it into bankruptcy.[240] The Texas Observer had to hold a fund-raiser in order to meet its litigation costs.[241]

To have a truly free flow of information, it is important to allow more than just the most powerful to speak. Yet in our present defamation system only the wealthiest institutions can publish confidently, and only the wealthiest individuals can defend their reputations completely.

V. Analysis of Suggested Solutions

The present system of libel litigation serves no one's interests well. The press fears losing First Amendment freedoms and suffers devastating litigation expenses, while potential plaintiffs can look forward to similarly high costs and little hope of eventual vindication. Valuable judicial resources are expended for small gain. Because First Amendment issues such as these attract much concern, commentators have advanced a variety of proposals to alleviate the situation. Unfortunately, the solutions proposed to date are inadequate: many are overly reliant on the likelihood of sweeping change; others serve special interests at the expense of the system as a whole; and still others simply fail to accomplish their objectives. Some proposals suffer from a combination of all three defects.

A. *Changing the Constitutional Standard*

One proposed solution is to abolish defamation law, at least in the "sphere of public issues and official conduct,"[242] and to provide absolute immunity to the press in libel suits. Press advocates argue that the *Sullivan* standard is not enough; if legislators, judges, and members of the executive branch have immunity for statements made "in the discharge of their public duties . . . [w]hy should the press have any less of a right?"[243]

---

[240] *See* Smolla, *supra* note 179, at 12–13 nn.72–73.

[241] *See* Smith, *supra* note 234, at 88. One court in West Virginia expressed fear of potential chill from libel suits because that state's newspapers have limited financial resources. Sprouse v. Clay Communication, Inc., 158 W. Va. 427, 211 S.E.2d 674, *cert. denied*, 423 U.S. 882 (1975).

[242] Oakes, *supra* note 61, at 720.

[243] Garbus, *The Limits for Libel*, N.Y. Times, July 29, 1983 at A23, col. 2; *see also* Lewis, *supra* note 199, at 624.

Dissatisfaction with *Sullivan* stems from both the conviction that political speech, even if defamatory, should not be regulated[244] and agreement with Justice Black's concern that juries who dislike the media can misapply the actual malice test to find liability.[245] The Supreme Court has given new impetus to the call for absolute immunity with its decisions in *Herbert*, *Hutchinson*, *Calder*, and *Keeton*.[246] To the extent that these cases preclude special constitutional procedural protections for the press, it makes sense for the press to look to changes in the substantive law for greater protection.

Abolishing all liability for libel would certainly reduce editorial chill and litigation expense, but, for a number of reasons, this is not a satisfactory approach to the problems of libel law. First, the Supreme Court has firmly entrenched its policy of balancing interests in First Amendment cases and shows no sign of accepting absolute immunity. The Court stated in *Herbert* that:

> [I]n the 15 years since *New York Times*, the doctrine announced by that case, which represented a major development and which was widely perceived as essentially protective of press freedoms, has been repeatedly affirmed as the appropriate First Amendment standard applicable in libel actions brought by public officials and public figures. . . . At the same time, however, the Court has reiterated its conviction—reflected in the laws of defamation of all the States—that the individual's interest in his reputation is also a basic concern.[247]

Moreover, as *Bose* indicates,[248] some members of the Court are more willing to cut back on the protections afforded by *Sullivan* than to offer new ones.[249] Similarly, suggestions that the *Gertz* public figure rule be abandoned in favor of a "context public figure"[250] scheme are unrealistic in light of the Court's restriction of the public figure doctrine in *Hutchinson* and *Wolston*.[251] Urging the Court to reconsider the basic *Sullivan* standard could

---

[244] A. MEIKLEJOHN, POLITICAL FREEDOM 26–27 (1960).

[245] *Sullivan*, 376 U.S. at 293 (Black, J., concurring); *see also id.* at 297 (Goldberg, J., concurring).

[246] *See supra* text accompanying notes 33–34, 55–79, 99–100.

[247] *Herbert*, 441 U.S. at 169 (citations omitted). Moreover, the Framers did not intend the First Amendment to abolish liability for libel. *Id.* at 158.

[248] *See supra* text accompanying notes 80–87.

[249] *See Bose*, 104 S. Ct. at 1967 (Rehnquist, J., dissenting); *see supra* note 87.

[250] Smolla, *supra* note 179, at 12.

[251] *See supra* text accompanying notes 33–36.

easily result in a less protective test than that which currently exists. Until the complexion of the Court becomes more hospitable to press clause concerns, sweeping change in the constitutional standard is unlikely, and advocates of absolute immunity should content themselves with the knowledge that even opinions such as *Calder*, seen as inimicable to press freedom, do not cut back on the substantive protections of *Sullivan*.

Second, not only is the Court unlikely to change the substantive standard, but the constituency supporting absolute immunity is small. The idea was most popular in the years immediately before *Sullivan*.[252] Since then, with the notable exception of the American Civil Liberties Union, which switched its position to favor absolute protection, support has dwindled.[253] Significantly, some larger publishers, including the Gannett Company and the New York Times, are opposed to a total prohibition on defamation suits.[254] They generally believe that *Sullivan* "strikes a fair balance"[255] and absolute immunity "goes further than we actually need."[256] Moreover, an absolute prohibition would leave truly injured plaintiffs without a remedy. Public opinion of the press is already low;[257] lobbying for absolute immunity would not raise it.

Third, the actual malice standard is still a heavy burden for plaintiffs to meet, and, as former Israeli Defense Minister Ariel Sharon's recent loss to Time has highlighted,[258] it routinely decides cases.[259] If the problem is that the *Sullivan* rule is too confusing for a jury to understand,[260] then the solution is (1) to attempt to get a bench trial, (2) to argue the case stressing actual

---

[252] Franklin, *Reflections*, *supra* note 89, at 1042 n.52.

[253] Franklin, *Good Names and Bad Law*, *supra* note 186, at 23 n.105, *citing* CIVIL LIBERTIES, Feb. 1983, at 2; CIVIL LIBERTIES, June 1983, at 11. *But see* NEWSWEEK, Feb. 4, 1985, at 55 (stating that "[n]ow legal specialists like columnist Anthony Lewis and groups like the American Civil Liberties Union are reviving the idea [of an absolute immunity for the press from libel suits by public figures]").

[254] Franklin, *Good Names and Bad Law*, *supra* note 186, at 26 n.124.

[255] N.Y. Times, May 7, 1983, at 22, col. 1.

[256] EDITOR & PUBLISHER, Apr. 23, 1983, at 84 (quoting Alice Neff Lucan, Gannett Co.'s in-house libel counsel).

[257] *See supra* notes 227–229 and accompanying text.

[258] *See* N.Y. Times, Jan. 25, 1985, at 1, col. 2.

[259] *See, e.g.*, Brewer v. Memphis Publishing Co., 626 F.2d 1238 (5th Cir. 1980), *cert. denied* 452 U.S. 962 (1981); Ali v. Daily News, 8 Media L. Rep. (BNA) 1844 (D.V.I. 1982); Barbarita v. Gannett Co. Inc., 8 Media L. Rep. (BNA) 1050 (N.Y. Sup. Ct. 1981), *aff'd*, 92 A.D.2d 599 (1983), *appeal denied*, 59 N.Y.2d 604 (1983).

[260] *See* Lewis, *supra* note 199, at 617. *Cf.* Smolla, *supra* note 179, at 21 (arguing that a jury's verdict will be based more on its rough sense of justice than upon the complex constitutional doctrine surrounding libel law).

malice, (3) to move for a directed verdict, (4) to ask for simple jury instructions, (5) to request a special verdict, (6) to move for a judgment notwithstanding the verdict, and (7) to prepare to appeal. Recent experience has shown that the trial court still has the discretion to ensure that the actual malice standard is followed. For example, the special verdict, as utilized in the Sharon litigation,[261] and the mini-summation, as utilized in the Westmoreland case,[262] are techniques that help to ensure the *Sullivan* standard's proper application. A different test may be less complex, but also less protective.

Fourth, there are other constitutional changes that the press can suggest that have a better chance of being realized because they do not ask the Court to revise extensively well-entrenched law. For example, the Supreme Court has never addressed the constitutionality of punitive damages in libel suits. Punitive damages have been presumed to be appropriate since *Gertz v. Robert Welch, Inc.*, when the Court held that states could not allow recovery of punitive damages without showing actual malice,[263] but the Court has never specifically examined the relationship between punitive damages and the First Amendment.[264] At least one state, however, has held that punitive damages in defamation actions violate its state constitution because punitive damages in a sense punish speech.[265] Punitive damages are often several multiples of the amount of actual damages,[266] and several

---

[261] In Ariel Sharon's recent suit against *Time* magazine, U.S. District Court Judge Abraham D. Sofaer instructed the jury to return its verdict in three separate stages. The jury was told to decide first whether *Time* had defamed Sharon; second, whether the defamatory statements were true; and third, whether *Time* had acted with actual malice. Judge Sofaer's instructions were credited with "ensuring that the jurors would keep to their appointed course rather than get hopelessly lost in the laws' conundrums." NEWSWEEK, Feb. 4, 1985, at 55. *See* Sharon v. Time, Inc., Civ. No. 83-4660 (ADS) (S.D.N.Y. 1983) (charge to the jury); *see also* NEWSWEEK, Feb. 4, 1985, at 52–54; NEWSWEEK, Jan. 28, 1985, at 46–48; N.Y. Times, Jan. 25, 1985, at B4, col. 1.

[262] In General Westmoreland's libel suit against CBS, Judge Pierre N. Leval allowed attorneys for each side to use "mini-summations" to help jurors understand the complex trial. The judge allotted a total of two hours to each side for the duration of the trial and required only that the mini-summations not be used to interfere unduly with the opposing side's presentation of evidence. Attorneys for each side could use their two hours in blocks as large or as small as they saw fit, and for any purpose that they desired. Attorneys for both sides found the procedure effective and approved its use in long, complex trials. Legal Times, Feb. 25, 1985, at 1, col. 2.

[263] 418 U.S. at 349–50.

[264] Note, Punitive Damages and Libel Law, 98 HARV. L. REV. (1985) (in press).

[265] Wheeler v. Green, 286 Ore. 99, 119–20, 593 P.2d 777, 789 (1979). *Cf.* Madison v. Yunker, 180 Mont. 54, 67, 589 P.2d 126, 133 (1978) (holding that punitive damages may be awarded only for violations of the actual malice standard).

[266] For example, William Tavoulareas in his suit against the Washington Post was awarded seven times his actual damages (verdict reinstated by the Court of Appeals for

commentators urge their abolition by either constitutional interpretation or legislative mandate.[267] Abolishing punitive damages still would allow plaintiffs to recover their real losses, but it would reduce their windfall incentive. Consequently, insurance rates should diminish to reflect diminished exposure. Finally, because punitive damages are a question of substantive law, instead of procedure, the Supreme Court might be willing to make an incremental substantive change while still affirming *Sullivan*.

B. *Creating Nonmonetary Remedies*

A second kind of solution would be the creation of nonmonetary remedies. "[T]he vindication of one's good name does not require colossal verdicts."[268] If the substantive law is inefficient and expensive, one solution is to create new remedies through legislative enactment. The key to evaluating the merit of the various proposals creating nonmonetary remedies is understanding the incentives that they offer litigants.[269]

Right to reply statutes are one alternative. Such statutes allow an individual to compel a publisher or broadcaster to provide a forum to respond to an allegedly defamatory statement. The reply must be as conspicuous as the original statement and must be provided free of charge.[270] The theory behind these statutes is that reputation is merely a form of public opinion and the public should be allowed to decide among competing viewpoints.[271] However, reply statutes suffer from several drawbacks. They do not help a potential plaintiff who has suffered

---

the District of Columbia Circuit, Washington Post v. Tavoulareas, Nos. 83-1603, 83-1604 (D.C. Cir. April 9, 1985)). Carol Burnett won four times her actual loss from the National Enquirer. *See* Goodale, *Is the Public 'Getting Even' With the Press in Libel Cases?*, 188 N.Y.L.J. 29 (1982).

[267] *See, e.g.*, Goodale, *supra* note 266; Franklin, *Reflections*, *supra* note 89, at 1048–49; Lewis, *supra* note 199, at 615; TIME, Mar. 19, 1984, at 64 (statement of Judge Irving Kaufman). *But cf.* 10 MEDIA L. REP. (BNA), Mar. 20, 1984 (news section) (statement of professor Diane Zimmerman that even if punitive damages were unconstitutional, juries could manipulate actual damages to give high awards).

[268] TIME, Mar. 19, 1984, at 64 (statement of columnist Anthony Lewis).

[269] Nonmonetary remedies in libel cases are not new; in the ninth century, under the laws of King Alfred of Wessex, a person found guilty of libel paid no damages, but suffered "no lighter penalty than the cutting off of his tongue." Garbus, *supra* note 19.

[270] *See, e.g.*, 1913 Fla. Laws 274 (codified at FLA. STAT. § 104.38) (repealed 1975).

[271] *See* Pedrick, *Senator McCarthy and the Law of Libel: A Study of Two Campaign Speeches*, 48 NW. U.L. REV. 135, 179 (1953); *see also* Pedrick, *Freedom of the Press and the Law of Libel: The Modern Revised Translation*, 49 CORNELL L.Q. 581, 606 (1964).

actual damages. Nor do they provide a defamed party with authoritative vindication. Without an express finding or admission of falsehood, it is unlikely that the right to reply will restore to many plaintiffs their good names. Most importantly, though, from the press's perspective, compulsory reply statutes are a tremendous intrusion on the editorial process. The Supreme Court, in *Miami Herald Publishing Co. v. Tornillo*,[272] recognized this threat to the First Amendment and, in a unanimous ruling, declared these statutes unconstitutional.

A second suggestion is to use the procedure available under the Federal Declaratory Judgment Act to determine falsity.[273] The Act, which empowers federal courts to decide the rights of interested parties and to grant appropriate relief, is recommended by many commentators for use in lieu of full-fledged libel actions.[274] A declaratory judgment proceeding would concentrate solely on the veracity of the published statement. Discovery would therefore be minimal, expenses would be low, and an individual would have the opportunity to clear his or her name. One commentator has suggested that the proceeding be called a "vindication action" and that media organizations that published falsities be required by the court to circulate the truth.[275] While declaratory judgments may be useful in some instances, for most plaintiffs their appeal would be limited.

Plaintiffs have little incentive to seek a declaratory judgment. First, proving truth is often difficult and expensive.[276] "Minimal discovery" is more optimistic than realistic. Second, while costs remain high, the potential outcome for an injured party is purely nonmonetary. Most plaintiffs would probably choose an incremental increase in litigation expenses for the possibility of a

---

[272] 418 U.S. 241 (1974). The statute invalidated in *Tornillo* "create[d] a right to reply to press criticism of a candidate for nomination or election" to public office. *Id.* at 247. In overturning the statute, the Supreme Court rejected the notion that the First Amendment creates a positive right of access to the media, *see id.* at 247–55, and held that even government regulation aimed at stimulating the exchange of ideas was an unconstitutional interference with the editorial independence of the press. *Id.* at 258.

[273] 28 U.S.C. §§ 2201–2202 (1973). Declaratory judgments are also available in the thirty-nine states that have adopted the Uniform Declaratory Judgment Act. *See* UNIFORM DECLARATORY JUDGMENT ACT ANN. 39 (1922).

[274] *See, e.g.*, Hulme, *Vindicating Reputation: An Alternative To Damages As A Remedy For Defamation*, 30 AM. U.L. REV. 375 (1981); *see also* DAVILA, LIBEL LAW AND THE PRESS (1971) (discussing West German system of no damages but forum for truth). *Cf.* Lewis, *supra* note 199, at 616 n.78 (discussing possible use of a streamlined procedure designed solely to determine falsity, but pointing out the probable conflict between *Tornillo* and a court's power to order a retraction).

[275] Hulme, *supra* note 274, at 391–92.

[276] Garbus, *supra* note 19, at 49.

high award.[277] Moreover, because requiring a media organization to circulate news of an adverse judgment is likely to be unconstitutional under *Tornillo*, a plaintiff could end up with a useless victory. Furthermore, a "vindication action" makes no allowances for a plaintiff who has suffered special damages, or, for that matter, for a plaintiff who wants revenge. A declaratory judgment also serves as a disincentive for potential defendants to settle or to retract because it removes the threat of a large award. In sum, a vindication action is likely to be ineffective and to create a host of new problems.

A third alternative has been suggested by professor Marc Franklin, who has proposed the creation of a "restoration action."[278] Franklin's remedy requires a plaintiff to present a publisher of a defamatory statement with evidence that the statement is false. If the publisher refuses to retract, the plaintiff may make an irrevocable election between a damage action and a restoration action. If the plaintiff chooses restoration, he or she must "present in court essentially the same evidence that was presented to the defendant, and must persuade the trier of fact of the falsity with convincing clarity."[279] A successful plaintiff is entitled to reasonable attorneys' fees, and a successful defendant is similarly entitled to recover fees if the plaintiff is found to have brought the action without a reasonable chance for success.[280]

Franklin succeeds in encouraging creative solutions to the current problems in libel law, but his restoration action itself is an impractical complication. Franklin designed the action in part to provide a remedy for those plaintiffs who cannot afford to bring a full-fledged action, yet a plaintiff who has suffered actual damages will not be made whole by a purely nonmonetary remedy. Wealthy plaintiffs are likely to ignore the proceeding and to persist in bringing suits for damages, thereby perpetuating the cycle of high awards and expenses. The restoration action will actually increase the already high level of litigation, yet provide no more satisfaction.

There are additional problems with professor Franklin's proposal. First, there is still the possibility of an empty victory because the press cannot be compelled to publish news of an

---

[277] *See* Lewis, *supra* note 199, at 616 n.78.
[278] *See generally* Franklin, *Good Names and Bad Law*, *supra* note 186.
[279] *Id.* at 36.
[280] *See id.* at 29–30.

adverse judgment. Second, Franklin would not permit discovery in a restoration action. This assumes, however, that everything that is needed to prove truth is in the possession of the plaintiff. This is not always the case. It is likely that over time the restoration action, if it is to have any meaning, would take on all the discovery trappings and expenses of a full-fledged proceeding. Third, the proposal attempts to encourage settlement, but actually would result in delay. By requiring plaintiffs to present proof of falsity to defendants, Franklin promotes discussion; however, because there is the possibility a plaintiff may elect a nonmonetary remedy, a publisher would be tempted to avoid settling until the plaintiff made the election and the publisher knew what was at stake. Fourth, *Gertz* may not allow attorneys' fees to be awarded without fault.[281]

These alternative nonmonetary remedies are just a few of the ideas responding to the perceived crisis in libel law. More work needs to be done to ensure the that the proposed cure is better than the illness.

C. *Amending the Federal Rules of Civil Procedure*

After *Herbert* was decided, the Advisory Committee on the Federal Rules of Civil Procedure was "besieged by requests from media organizations for prompt action."[282] Press organizations wanted the Federal Rules to be altered to give them special protections against intrusive discovery in defamation actions. Various suggestions were offered, including narrowing the standard of relevance, formalizing a bifurcated discovery procedure, and requiring judicially supervised discovery conferences.[283] The Advisory Committee rejected these proposals. The reasoning behind that decision remains valid and recent developments lend it further justification.

One solution offered was for district judges to apply a stricter standard of relevance when evaluating requests for discovery into the editorial process.[284] The impetus for the suggestion

[281] Franklin himself noted this problem but did not discuss its consequences. *Id.* at 47.

[282] Discovery From Media Defendants In Privacy and Defamation Actions, Discussion Guide for the Advisory Committee on the Federal Rules of Civil Procedure, Topic Q, at 1 (Dec. 1979) (on file at HARV. J. ON LEGIS.) [hereinafter cited as Discussion Guide].

[283] *Id.* at 8.

[284] *Id.* at 9–10.

stemmed from *Herbert* itself. In his concurring opinion in *Herbert*, Justice Powell urged courts to "measure the degree of relevance required" by balancing the private needs of the parties with First Amendment implications.[285] Justice Marshall, dissenting, specifically called for a "strict standard of relevance" in defamation suits;[286] and Justice Stewart, also dissenting, apparently applied a narrow interpretation of Rule 26(c) to find plaintiff's questions aimed at actual malice "totally irrelevant."[287]

The Advisory Committee, however, reasoned that the *Herbert* majority implicitly rejected a stricter standard of relevance.[288] Such a standard would create indirectly the editorial process privilege that the majority refused to create directly.[289] The Advisory Committee also noted professor Friedenthal's analysis of additional problems with a special relevance standard.[290] First, Friedenthal was concerned that narrowing relevance would increase expensive and time-consuming motion practice.[291] Recent empirical evidence has supported this position. A Federal Judicial Center study indicated that discovery motions accounted for 17.9% of all pretrial motions,[292] and 59.5% of these were motions to compel.[293] Creating a special exception for libel cases narrowing the "liberal construction"[294] of the Federal Rules familiar since *Hickman v. Taylor*[295] would only increase the potential for dispute. Second, a pretrial limitation on discovery would have limited utility without a parallel restriction on the admissibility of evidence at trial.[296] One without the other would cause a trial to grind to a halt in order to conduct more discovery. Finally, a narrower federal standard would encourage forum

---

[285] *Herbert*, 441 U.S. at 179 (Powell, J., concurring).
[286] *Id.* at 206 (Marshall, J., dissenting).
[287] *Id.* at 202 (Stewart, J., dissenting); *see* FED. R. CIV. P. 26(c).
[288] *See generally* Discussion Guide, *supra* note 282, at 9.
[289] *Id.*
[290] Friedenthal, Herbert v. Lando: *A Note on Discovery*, 31 STAN. L. REV. 1059, 1062–64 (1979).
[291] *Id.* at 1062–63.
[292] *See* P. CONNOLLY & P. LOMBARD, JUDICIAL CONTROLS AND THE CIVIL LITIGATIVE PROCESS: MOTIONS (table 19) (Federal Judicial Center Study 1980).
[293] *Id.* at table 22.
[294] *See, e.g.*, Burns v. Thiokol Chemical Corp., 483 F.2d 300 (5th Cir. 1973); *see generally* 8 C. WRIGHT & A. MILLER, FEDERAL PRACTICE & PROCEDURE: CIVIL § 2001, at 17, § 2008, at 41–51 (1982).
[295] 329 U.S. 495 (1947).
[296] Friedenthal, *supra* note 290, at 1063.

shopping by creating an incentive for plaintiffs to litigate in state courts that adhered to the old test.[297]

Analysis of the Advisory Committee's decision to reject a strict standard of relevance is informed by the discussions surrounding the 1979 amendments to the Federal Rules.[298] Widespread concern that "discovery abuse" was rampant[299] led to the suggestion that Rule 26(c) be amended to limit the scope of discovery to matters relevant to "issues raised" by the claims or defenses instead of the broader existing test of "any matter . . . relevant to the subject matter."[300] After extensive hearings and debate,[301] the Advisory Committee decided not to recommend altering the relevance standard.[302]

The Advisory Committee noted that empirical evidence did not indicate that discovery was in fact being abused.[303] There was fear that motion practice interpreting the meaning of the changed language would consume any possible efficiencies.[304] Most importantly, such a revision would require a return to the "byzantine technicalities" of pre-Federal Rule pleading.[305]

Similarly, a revised relevance test for defamation actions is unwarranted in light of *Herbert*'s limited actual chilling effect. A greater danger to media defendants would result from the increased expense of motion practice to define the new standard and the difficulty in constructing satisfactory pleadings.

Another proposal was offered by the Radio Television News Directors Association (RTNDA).[306] RTNDA called for the amendment of Rule 26 to include a subdivision (g) governing

---

[297] *Id.* at 1064.

[298] *See generally Amendments to the Federal Rules of Civil Procedure*, 85 F.R.D. 521 (1979) [hereinafter cited as *1979 Amendments*].

[299] *See, e.g.*, Brazil, *Civil Discovery: How Bad Are the Problems?*, 67 A.B.A.J. 450 (1981); Shapiro, *Some Problems of Discovery in an Adversary System*, 63 MINN. L. REV. 1055 (1979); Note, *The Emerging Deterrence Orientation in the Imposition of Discovery Sanctions*, 91 HARV. L. REV. 1033 (1978).

[300] FED. R. CIV. P. 26(b); *see generally* Friedenthal, *A Divided Supreme Court Adopts Discovery Amendments to the Federal Rules of Civil Procedure*, 69 CALIF. L. REV. 806 (1981) (discussing 1979 Amendments and the Advisory Committee's decision to leave Rule 26(b) intact).

[301] *See 1979 Amendments*, *supra* note 298, at 539.

[302] *Id.* at 542.

[303] *Id.* at 526, *citing* P. CONNOLLY, E. HOLLEMAN & M. KUHLMAN, JUDICIAL CONTROLS AND THE CIVIL LITIGATIVE PROCESS: DISCOVERY (Federal Judicial Center Study 1978).

[304] *1979 Amendments*, *supra* note 298, at 541.

[305] Pollack, *Discovery—Its Abuse and Correction*, 80 F.R.D. 219, 220 (1978); *see also* Becker, *Modern Discovery: Promoting Efficient Use and Preventing Abuse of Discovery in the Roscoe Pound Tradition*, 78 F.R.D. 267, 267–69, 275 (1978).

[306] Discussion Guide, *supra* note 282, at 10.

"Defamation and Privacy Cases."[307] The amendment would mandate conferences before commencement of discovery whenever a defamation or privacy plaintiff sought discovery under *Herbert*. At the conference, the parties would agree to a discovery schedule requiring that discovery initially be limited to issues other than actual malice. Further discovery would be permitted only if the plaintiff could make a prima facie showing of all elements of the cause of action other than actual malice.[308]

The Advisory Committee rejected this proposal because it seemed to use rulemaking to circumvent the directives of the Supreme Court. Justice Brennan's dissent in *Herbert* suggested a similar bifurcated procedure,[309] but the majority found the plan unpersuasive,[310] preferring to leave the order of discovery within the discretion of the trial court.[311] In addition, the RTNDA amendment would now be superfluous. Rule 26(f), which had not been approved when the RTNDA made its proposal,[312] already permits discovery conferences.[313] The only difference is that the RTNDA proposal, by compelling the conference, would eliminate the trial judge's discretion and the parties' freedom to work out a solution on their own. Because *Herbert* has not been as intrusive as originally feared, no further changes appear justified until empirical evidence indicates that Rule 26(f) is insufficient to control orderly discovery.

The last alternative considered and rejected without comment by the Advisory Committee was to add a paragraph to its Note to Rule 26(f) similar to the following:

> In defamation and privacy actions in which the plaintiff seeks to inquire into matters relevant to the element of "actual malice," discovery abuse by plaintiffs may pose a particular threat to first amendment interests. To reduce press self-censorship induced by media fears of litigation expense, district judges should make certain that the discovery provisions are employed so as to ensure protection of the editorial process from unnecessary or unjustified intrusions. In

---

[307] *Id.*

[308] *Id.* at 11. RTNDA also proposed in the alternative that the Federal Rules of Evidence be amended to grant libel defendants the privilege to refrain from testifying until the plaintiff made a prima facie showing of all elements other than actual malice. The proposal was not adopted. *Id.*

[309] *Herbert*, 441 U.S. at 198–99 (Brennan, J., dissenting).

[310] *Id.* at 174–75 n.23 (rejecting use of a bifurcated trial as "overly burdensome" and subject to "intolerable delay").

[311] *Id.* at 177.

[312] Discussion Guide, *supra* note 282, at 12.

[313] *Id.* at 12; *see also* FED. R. CIV. P. 16.

> particular, judges may hold discovery conferences, schedule discovery issue by issue, and grant summary judgment on issues other than actual malice in the interest of safeguarding first amendment values.[314]

This was designed to increase judicial sensitivity to the First Amendment. Case law, however, indicates that the judiciary is already highly sensitive to First Amendment interests.[315] Moreover, it is unlikely that more prodding is needed as the press has not been reticent in vocalizing its need for protection.[316] The Rules should not be cluttered with gratuitous homilies.

There is an additional reason to reject amending the Federal Rules: the Supreme Court is unlikely to approve any such changes. The Court plainly indicated in *Herbert*, *Calder*, and *Keeton* that the media is not constitutionally entitled to special procedural protections beyond the substantive protections afforded by *Sullivan*. Granted, the legislature is empowered to create additional, nonconstitutional protections, but in the present situation where the Supreme Court so strongly rejected unique procedures for defamation actions, a change in the Rules would flout the spirit of the opinions.[317] Because the Supreme Court must approve changes in the Federal Rules before presenting them to Congress,[318] there is the possibility that the Court would not affirm those changes.[319]

Finally, although the various amendments would reduce intrusiveness, as pointed out in Part III, intrusive discovery is not as great a problem for media defendants as is litigation expense. To gain marginal protection from intrusion, the press would have to incur additional expenses from extended motion practice, mandatory meetings, and formalized schedules. Other alternatives are more responsive to both the press's true concerns and the system's inherent needs.

---

[314] Discussion Guide, *supra* note 282, at 16.

[315] *See supra* notes 119–69 and accompanying text.

[316] See *supra* notes 89–118 and accompanying text.

[317] A similar situation arose in congressional hearings concerning the proposed Federal Rules of Evidence. Journalists sought to make the subpoena process "as cumbersome as possible" in order to achieve indirectly the protection from compulsory testimony which the Supreme Court refused to grant directly. Congress rejected the suggestion. *See* 23 C. WRIGHT & K. GRAHAM, FEDERAL PRACTICE AND PROCEDURE § 5426 (1982).

[318] *See* 28 U.S.C. § 2072 (1982).

[319] *1979 Amendments*, *supra* note 298, at 521. For example, Justices Powell, Rehnquist, and Stewart believed that the proposed discovery reforms were insufficient and dissented from the 1980 amendments to the rules. *Id.*

VI. Proposals for Change

While the press is legitimately concerned about the chilling effect of litigation expenses under current libel law, it must look beyond its own interests and seek systemic solutions to the current spate of libel suits if it is to find effective and lasting relief. Previous attempts to remedy defamation law have suffered from their focus on the press's special interest in reducing interference with its editorial processes. By working with the judiciary and the public on shared problems of expense and efficiency, the press can suggest mutually beneficial improvements. The most important step that can be taken, both to reduce interference with the editorial process and to make libel law more efficient, is the adoption of a national statute aimed at the best solution of all—avoiding litigation.[320] In the interim, the press should make better use of existing procedural rules and strive to preserve public access to discovery materials.

A. *Make The Best Possible Use of Existing Discovery Rules*

In *Herbert*, the Supreme Court affirmed the ample powers of the trial judge to manage proceedings,[321] and *Herbert* has been cited frequently in both the media and nonmedia cases as a call for judicial control.[322] In 1983, the Federal Rules of Civil Procedure were amended to promote stronger judicial case management[323] and to discourage overly time-consuming and expensive discovery.[324] Rule 16 was amended to provide for more wide-ranging, discretionary pretrial conferences for scheduling discovery, eliminating frivolous claims and defenses, and expediting the resolution of cases.[325] Rule 26 was amended to increase the judge's control over the scope of discovery and to

---

[320] Although throughout this Note I have urged the press to take various actions, I do not mean to imply that libel is purely the press's problem. The press is merely both an interested party and a force with unique abilities to promote change.

[321] *Herbert*, 441 U.S. at 177.

[322] *See, e.g.*, Marrese v. American Academy of Orthopedic Surgeons, 706 F.2d 1488, 1495 (7th Cir. 1983); J.M. Cleminshaw Co. v. Norwich, 93 F.R.C. 338, 359 (D. Conn. 1981); *In re* Arthur Treacher's Franchisees Litigation, 92 F.R.D. 429, 439–40 (E.D. Pa. 1981).

[323] *See* Fed. R. Civ. P. 16 advisory committee note.

[324] *See* Fed. R. Civ. P. 26 advisory committee note.

[325] Fed. R. Civ. P. 16.

state that discovery shall be limited where it is overly burdensome, expensive, duplicative, or unnecessary.[326]

The managerial powers of the judge cited in *Herbert*, especially as bolstered by the 1983 Amendments, provide media litigants with significant tools with which to limit the burden of discovery.[327] Judges should be encouraged, both as a matter of general policy and in libel suits in particular, to limit discovery into the editorial process to necessary and cost-effective areas.

While media defendants should generally encourage tight judicial control of discovery, media litigants should not be blind to the beneficial uses of broad discovery, and they should take advantage of broad discovery whenever possible. Broad discovery rules provide media with a unique opportunity to acquire voluminous information. In fact, for the media defendant, discovery may be an effective weapon. A plaintiff in a defamation action often has more to fear from discovery than does the defendant. While *Herbert* allows discovery into the internal practices of a media organization, media defendants can inquire at length into the truth of the allegedly libelous statements, and if there is any basis to the defamatory statements, discovery will provide defendants with the opportunity for a full-fledged inquiry.[328]

A media defendant is also uniquely able to disseminate its discovery findings; both its ready access to the channels of mass communication and its familiarity with the processes of news dissemination give it an advantage over most defamation plaintiffs. For example, in *Westmoreland v. C.B.S.*, C.B.S. not only used its own channels for communication, it took the additional step of hiring a public relations firm "to tell reporters what the network lawyers learned from General Westmoreland's aides."[329]

---

[326] FED. R. CIV. P. 26.

[327] Strong judicial management of discovery generally results in significant savings of time and expense for little or no sacrifice of substantive rights. The case disposition time of a managerial judge in a strong control jurisdiction has been estimated to be two years quicker than that of a nonmanagerial judge in a least-controlling jurisdiction. P. CONNOLLY, E. HOLLEMAN & M. KUHLMAN, *supra* note 303, at 69. Significantly, although the case proceeds swiftly, the number of discovery requests and the sequence of requests remains unchanged, and the imposition of controls does not result in greater use of compelling motions. *Id.* at 66.

[328] The press should be concerned about the potential limits to dissemination if for no other reason than in libel suits the threat of "reverse discovery" is a deterrent to many potential plaintiffs. LIBEL LITIGATION 1981, *supra* note 127, at 183.

[329] N.Y. Times, May 22, 1984, at A19, col. 2.

In addition, the benefits to the media of wide-ranging discovery are not limited to cases where the media is a party. The Federal Rules never specifically address third party access to discovery materials, but several courts have determined that absent compelling reasons for denying access to the proceedings, pretrial discovery is available to the public.[330] If a court attempts to restrict third party access or dissemination through a protective order under Rule 26(c),[331] the nonparty press is entitled to a hearing before the protective order is issued.[332] Even if the protective order is granted, whenever information is introduced in evidence at trial it becomes part of the public record.[333] The construction of the Federal Rules to allow use of discovery materials outside the context of litigation has made millions of documents available for inspection by the press and the public.[334] For example, the ultimate utility of the Westmoreland case may be its contribution to providing a more complete record of the events during the Vietnam War.

The controversy surrounding the 1980 amendment of Rule 5 shows that the press has begun to recognize the importance of open discovery. Rule 5 had provided that all discovery materials produced must be filed with the court.[335] The Advisory Committee to the Federal Rules, concerned about both the high cost to litigants of making multiple copies and the courts' increasing storage problems, proposed amending Rule 5 to require filing only if ordered by the court.[336] Negative response to the proposed change, however, led the Advisory Committee to reconsider; Rule 5(d) was amended to allow the court to order "on motion of a party or on its own initiative" that discovery materials not be filed. The presumption in favor of filing was re-

---

[330] American Tel. & Tel. Co. v. Grady, 594 F.2d 594 (7th Cir.), *cert. denied*, 440 U.S. 971 (1979); Johnson Foils, Inc. v. Huyck Corp., 61 F.R.D. 405 (N.D.N.Y. 1973); *see* United States v. IBM Corp., 66 F.R.D. 219, 220 (S.D.N.Y. 1974).

[331] *See infra* notes 341–54 and accompanying text.

[332] *In re* Knoxville News-Sentinel Co., 723 F.2d 470 (6th Cir. 1983).

[333] Cianci v. New Times Publishing Co., 486 F. Supp. 368 (S.D.N.Y. 1979).

[334] *See* Rifkind, *Are We Asking Too Much of Our Courts?*, 70 F.R.D. 96, 107 (1976) ("A foreigner watching the discovery proceedings in a civil suit would never suspect that this country has a highly-prized tradition of privacy."); *see also* Howard & Crowley, *Pleading, Discovery, and Pretrial Procedure for Litigation Against Government Spying*, 55 U. DET. J. URB. L. 931, 962 (1978) (dissemination may be a remedy in its own right to force government officials to stop engaging in unlawful activity).

[335] Fed. R. Civ. P. 5, 398 U.S. 981 (1969); *see* 4 C. WRIGHT & A. MILLER, FEDERAL PRACTICE & PROCEDURE § 1152, at 596–98 (1969).

[336] 4 C. WRIGHT & A. MILLER, FEDERAL PRACTICE & PROCEDURE § 1142, at 214 (Supp. 1983).

tained.[337] The Advisory Committee noted "such materials are sometimes of interest to those who may have no access to them except by a requirement of filing, such as members of a class, litigants similarly situated, or the public generally."[338]

The press responded to the procedural change with vigor. A *New York Times* editorial belittled the court storage problem and interpreted the change as giving judges "the power to prevent public access to a huge number of documents that now belong to the record."[339] Commentators expressed concern that Rule 5(d) does not specify what guidelines the court should use in deciding whether to issue an order, or what interests would compel quashing such an order.[340] The extent to which Rule 5(d) will affect filing is not yet determined; much depends on the discretion of the trial court. The press should take an interest in Rule 5(d)'s implementation and request a hearing or notify the public whenever it feels the free flow of information is unduly sacrificed to efficiency.

Perhaps an even more threatening obstacle to the press's effective use of discovery, both as a litigant and as a third party, is the use of protective orders by courts attempting to protect parties subject to discovery. As a general rule, an individual may use the fruits of discovery for any lawful purpose.[341] However, under Rule 26(c), upon a party's showing of good cause, the court can issue a protective order that discovery be restricted or not be had at all.[342] Orders are frequently imposed, for example, when trade secrets and confidential information are at risk.[343] While before 1984 it was not clear whether a court's power to issue protecive orders applicable to discovery was limited by First Amendment concerns,[344] in a 1984 case,

---

[337] 85 F.R.D. 521, 525 (1979).

[338] *Id.*

[339] *Paper Justice*, N.Y. Times, July 22, 1980, *cited in* 4 C. WRIGHT & A. MILLER, FEDERAL PRACTICE & PROCEDURE: § 1152, at 220 (Supp. 1983).

[340] *See* 4 C. WRIGHT & A. MILLER, FEDERAL PRACTICE & PROCEDURE § 1152, at 220 (Supp. 1983).

[341] Leonia Amusement Corp. v. Loew's, Inc., 18 F.R.D. 503, 508 (S.D.N.Y. 1955).

[342] FED. R. CIV. P. 26(c).

[343] *See, e.g.*, Fed. Open Market Committee of Federal Reserve System v. Merrill, 443 U.S. 340 (1979) (Federal Reserve System compelled to disclose all information except that which might harm the economy); National Utility Service, Inc. v. Northwestern Steel & Wire Co., 426 F.2d 222, 227 (7th Cir. 1970) (plaintiff permitted to request a protective order for trade secrets).

[344] *Compare* International Paper Products v. Koons, 325 F.2d 403, 407–08 (2d Cir. 1963) (First Amendment does not limit a trial court's discretion to restrict the dissemination of information produced through pretrial discovery) *with In re* Halkin, 598 F.2d 176 (D.C. Cir. 1979) (First Amendment protects dissemination of information obtained

Seattle Times Co. v. Rhinehart,[345] the Supreme Court granted trial judges broad discretion to limit the use of information gained through discovery and indicated that protective orders covering information learned through discovery did not merit heightened First Amendment scrutiny.[346]

Justice Powell, writing for the majority in *Rhinehart*, reasoned that because information is obtained through discovery only as "a matter of legislative grace," a trial court can properly condition its order granting a litigant access to information by placing restraints on the way the disclosed material is used.[347] Moreover, the Court found that a protective order was not a "classic

---

through discovery, and to grant a protective order, the court must make an initial inquiry into the nature of the speech interests implicated and then examine three criteria: first, that the harm threatened by dissemination be substantial and serious; second, that the order be drawn narrowly; and third, that the protective order be the least restrictive alternative.)

[345] 104 S. Ct. 2199 (1984).

[346] *Id.* at 2207. The plaintiff in *Rhinehart* was the spiritual leader of a religious group centered in the State of Washington. Over a period of years the Seattle Times published a total of eleven articles about Rhinehart's group, including articles that described seances Rhinehart was paid to conduct, referred to Rhinehart's vacated sodomy conviction, and detailed an "extravaganza" sponsored by Rhinehart at the state penitentiary where the paper said a "chorus line of girls shed their gowns and bikinis and sang." *Id.* at 2202. Rhinehart and other members of the group brought a defamation action in state court. Rhinehart was joined in his suit by five female members of the Foundation who had participated in its presentation at the penitentiary. In addition to the Seattle Times, defendants included the Walla Walla Union-Bulletin, the authors of the articles and the spouses of the authors. *Id.*

Defendants promptly initiated extensive discovery. They deposed Rhinehart and served him with lengthy interrogatories. *Id.* at 2203. When Rhinehart refused to disclose financial information regarding the sources and uses of his foundation's wealth, the identities of its donors during the preceding ten years, and a list of its members during that period, the defendants moved to compel discovery under the state counterpart to Federal Rule of Civil Procedure 37. *Id.*

The trial court essentially granted the motion to compel. The motion required the plaintiffs to identify all donors and the amounts they contributed in the five years preceding the complaint. The plaintiffs also had to disclose such information as was necessary to support their claim of decline in membership. *Id.* at 2203. The plaintiffs, however, subsequently moved for a more specific protective order, and after considering affadavits of several group members averring that public release of the donor lists would hurt membership and subject members to harassment, *id.* at 2204, the court issued an order that limited the defendants' use of information learned through discovery regarding the financial affairs of the plaintiffs, the names and addresses of group members, and the group's contributions to "such use as is necessary to prepare and try the case." *Id.* at 2205 n.8. The order prohibited the defendants from publishing this information or making it available to any other news media, but the order specifically did not apply to information acquired from sources other than discovery. *Id.*

On appeal, the Supreme Court of Washington affirmed the court's protective order. *Rhinehart*, 654 P.2d 673 (Wash. 1982). The Washington court reasoned that although a protective order may fall within the definition of a "prior restraint," the need to preserve "the integrity of the discovery processes is sufficient to meet the 'heavy burden' of justification." *Id.* at 690. In a unanimous decision, the Supreme Court of the United States affirmed. *Rhinehart*, 104 S. Ct. 2199.

[347] Rhinehart, 104 S. Ct. at 2207–08.

prior restraint that requires exacting First Amendment scrutiny" because such an order did not limit the dissemination of material acquired outside the discovery process.[348] The Court also found that the issuance of protective orders generally requires no heightened First Amendment scrutiny.[349] The Court cited *Herbert v. Lando* and its concern for discovery abuse[350] and concluded that the trial court in its discretion has the responsibility for deciding when a protective order was necessary and what degree of protection was required.[351]

While protective orders probably cannot shield the most important information gleaned through discovery for long,[352] the Court's characterization of discovery as essentially "private" may have far-ranging implications.[353] With First Amendment concerns entitled to no special consideration, trial courts may make greater use of protective orders. But by leaving trial judges free to assess the proper weight to be given First Amendment concerns when issuing protective orders, the Court ensured some flexibility in the system. The press should use this flexibility to limit the damage it suffers from protective orders, both by being attuned to the broad range of access issues involved in litigation[354] and by using its abilities to inform the public and the judiciary what the proper weight should be. Nonetheless, protective orders may prove a growing obstacle to the media's use of discovery.

---

[348] *Id.* at 2208, *citing* Gannett Co. v. DePasquale, 443 U.S. 368, 399 (1979) (Powell, J., concurring).

[349] *Rhinehart,* 104 S. Ct. at 2209. The Court noted that heightened scrutiny would ". . . necessitate burdensome evidentiary findings and could lead to time-consuming interlocutory appeals." *Id.* at 2209 n.23.

[350] *Id.* at 2208 n.20.

[351] *Id.* at 2209. Justice Brennan, in a concurring opinion joined by J. Marshall, emphasized his view that protective orders are subject to First Amendment scrutiny, yet based his agreement with the Court on the plaintiff's interests in privacy and religious freedom, which, he argued, overcame the defendant's First Amendment interests. *Id.* at 2210 (Brennan, J., concurring).

[352] *See* N.Y. Times, May 22, 1984, at A19, col. 2 (statement of James C. Goodale, Esq.). This information will most likely come to light at trial.

[353] *Rhinehart,* 104 S. Ct. at 2207–08. The Court reasoned that protective orders do not violate the Constitution when they apply to discovery materials because those materials are "not public components of a civil trial." *Id.* at 2207. To support its position, the Court noted that FED. R. CIV. P. 5(d) permits a court to restrict the public filing of discovery documents. *Id.* at 2207–08, n.19. Thus, the fears of the press that the 1980 Amendment to FED. R. CIV. P. 5(d) would be used to limit the information available to the public seems to have been realized. *See supra* notes 335–40 and accompanying text.

[354] *See, e.g.,* Press-Enterprise Co. v. Superior Court of California, 104 S. Ct. 819 (1984) (right of access at voir dire proceeding); Gannett v. DePasquale, 443 U.S. 368 (1979) (no right of access at preliminary hearing); Times Newspapers, Ltd. v. McDonnell Douglas Corp., 387 F. Supp. 189 (C.D. Ca. 1974) (no right of access to deposition).

B. *A National "Correction Statute"*

The problem with libel litigation is that, in addition to being expensive and intrusive, it does not work. It does not remedy the wrong it was created to right. A plaintiff whose reputation has been injured can look forward to years of litigation with little likelihood of ultimate success. Even if a story is false, without a showing of actual malice a public figure's reputation remains unvindicated. Moreover, it is doubtful that a judicial pronouncement several years after an injury can ever realistically remedy an individual's public image. Monetary damages are a consolation prize, but they are external to the true injury. Defendants, meanwhile, suffer the chilling effect of expensive litigation, and the judicial system expends valuable resources for statistically meager results.

This is particularly troublesome because, unlike many other tort actions that depend on monetary damages as a substitute for a lost limb or a lost life that can never actually be *restored*, a defamation plaintiff *can* be made whole. A libeled plaintiff is injured when his or her reputation is tarnished in a public forum. When a potential defendant is a media organization, however, the organization has continued access to that forum and thus the capacity to rectify the harm. A public injury demands a public remedy, but litigation is an expensive and ineffective solution. Most "solutions" to the perceived crisis in defamation law have the same flaw—they presuppose litigation. Changes in the substantive law or Federal Rules may be incrementally helpful, but the true answer is to avoid litigation completely. Potential litigants need stronger incentives to settle their differences without resorting to judicial involvement.[355]

I propose a national Correction Statute applicable to all media organizations in any action for damages for publication of a defamatory statement.[356] The statute would limit plaintiff's recovery to special damages unless the plaintiff can allege and prove that he or she made a sufficient request for correction and that the media organization failed to make conspicuous and timely publication of the correction.[357]

---

[355] *But see* Fiss, *Against Settlement*, 93 YALE L.J. 6 (1984) (arguing that settlement is often unfair and exacerbates existing inequalities).

[356] *See also* Note, *An Alternative to the General-Damage Award for Defamation*, 20 STAN. L. REV. 504 (1968) (proposes model state statute that provides for compulsory reply or retraction).

[357] The Correction Statute appears as an appendix to this Note.

The statute is not special interest legislation, but is designed to promote the shared interest of plaintiffs, defendants, and the public in the truth. Studies indicate that a libel victim is usually more interested in quick vindication than monetary damages.[358] The desire for a monetary windfall or to punish defendants with a high award commonly develops later as litigation drags on.[359] A correction soon after the original publication calculated to attract similar attention from the same audience would restore plaintiff's reputation and eliminate the incentive to sue. Moreover, correction provides an alternative remedy for the plaintiff who cannot afford to finance a full-fledged defamation action.

Similarly, the press has no interest in perpetuating fallacies. Professional standards codes clearly stress the importance of correcting errors.[360] Nevertheless, the press is widely perceived as, and is criticized for, failing to admit to mistakes.[361] In the past, this may have been a result of the small likelihood that a plaintiff would win or be awarded much money if a suit were brought. In recent years, however, potential liability has skyrocketed, and publishers are three times more likely to print a correction than they were ten years ago.[362] Enactment of the Correction Statute would encourage this trend and would improve public opinion of the press. The press could simultaneously serve its individual interest in avoiding liability, its professional interest in good reporting, and the public interest in truth.

---

[358] TIME, Mar. 19, 1984 at 64, *citing* Gilbert Cranberg, Director of the Libel Research Project, University of Iowa.

[359] *Id.*

[360] *See, e.g.*, Franklin, *Good Names and Bad Law*, *supra* note 186, at 31 n.138, *citing* J. HULTENG, PLAYING IT STRAIGHT 77–86 app. (1981). Examples of standards adopted by professional organizations include:

(1) Associated Press Managing Editors, Code of Ethics (newspaper "should admit all substantive errors and correct them promptly and prominently");
(2) United Press International—A Policy Statement ("[c]orrect all errors swiftly and fully, showing what is being corrected and why");
(3) The Society of Professional Journalists, Sigma Delta Chi, Code of Ethics (it is "the duty of news media to make prompt and complete correction of their errors");
(4) American Society of Newspaper Editors, Statement of Principles ("[s]ignificant errors of fact, as well as errors of omission, should be corrected promptly and prominently").

Id.

[361] *The Eight Most Common Complaints*, EDITOR & PUBLISHER, Mar. 19, 1983, at 40.

[362] *More Papers Admitting Their Errors*, L.A. Times, Aug. 18, 1983, at 1, col. 1. A 1973 survey of newspapers with a circulation over 100,000 by the American Newspaper Publishers Association reported only twenty-four percent regularly ran correction notices; by 1983, however, the percentage had tripled.

It is vital that there be a national statute to ensure uniformity because the Supreme Court's recent decisions in *Calder* and *Keeton* provide plaintiffs with an unprecedented opportunity for forum shopping. National media organizations and their employees can now be sued in almost any state. While many states have some form of correction statute,[363] the statutes are often weak,[364] overly complex,[365] or limited in scope.[366] Plaintiffs suing defendants with nationwide circulations are free to pick and to choose among these laws, in effect nullifying those correction statutes unfavorable to plaintiffs as well as those shield laws that might protect defendants.[367] There may be a concern that a national statute enacted by Congress under its Commerce Clause[368] powers would not apply to intrastate defendants who can least afford litigation. A national statute, however, would still be important because (1) "interstate commerce" traditionally has been interpreted broadly;[369] (2) the national Correction Statute would be a model for state counterparts; and (3) intrastate defendants are by definition not subject to the evils of forum shopping, which the statute is designed to alleviate.

A national statute preempting state law would also circumvent the problem of state courts striking down similar statutes under

[363] *See, e.g.*, ALA. CODE §§ 6-5-184 to -186 (1977); ARIZ. REV. STAT. ANN. §§ 12-653.01–.05 (1982); CAL. CIV. CODE § 48a (1982); CONN. GEN. STAT. § 52-237 (1983); FLA. STAT. §§ 770.01–.02 (1983); GA. CODE § 51-5-11 (1982); IDAHO CODE § 6-712 (1979); IND. CODE ANN. §§ 34-4-14-1, -15-1 (West 1983); IOWA CODE §§ 659.2, 659.3 (1983); KY. REV. STAT. ANN. §§ 411.051–.062 (Bobbs-Merrill 1972); ME. REV. STAT. ANN. tit. 14, § 153 (Supp. 1984–85); MASS. GEN. LAWS ANN. ch. 231, § 93 (West 1959); MICH. COMP. LAWS § 600.2911(2)(b) (1979); MINN. STAT. § 548.06 (1982); MISS. CODE ANN. § 95-1-5 (1973); MONT. CODE ANN. §§ 27-1-818 to -1-821 (1983); NEB. REV. STAT. § 25-840.01 (1979); NEV. REV. STAT. §§ 41.331–.338 (1969); N.J. STAT. ANN. § 2A:43-2 (West 1952); N.C. GEN. STAT. §§ 99-2 (1979); N.D. CENT. CODE § 14-02-08 (Supp. 1983); OHIO REV. CODE ANN. §§ 2739.13–.18, 2739.99 (Page 1981); OKLA. STAT. tit. 12, § 1446a (1981); OR. REV. STAT. §§ 30.155–.175 (1983); S.D. CODIFIED LAWS ANN. §§ 20-11-7 to 20-11-8 (1979); TENN. CODE ANN. § 29-24-103 (1980); TEX. REV. CIV. STAT. ANN. art. 5431 (Vernon 1958); UTAH CODE ANN. §§ 45-2-1, 45-2-1.5 (1981); VA. CODE §§ 58.01-46 to -48; W. VA. CODE § 57-2-4 (1966); WIS. STAT. § 895.05(2) (1979–1980).

[364] *See, e.g.*, W. VA. CODE § 57-2-4 (1966) (timely offer of apology is mitigation of damages).

[365] *See, e.g.*, MISS. CODE ANN. § 95-1-5 (1973). The statute places upon the defendant the burden of proving that the defamatory statement was made in good faith, that it was the result of an honest mistake of fact, that there were reasonable grounds for believing the statement was true, and that a retraction was published within ten days of written notice of the falsehood. The statute does not apply to candidates for public office or to statements of opinion.

[366] *See, e.g.*, MINN. STAT. § 548.06 (1982) (applies only to newspapers).

[367] *See supra* text accompanying notes 148–159.

[368] U.S. CONST. art. I, § 8.

[369] *See, e.g.*, Wickard v. Filburn, 317 U.S. 111 (1942).

state constitutions[370] and would prevent state courts from eviscerating state statutes through interpretation.[371]

The Correction Statute proposed by this Note also would circumvent the First Amendment stumbling blocks raised by the Supreme Court in *Miami Herald Publishing Co. v. Tornillo*.[372] In *Tornillo* the Court held unconstitutional a Florida statute giving political candidates a "right to reply" to derogatory statements by newspapers. Under the statute, any newspaper assailing the personal characteristics of a candidate, or attacking a candidate's record while in office, or providing any person free space to do so, could be forced to publish, free of cost, any reply statement of similar length that the candidate cared to make.[373] In an opinion written by Chief Justice Burger, the Court held that the statute unconstitutionally interfered with the "exercise of editorial control and judgment"[374] because it compelled "editors or publishers to publish that which 'reason' tells them should not be published."[375] A broad reading of the Court's language could give rise to the inference that any statute that pressures a publisher or broadcaster to make statements it would not otherwise make is unconstitutional. At least one commentator has noted this possible interpretation of the case.[376]

Even if this interpretation has value in some circumstances, it is invalid if applied to a statute "that permit[s] defamers to

---

[370] *See, e.g.*, Madison v. Yunker, 589 P.2d 126 (Mont. 1978) (holding that a state correction statute violated the state constitution insofar as it barred a plaintiff from seeking remedy in the courts for a violation of the law); *see also* Byers v. Meridian Printing Co., 84 Ohio St. 408, 95 N.E. 917 (1911) (holding that a state correction statute violated the state constitution by depriving plaintiffs of a remedy for a violation of substantive law without due process of law); Hanson v. Krehbiel, 68 Kan. 670, 75 P. 1041 (1904) (holding that a state correction statute prohibiting the recovery of punitive damages from a defendant who took specified steps to correct a misstatement violated the state constitution by depriving plaintiffs of the right to a remedy by due course of law).

[371] *See, e.g.*, Gersten v. Newark Morning Ledger Co., 52 N.J. Super. 152, 145 A.2d 56 (1958); Miami Herald Publishing Co. v. Brown, 66 So.2d 679 (Fla. 1953).

[372] 418 U.S. 241.

[373] 1913 Fla. Laws 274 (codified as amended at FLA. STAT. § 104.38) (repealed 1975), *reprinted in Tornillo*, 418 U.S. at 244–45 n.2.

[374] *Tornillo*, 418 U.S. at 258.

[375] *Id.* at 256.

[376] C. MORRIS & C.F. MORRIS, MORRIS ON TORTS 380 (2d ed. 1980) ("The Burger opinion may imply that all retraction statutes are also unconstitutional: any legal pressure put upon a newspaper to publish a retraction is an 'intrusion into the function of editors' and may distort 'the exercise of editorial judgment.'"). *But see Tornillo*, 418 U.S. at 258 (Brennan, J., concurring) ("the Court's opinion . . . addresses only 'right of reply' statutes and implies no view upon the constitutionality of 'retraction' statutes affording plaintiffs able to prove defamatory falsehoods a statutory action to require publication of retraction").

ignore appeals for retraction."[377] The possibility of incurring punitive damages will, of course, sometimes motivate the media to publish corrections it would not publish otherwise, but the offer of increased protection from catastrophic damages offered by this Correction Statute cannot be equated with the compelled publication struck down by *Tornillo*. The statute challenged in *Tornillo* required newspapers to publish responses—that is, to "open their columns, gratis, for replies."[378] In sharp contrast, this Correction Statute makes publication of a correction notice the choice of the alleged defamer: the media defendant need not publish anything, and it may ignore the plaintiff's request for any reason. Nor does the statute provide any penalties that do not already exist in state libel law. If it constitutes unconstitutional coercion, then it is only because the already existing threat of punitive damages is too great an interference with the editiorial process.

In addition, publication of a correction notice is not nearly as intrusive as the replies required in the *Tornillo* statute. While the latter were required for all derogatory speech, whether or not defamatory, the Correction Statute will almost certainly operate only on clearly defamatory statements. A correction is publishable only at the option of the alleged defamer, and it presumably will be issued only where the publisher believes it will be held liable by a jury. In addition, the alleged defamer may fulfill its obligation by publishing a self-drafted correction notice, while under the *Tornillo* statute newspapers were forced to publish replies drafted by the candidates. The voluntary nature of the correction required by this statute, as well as its limited intrusiveness and the fact that it represents greater protection for the media should circumvent the *Tornillo* decision and make the statute constitutional.

Proponents of the Correction Statute should also stress that it provides plaintiffs with a new remedy, rather than taking away an old one. This approach has been used successfully to defend similar state statutes from constitutional attack. For example, in a thoughtful opinion upholding an Oregon retraction statute, the court in *Davidson v. Rogers*[379] held that the state constitution's guarantee of a remedy "does not specify that the remedy

---

[377] C. Morris & C.F. Morris, *supra* note 376, at 380.
[378] *Id.*
[379] 281 Or. 219, 574 P.2d 624 (Or. 1978) (Holman, J.).

need be the same as was available at common law at the time of the adoption of the constitution."[380] The court viewed retraction as a substitute remedy that "[a]s a practical matter, . . . can come nearer to restoring an injured reputation than can money."[381] It concluded:

> If the specific remedies available at common law were frozen at the adoption of Oregon's Constitution, the legislature would have been helpless to enact limitations upon actions such as those protected by the Workman's Compensation Law and the guest passenger statute, or to concern itself with other similar matters about which it is usual for legislatures to take action.[382]

The Correction Statute is in effect taking the protections afforded under the common law[383] and updating them to reflect new circumstances.

The title of the statute is "Correction" and not "Retraction" because terminology is extremely important if all interested parties are to recognize the statute as a remedy, not a punishment. *Retraction* has a more antagonistic connotation than *correction.* The goal is not for the press to admit a mistake, but to correct a public misconception. Following this reasoning, I reject those state statutes that insist "RETRACTION" be printed in large type as some sort of literary dunce cap.[384] Similarly, the Correction Statute uses the phrase "request a correction" rather than "demand a correction."[385] Power struggles among the parties do little to further the public interest in truth.

Using nonantagonistic terminology also has an important practical component in that the statute does not compel the press to make a correction. Although one state statute has a mandatory retraction provision,[386] this would probably be unconstitutional under *Miami Herald Publishing Co. v. Tornillo.*[387] Moreover, compulsory correction goes against the philosophy

---

[380] *Id.* at 625.

[381] *Id.* In a concurring opinion, Judge Linde rejected the alternative remedy theory but upheld its constitutionality under the state's power to restrict the financial scope of recovery. *Id.* at 626. Judge Lent, in dissent, rejected any limitation on the damage remedy. *Id.* at 626–33.

[382] *Id.* at 625.

[383] At common law, a retraction mitigated damages. *See* W. PROSSER, HANDBOOK OF THE LAW OF TORTS § 116, at 799 (4th ed. 1971).

[384] *See, e.g.,* MINN. STAT. § 548.06 (1982) (requiring 18 point type or larger).

[385] *See, e.g.,* KY. REV. STAT. ANN. §§ 411.051–.062 (Bobbs-Merrill 1972) (punitive damages recoverable only after "sufficient demand for correction").

[386] OHIO REV. CODE ANN. §§ 2739.13–.99 (Page 1981).

[387] 418 U.S. 241 (1974). *See also supra* notes 372–76 and accompanying text.

of the statute in promoting truth through enlightened self-interest.

The Correction Statute should encompass all media organizations that have continued access to the forum where the original statement was made. Some states limit their coverage to newspapers,[388] on the belief that only newspapers with the pressure of a daily publication schedule need protection from intrusion. Under such a law, however, rights depend upon a categorical determination by the judiciary of what publications qualify as newspapers. For example, in *Burnett v. National Enquirer*,[389] the Enquirer published a retraction but the court held that the state retraction statute was inapplicable. The court found that Burnett was not limited to special damages as the correction statute provided because the *Enquirer* was a magazine and not a newspaper.[390] More troublesome, however, is that limiting the scope of the statute to newspapers implies that the statute is some sort of privilege from suit based on the exigencies of publication. This obscures its true role as a more efficient and effective remedy for injury and a means to promote investigation into the truth. Any media organization that has continued access to the forum of injury should be entitled to correct an error.[391]

The request for correction should be signed and in writing, and it should specify the statement or statements claimed to be false or defamatory, explain why they are false, and state the truth. The request should contain substantially the same evidence as to truth as would be presented at trial with the names of sources and references listed. Plaintiffs have the burden at trial of showing that they complied with the statute in good faith. This will ensure that the media organization has enough information so that it can accurately assess the truth.

A correction will be sufficient if it is a good faith effort to state the true facts in as conspicuous a manner as the original statements. Merely stating that the original report was incorrect is insufficient to reveal the truth. Media organizations, furthermore, should be careful not to repeat the defamatory statements

---

[388] *See, e.g.*, MINN. STAT. § 548.06 (1982).

[389] 144 Cal. App. 3d 991, 193 Cal. Rptr. 206 (1983), *appeal dismissed*, 104 S. Ct. 1260 (1984).

[390] *Id.* at 214.

[391] Book publishers and private citizens do not have the requisite continued access needed to come under the statute, but any attempts they make to mitigate damages will be recognized under the statute. See Appendix, Section 11.

in the correction if this could reasonably be interpreted as perpetuating the falsity. The media organization also has the option of giving the injured party an opportunity to reply in a similarly conspicuous forum. This alternative promotes the marketplace of ideas and may be an attractive alternative to the organization if it does not want to admit that it committed an error. If the injured party, however, declines to make the reply, the media organization must still make the correction within the specified time period to fulfill the requirements of the statute.

The more time that elapses between publication and correction, the less likely it is that the correction will restore a plaintiff's reputation. Therefore, time limits are imposed both on when the request for correction should be made and when the correction must be circulated. The plaintiff simply requests correction within twenty days of publication or widespread circulation of the original publication. This should provide plaintiffs with an adequate opportunity to prepare evidence as to the truth.[392] Some states insist that a defendant make a correction as quickly as within one business day for a broadcast or within three days for a newspaper.[393] The Correction Statute recognizes the importance of speed, but chooses to use a slightly longer correction period to facilitate proper investigation of the truth.

Malice is irrelevant to the Correction Statute. The media organization's intention at the time of original publication does not matter as long as the correction is properly made.[394] To permit inquiry as to malice would only complicate the proceeding and jeopardize the statute's effectiveness as a tool for the truth. Similarly, neither the failure to publish a correction nor publishing a correction will negate or promote an inference of malice. Moreover, in order to encourage a full investigation of the truth, there is a limited privilege preventing discovery of the investigation process or its fruits from the time that the request

---

[392] Some states require that retractions be requested at a specified time before filing suit. *See, e.g.*, TENN. CODE ANN. § 29-24-103 (1980) (five days before filing). This, however, could be many years after the defamatory statement was published, and so it would not help an injured plaintiff to repair his reputation, nor would it promote the investigation into the truth. Other states specify that the request must be made a certain time after "knowledge" of the publication. *See, e.g.*, NEB. REV. STAT. § 25-840.01 (1979) (twenty days after knowledge); NEV. REV. STAT. § 41.336 (1969) (ninety days after knowledge). Calculating when "knowledge" occurred, however, can be difficult. By using the "twenty days of publication or widespread circulation" test, the Correction Statute emphasizes when the plaintiff should have known.

[393] KY. REV. STAT. ANN. §§ 411.051–.062 (Bobbs-Merrill 1972).

[394] *See* CAL. CIV. CODE § 48a (1982).

for correction is presented to the time that the decision whether to make the correction is made.

Further, because truth is the primary goal of the statute, if a media organization realizes on its own that it published a misrepresentation, and it corrects that error within twenty days of publication or widespread circulation, plaintiff will also be limited to special damages in a subsequent suit for libel.[395]

The Correction Statute recognizes, however, that in certain instances stating the truth is not enough. Specifically, if a candidate for public office is defamed and the correction is not made within a reasonable time before the election, the candidate may sue for general damages.[396] Similarly, despite a timely correction, an injured party may still suffer pecuniary loss. The Statute entitles the party to sue for special damages, and it encourages media organizations that have made corrections to make a good faith effort to settle claims for special damages out of court.[397]

The Correction Statute will ensure that potential plaintiffs and defendants meet to work out their differences and to disseminate the truth before bringing suit. Some courts have criticized similar statutes as giving defendants two chances to be *wrong* before the start of litigation—first, by publishing a defamatory falsehood, and second, by refusing to correct the error.[398] The statute instead should be viewed as instituting an opportunity to be *right* so that litigation can be avoided.

VII. Conclusion

The eleven month period between the Supreme Court's decisions in *Calder* and *Keeton* and the settlement of the West-

[395] *Id.*

[396] *Cf.* ARIZ. REV. STAT. §§ 12-653.01-.05 (1982) (candidate may recover special, general, and exemplary damages if libelous publication is made within thirty days of election and publication is "designed to in any way influence the results of such election"); MISS. CODE ANN. § 95-1-5 (1973) (candidate not limited to actual damages if libelous publication is made within ten days of election or if made in editorial or regular opinion column); N.D. CENT. CODE § 14-02-08 (Supp. 1983) (editorial retraction at least three days before election for daily and at least ten days before election for weekly); S.D. CODIFIED LAWS ANN. §§ 20-11-7, -8 (1979) (editorial retraction at least three days before election for daily and at least ten days before election for weekly); UTAH CODE ANN. §§ 45-2-1, -1.5 (1981) (editorial retraction at least three days before election or nominating convention for weekly and at least five days before election or nominating convention for daily).

[397] The statute recognizes that plaintiffs may be reluctant to go to court if their sole recovery is special damages. The statute cannot require a settlement, but it can encourage it.

[398] *See, e.g.*, Bank of Oregon v. Independent News, 670 P.2d 616, 627 (Or. App. 1983), *aff'd*, 693 P.2d 35 (Or. 1985).

moreland case has been the most significant time in the history of defamation law since the Court decided *New York Times v. Sullivan*. This period saw the extension in *Calder*, *Keeton*, *Bose*, and *Rhinehart* of the reasoning first put forth in *Herbert* and *Hutchinson*. In those cases, the Court has made it clear that the First Amendment does not entitle the press to special procedural protection in defamation cases beyond that afforded by *Sullivan*. The year was also a time when celebrated trials[399] brought to the limelight the inefficiencies of defamation litigation, where lengthy and expensive pretrial and trial procedures fostered by the constitutional interpretation of *Calder*, *Keeton*, *Hutchinson*, and *Herbert* yield inconclusive results because of protections afforded by *Sullivan* and *Bose*.[400] The public has been saturated with discussions of the inadequacies of modern libel law and made aware that libel suits cost both defendants and plaintiffs a great deal of money, that they often fail to vindicate a plaintiff, and that they tend to chill the media's editorial processes regardless of outcome.[401] The newly heightened awareness of these problems makes this the time to push forward significant improvements in libel law.

Absolutist rhetoric arguing for unique privileges for the Fourth Estate has little use at a time when the Supreme Court has so firmly stated it will not grant the press any special procedural protections or change the substantive constitutional standard. Although the avenue of special protections from the courts appears to be closed, many other paths exist to improve the current system; and the press is uniquely situated to promote progress along those paths. Once the press recognizes that relief from burdensome litigation will only be possible through systemic changes that serve all interests, it can devote some of its energies to promoting these changes. The most important such change is adopting a national Correction Statute that would reduce the new potential for forum shopping, while providing incentives to avoid litigation and to promote the truth. There is surprisingly little legislation covering defamation law.[402] The time is now ripe for change.

---

[399] *See supra* notes 130–32, 201, 258, 329 and accompanying text & note 261.

[400] *See, e.g.*, N.Y. Times, Jan. 25, 1985, at B4, col. 1; NEWSWEEK, Feb. 4, 1985, at 55.

[401] *See, e.g.*, Wash. Post, Feb. 19, 1985, at A10, col. 1, A11, col. 5; N.Y. Times, Feb. 18, 1985, at 1, col. 1; N.Y. Times, Feb. 19, 1985, at B6, col. 1, B7, col. 1; NEWSWEEK, Feb. 4, 1985, at 52–58.

[402] *See generally* R. SACK, LIBEL, SLANDER, AND RELATED PROBLEMS (1980).

APPENDIX

DEFAMATION CORRECTION STATUTE

SECTION 1.

In any action against a media organization for damages for the publication of a defamatory statement through any medium, the plaintiff shall recover no more than special damages unless the plaintiff shall allege and prove that the plaintiff made a timely sufficient request for correction and that the media organization failed to make conspicuous and timely publication of the correction.

SECTION 2.

A "media organization" shall be interpreted broadly to include all publishers and broadcasters that have continued access to a medium of expression.

SECTION 3.

A "sufficient request for correction" is a good faith request for correction that is in writing; that is signed by the plaintiff or his or her duly authorized attorney or agent; that specifies the statement or statements claimed to be false and defamatory, states how they are false, sets forth the true facts, is accompanied with substantially the same evidence as to truth that would be presented at trial, including the identity of sources or references, such that the media organization can reasonably assess the truth; and that is delivered to the defendant within a timely period.

SECTION 4.

A "timely" request for correction is a request made within twenty (20) days of publication or twenty (20) days after the publication was widely circulated, whichever is later.

SECTION 5.

A "correction" is either (a) the good faith publication of the true facts, or (b) the good faith publication of the plaintiff's statement of the true facts in a reply written by the plaintiff but exclusive of any portions that are defamatory of another, obscene, or otherwise improper for publication. If the request for correction has specified two or more statements

as false and defamatory, the correction may deal with some of such statements pursuant to (a) above and with other of such statements pursuant to (b) above. The defendant has the option of using either (a) or (b); however, if the plaintiff declines to make a reply under (b) the defendant must still make a correction pursuant to (a) in order to fulfill the requirements of the statute.

SECTION 6.

A "conspicuous" publication in a visual or sound television or radio broadcast is a good faith publication that is broadcast at substantially the same time of day, and with the same sending power, as the statement(s) specified to be false and defamatory in the request for correction. A "conspicuous" publication for a print publisher is a good faith publication that is printed in substantially the same manner as the statement(s) specified to be false and defamatory in the request for correction. A publication in a particular manner that is agreeable to the plaintiff shall in any event be deemed "conspicuous."

SECTION 7.

A "timely" publication for a media publication published with a frequency of less than thirty (30) days is within twenty (20) days of receipt of the request for correction. A "timely" publication for a media publication published with a frequency of thirty (30) days or greater is either within twenty (20) days of receipt of the request for correction or in the next edition of the publication, whichever is later, provided that publication is within six (6) months of the request for correction. A publication on a particular day that is agreeable to the plaintiff shall in any event be deemed "timely."

SECTION 8.

A correction of a statement(s) that is (are) false and defamatory about a candidate for public office is presumed not to be "timely" unless it is published at least three (3) days before the election or on a particular day agreeable to the plaintiff.

SECTION 9.

A good faith conspicuous correction published by the defendant before the plaintiff makes a request for correction shall have the same force and effect as though such correction had been published in a timely manner after a request by the plaintiff.

SECTION 10.

There is a privilege from discovery into the investigative and editorial process of all materials, communications, and thoughts of the defendant or any of its employees from the time the request for correction is delivered and the decision whether to publish the correction is made.

SECTION 11.

A book publisher, media organization that publishes with a frequency of greater than six (6) months, or any person or organization, which publishes a statement(s) that is (are) false and defamatory, and then makes a good faith publication of the true facts in some medium of expression, shall have that publication deemed to mitigate damages.

SECTION 12.

"Special damages" are pecuniary damages that the plaintiff alleges and proves that the plaintiff has suffered with respect to the plaintiff's property, business, trade, profession, or occupation (including such amounts of money as the plaintiff alleges and proves he has expended, exclusive of attorneys' fees, as a proximate result of the alleged defamation), and no other.

SECTION 13.

The media organization and the plaintiff shall make a good faith effort to settle a claim for special damages out of court.

SECTION 14.
This statute is enacted through Congress's power under the Commerce Clause and is preemptive of state law.

NOTE

TELEPHONE PORNOGRAPHY: FIRST AMENDMENT CONSTRAINTS ON SHIELDING CHILDREN FROM DIAL-A-PORN

JOHN C. CLEARY*

Dial-a-porn services offer sexually explicit entertainment to all callers. In 1983, public concern with children's access to dial-a-porn led to an amendment of the Communications Act of 1934. The amendment restricts transmission by telephone of obscene or indecent speech to persons under eighteen years of age and requires the Federal Communications Commission (FCC) to promulgate regulations by which dial-a-porn sponsors could screen out underaged callers. The United States Court of Appeals for the Second Circuit held that the screening regulations, subsequently promulgated by the FCC, were unconstitutional content-based regulations of speech.

In this Note, Mr. Cleary examines the constitutionality of the new federal statute regulating dial-a-porn. He analyzes the free speech rights of children and of adults and the limits of such rights in relation to obscene or indecent communications. The author proposes that certain content-based regulations of speech may be permissible when a communication mode's potential audience cannot be segregated into adults and children. The author thus concludes that such regulations should be permissible in the dial-a-porn context so that children may be shielded from telephone pornography.

The First Amendment provides that "Congress shall make no law . . . abridging the freedom of speech . . . ."[1] Enforcing this command for audio and video technologies as diverse as sound trucks, radio, television, billboards, and telephones is an increasingly problematic task for the legal system.[2] The intrusive features of modern communication technologies implicate significant nonspeech interests, such as abating nuisances, protecting unwilling listeners, and sheltering children. Reconciling these interests with constitutional guarantees challenges the creativity and resourcefulness of the legal system and ultimately measures society's commitment to freedom of speech itself.

---

* Associate, LeBoeuf, Lamb, Leiby & MacRae, New York, N.Y. B.S., Cornell University, 1981; J.D., Harvard University, 1984.

[1] U.S. CONST. amend. I.

[2] *See generally* I. DE SOLA POOL, TECHNOLOGIES OF FREEDOM 1–10, 129–35 (1983) (noting that new communication technologies do not receive all of the legal immunities that traditional communication modes have received).

The use of automatic telephone answering equipment to offer sexually oriented entertainment to all callers, including children, places free speech in conflict with other societal interests. In 1983, Congress enacted legislation to protect children from exposure to such services, making the transmission of sexually oriented messages to minors over such a system a federal crime.[3] Because sponsors of these messages cannot ascertain the ages of callers, the statute may force sponsors to stop communicating with adults or to censor the content of their messages to retain "only what is fit for children."[4] Due to this self-censorship, the new law may violate the longstanding rule of *Butler v. Michigan* that state efforts to protect children cannot limit adult communication to "only what is fit for children."[5]

This Note will analyze the free speech principles underlying the *Butler* doctrine and the limitations placed on *Butler* by the Supreme Court in the context of radio broadcasting.[6] Modifications of *Butler* for application to communication modes whose potential audience cannot be segregated into adults and children will be proposed. This proposal will be applied to the problem of regulating child access to dial-a-porn.

I. Sexually Oriented Telephone Entertainment Services

A. *Factual Background*

The ordinary telephone can now bring an array of sexually oriented entertainment services to virtually every household, office, and telephone booth in America. These services are advertised in newspapers and magazines and cater to male and female, heterosexual and homosexual audiences.[7]

---

[3] Federal Communications Commission Authorization Act of 1983, Pub. L. No. 98-214, § 8(a), 1984 U.S. Code Cong. & Ad. News (97 Stat.) 1467, 1469–70 (to be codified at 47 U.S.C. § 223). The statute is reprinted in the Appendix to this Note. The statute criminalizes sexually oriented communications to "any person under eighteen years of age." 47 U.S.C.A. § 223 (b)(1)(A) (West Supp. 1984). This group will be referred to as minors or children throughout this Note without distinguishing between the two and without intending to implicate any other federal or state definitions of minority or childhood.

[4] Butler v. Michigan, 352 U.S. 380, 383 (1957).

[5] *Id.* at 383.

[6] *See* FCC v. Pacifica Foundation, 438 U.S. 726 (1978).

[7] *See For a Good Time Call . . .* , Forbes, Mar. 28, 1983, at 46; Levey, *A Garbage Call Should Trigger a Separate Check*, Wash. Post, June 16, 1983, at B20, col. 1; *Aural Sex*, Time, May 9, 1983, at 39.

Telephone sex entertainment services are available in two basic modes. Only the mode known as dial-a-porn will be addressed in detail.[8] Dial-a-porn utilizes the telephone system's Mass Announcement Network Service, informally referred to as the dial-it service.[9] Dial-it sponsors can communicate prerecorded messages to 50,000 callers per hour without a busy signal.[10] These messages typically offer callers time and temperature information, weather forecasts, sports scores, or other uncontroversial material. Dial-a-porn messages, on the other hand, generally feature a female voice delivering a "description or depiction of actual or simulated sexual behavior."[11] Callers reach the telephone company's dial-it equipment without the intervention of a live operator and pay only their usual local or long distance telephone rates.[12] Dial-a-porn sponsors, frequently publishers of pornographic magazines, receive cash payments from telephone company revenues, calculated according to local telephone tariffs.[13] The dial-a-porn service in New York City received 800,000 calls daily in May 1983 and 180,000,000 calls

---

[8] A second "for pay" mode of offering telephone sex services requires the customer to pay the sponsor directly, generally by disclosing a credit card number and authorizing appropriate charges, in addition to paying the usual local or long distance telephone rates. *See A Dial-a-Porn Protest*, NEWSWEEK, Sept. 26, 1983, at 40 (charges up to $35 per call); Harv. Crimson Weekly Mag., Mar. 1–7, 1984, at 9, col. 1 (charges range from $15 to $30). Callers reach a live operator who returns the call after verifying the credit card number. *See id.* The performances are conversations lasting up to forty-five minutes between the customer and the performer and, as advertised, are intended to stimulate autoerotic conduct by the customer. *See id.* Access by children is limited because of the credit card requirement. *See* Enforcement of Prohibitions Against the Use of Common Carriers for the Transmission of Obscene Materials (Report and Order), 49 Fed. Reg. 24,996, 25,000–01 (1984) [hereinafter cited as Report and Order], *set aside*, Carlin Communications, Inc. v. FCC, 749 F.2d 113 (2d Cir. 1984).

[9] Enforcement of Prohibitions Against the Use of Common Carriers for the Transmission of Obscene Materials (Notice of Inquiry), 48 Fed. Reg. 43,348, 43,349 (1983) [hereinafter cited as Notice of Inquiry]; Report and Order, *supra* note 8, at 24,996 n.6.

[10] Notice of Inquiry, *supra* note 9, at 43,349 n.3; *see also* 129 CONG. REC. H10,559 (daily ed. Nov. 18, 1983) (statement of Rep. Bliley (R-Va.)).

[11] Notice of Inquiry, *supra* note 9, at 43,349; *see also* N.Y. Times, Dec. 15, 1983, at A25, col. 6.

[12] In the Matter of Enforcement of Prohibitions Against the Use of Common Carriers for the Transmission of Obscene Materials (Further Notice of Inquiry and Notice of Proposed Rulemaking), 49 Fed. Reg. 2124 (proposed Jan. 18, 1984) [hereinafter cited as Further Notice of Inquiry and Notice of Proposed Rulemaking]; *see* N.Y. Times, Dec. 15, 1983, at A25, col. 6. The recent restructuring of the Bell system and related developments have caused some newly independent local telephone companies to begin charging their customers $.50 to $1.00 for each call to dial-it services. *See, e.g.*, Wash. Post, Dec. 28, 1984, at B1, col. 1.

[13] *See* Notice of Inquiry, *supra* note 9, at 43,349 n.7. For example, the intrastate tariff governing dial-it services in New York allots 2¢ of revenue to a dial-it service sponsor for each local or long distance call received. New York Telephone P.S.C. Tariff No. 900, § 13, at 25, *cited in* Notice of Inquiry, *supra* note 9, at 43,349 n.7.

in the year ending February 1984.[14] At the tariff rate of two cents per call, these calls generated revenues for the dial-a-porn sponsor of $16,000 per day and $3,600,000 per year, respectively.[15]

B. *Early Efforts to Control Dial-A-Porn Services*

Public concern with child access to dial-a-porn services led to early efforts to prohibit such services. The dial-a-porn service run in New York by Car-Bon Publishers, Inc., and High Society Magazine, Inc., publishers of an adult magazine named *High Society Live!*, was the target of several of these efforts.[16] Peter F. Cohalan, the county executive of Suffolk County, New York, sued New York Telephone Company, the Federal Communications Commission (FCC), and the sponsors, but the suit was later removed from New York state court to federal court and dismissed for lack of jurisdiction.[17] In March 1983, Cohalan filed a formal complaint with the FCC,[18] charging that New York Telephone Company violated section 223 of the Communications Act of 1934 by knowingly permitting a "telephone under [its] control" to be used to make "any comment, request, suggestion or proposal which is obscene, lewd, lascivious, filthy, or indecent."[19] The FCC referred the matter to the Department

---

[14] *See* Carlin Communications, Inc. v. FCC, 749 F.2d 113, 114 (2d Cir. 1984).

[15] Notice of Inquiry, *supra* note 9, at 43,349 n.7 (citing New York Telephone P.S.C. Tariff No. 900, § 13, at 25); *see also* Carlin Communications, Inc. v. FCC, 749 F.2d 113, 115 (2d Cir. 1984).

[16] Notice of Inquiry, *supra* note 9, at 43,349 nn.3–4. In January 1983, these joint sponsors received a dial-a-porn number from New York Telephone and advertised the number in their magazine. *Id.* at nn.3–4. This service changed messages at least once a day and was available 24 hours a day. *Id.*

[17] Carlin Communications, Inc. v. FCC, 749 F.2d 113, 115 n.4 (2d Cir. 1984).

[18] *In re* Peter F. Cohalan and the County of Suffolk, New York v. New York Telephone Company, FCC File No. E-83-14 (Mar. 31, 1983), *cited in* Report and Order, *supra* note 8, at 24,996 & n.4.

[19] *Id.* § 223(1)(A)(2). The text of original section 223, with minor changes not relevant here, is now section 223(a) of the statute. *See* Appendix, *infra*. The original section was enacted to respond to the problem of "obscene, abusive, or harassing telephone calls" inflicted on others and not to the voluntary reception of sexually explicit communications. H.R. REP. No. 1109, 90th Cong., 2d Sess. 2, *reprinted in* 1968 U.S. CODE CONG. & AD. NEWS 1915, 1915. The 1968 addition of section 223 to the Communications Act of 1934 was the first federal attempt to regulate obscene or harassing interstate communications. Amendment to the Communications Act of 1934, Pub. L. No. 90-299, 82 Stat. 112 (codified at 47 U.S.C. § 223 (1982)).

of Justice because of possible criminal penalties.[20] The Department of Justice decided not to prosecute New York Telephone and sent the case back to the FCC for administrative action in light of the complex statutory and First Amendment issues presented by the case.[21] In September 1983, the FCC published a Notice of Inquiry in the *Federal Register* seeking public comment on whether section 223 applied to dial-a-porn services or could be used to impose liability on common carriers such as New York Telephone.[22] The FCC also sought comments on its authority to regulate the content of telephone communications in general and to regulate dial-a-porn in particular.[23]

II. Federal Response to the Dial-A-Porn Problem

A. *Background*

During the FCC's consideration of the dial-a-porn issue, Representative Thomas J. Bliley (R-Va.)[24] introduced a far-reaching amendment to section 223 as a rider to an FCC appropriations bill then pending before the House Energy and Commerce Committee.[25] As adopted by the committee and reported to the full House, this amendment would have repealed the provision in section 223, which New York Telephone was alleged to have

---

[20] The section provided that violators "shall be fined not more than $500 or imprisoned not more than six months or both." 47 U.S.C.A. § 223 (West Supp. 1984). The sanctions in this section were increased to $50,000 fines in the 1983 amendment. Federal Communications Act of 1983, Pub. L. No. 98-214, 1983 U.S. Code Cong. & Ad. News (97 Stat.) 1469 (to be codified at 47 U.S.C. § 223(a)(1)); *see* Notice of Inquiry, *supra* note 9, at 43,349, 43,350 & nn.1 & 11.

[21] Notice of Inquiry, *supra* note 9, at 43,350 nn.11–12.

[22] Notice of Inquiry, *supra* note 9.

[23] Approximately 20% of the calls to the dial-a-porn service in New York City were reported to be interstate calls. *See* Notice of Inquiry, *supra* note 9, at 43,349; *see also* Carlin Communications, Inc. v. FCC, 113, 115 (2d Cir. 1984). The proper allocation of telephone regulatory authority between state and federal jurisdictions, and the Commerce Clause and federalism implications of such an allocation, are beyond the scope of this Note.

At least one state is contemplating a legislative response to the dial-a-porn problem. Two bills introduced in the Maryland House of Delegates in January 1985, Md. H. Delegates Bills No. 531 (Jan. 21, 1985), No. 607 (Jan. 24, 1985), would amend the Maryland obscenity statute, Md. Ann. Code art. 27, § 419 (1982 & Supp. 1984), to prohibit willful or knowing telephone transmission of obscene messages to minors.

[24] Rep. Bliley had filed comments in support of the dial-a-porn complaint before the FCC and had participated in the proceedings that followed. Notice of Inquiry, *supra* note 9, at 43,350.

[25] H.R. 2755, 98th Cong., 1st Sess., 129 Cong. Rec. H10,209 (daily ed. Nov. 17, 1983) (statement of Rep. Bliley).

violated in the FCC proceeding,[26] and would have added a new subsection, providing:

> Whoever . . . , by means of telephone, makes (directly or by recording device) any comment, request, suggestion, or proposal which is obscene, lewd, lascivious, filthy, or indecent, regardless of whether the maker of such comment placed the call, . . . shall be fined not more than $50,000 or imprisoned not more than six months, or both.[27]

The same penalties were provided for anyone who "knowingly permit[ted]" his telephone to be used to violate the new subsection.[28] Violations "for commercial purposes" would have exposed a defendant to additional civil fines of up to $50,000 per day of violation.[29]

On the second day of debate in the House, the text of this provision was replaced by language drafted by the House Judiciary Committee in an attempt to comply more closely with Supreme Court decisions restricting the regulation of sexually explicit communications.[30] The new language significantly limited the scope of the proposed statute by prohibiting (1) only "obscene or indecent" speech, (2) only transmissions to persons under eighteen years of age, and (3) only speech made "for commercial purposes."[31] The new provision also required the FCC to promulgate regulations specifying procedures by which a potential defendant could screen out callers under eighteen years of age[32] and allowed compliance with such regulations to be a "defense to a prosecution" under the new statute.[33] With

---

[26] *In re* Peter F. Cohalan and the County of Suffolk, New York v. New York Telephone Company, FCC File No. E-83-14 (Mar. 31, 1983), *cited in* H.R. REP. No. 356, 98th Cong., 1st Sess. 4 (1983). Later versions of the statute, however, did not repeal the provision allegedly violated by New York Telephone. Sen. Paul S. Trible (R-Va.), a sponsor of the legislation, explained "it is not the intent of Congress that a common carrier be prosecuted under this amendment when it is otherwise abiding by the law . . . ." 129 CONG. REC. S16,867 (daily ed. Nov. 18, 1983) (statement of Sen. Trible). This provision is now sections 223(a)(1)(A) and 223(a)(2) of the amended statute. *See* Appendix *infra*.

[27] H.R. 2755, 98th Cong., 1st Sess., H.R. REP. No. 356, *supra* note 26, at 3.

[28] *Id.*

[29] *Id.*

[30] 129 CONG. REC. H10,559 (daily ed. Nov. 18, 1983) (statement of Rep. Bliley). *See infra* notes 48–52 and accompanying text.

[31] 129 CONG. REC. H9356 (daily ed. Nov. 8, 1983) (text of amendment to H.R. 2755 proposed by Rep. Rodino (D-N.J.)).

[32] *Id.*

[33] *Id.*

minor additional modifications,[34] this new version of the bill amending section 223 passed both houses of Congress on November 18, 1983.[35] The President signed the bill on December 8, 1983.[36]

The FCC adopted its screening regulations on June 4, 1984.[37] The regulations, which sought to screen children out of the dial-a-porn audience by confining dial-a-porn to certain "adult hours," made "[o]perating [a dial-a-porn service] only between the hours of 9:00 p.m. and 8:00 a.m. Eastern Time" a "defense to prosecution" under the new dial-a-porn law.[38]

The Court of Appeals for the Second Circuit, however, held the screening regulations unconstitutional and set them aside.[39] The court did not decide the constitutionality of the underlying dial-a-porn statute.[40] Rather, it ruled that under the heightened scrutiny appropriate for any content-based regulation of speech, "the FCC has failed adequately to demonstrate that the regulatory scheme is well tailored to its ends or that those ends could not be met by less drastic means."[41] In requiring the FCC to reconsider alternatives discarded during its accelerated consideration of possible screening regulations,[42] the court insisted on "a record that shows, convincingly, that the regulations were chosen after thorough, careful, and comprehensive investigation and analysis."[43]

---

[34] The Judiciary Committee's proposal would have delayed the effective date of the new statute until the FCC promulgated screening regulations. *Id.* The statute as enacted deleted this postponement. Federal Communications Commission Authorization Act of 1983, Pub. L. No. 98-214, § 8(c), 1984 U.S. CODE CONG. & AD. NEWS (97 Stat.) 1467, 1469–70 (to be codified at 47 U.S.C. § 223).

[35] 129 CONG. REC. H10,562, S16,867 (daily ed. Nov. 18, 1983).

[36] 19 WEEKLY COMP. PRES. DOC. 1674 (Dec. 8, 1983); 129 CONG. REC. H10,664–65 (daily ed. Dec. 14, 1983).

[37] Report and Order, *supra* note 8, at 24,996.

[38] Report and Order, *supra* note 8, at 25,003 app. E (to be codified at 47 C.F.R. § 64.201(a)), *set aside*, Carlin Communications, Inc. v. FCC, 749 F.2d 113 (2d Cir. 1984). The regulations also exempted all "for pay" telephone sex services from prosecution under the new dial-a-porn law so long as they "requir[e] payment by credit card before transmission of the message(s)." *Id.* (to be codified at 47 C.F.R. § 64.201(b)). Any dial-a-porn service imposing such a payment requirement would also be exempt from prosecution. Such a service, however, would not be dial-a-porn as defined in this Note. *See supra* notes 9–15 and accompanying text.

[39] Carlin Communications, Inc. v. FCC, 749 F.2d 113 (2d Cir. 1984).

[40] *Id.* at 113.

[41] *Id.* at 121.

[42] *See* Report and Order, *supra* note 8, at 24,998–25,000. The infeasibility of these and other screening methods is addressed more thoroughly at *infra* text accompanying notes 194–209. The Second Circuit indicated that the FCC should reconsider all of them, including time channeling, screening and blocking, and access and identification codes. Carlin Communications, Inc. v. FCC, 749 F.2d 113, 121–23 (2d Cir. 1984).

[43] Carlin Communications, Inc. v. FCC, 749 F.2d 113, 123 (2d Cir. 1984).

In March 1985, the FCC issued a new notice of proposed rulemaking in its effort to promulgate acceptable screening regulations.[44]

B. *Coverage and Impact of New Statute*

As amended, section 223 prohibits anyone from "knowingly . . . by means of telephone, mak[ing] (directly or by recording device) any obscene or indecent communication for commercial purposes to any person under eighteen years of age . . . , regardless of whether the maker of such communication placed the call."[45] The new law, which also applies to those who "knowingly . . . permit[]"[46] their telephones to be used for such purposes, is enforceable by an array of criminal, civil, administrative, and injunctive sanctions.[47]

The legislative history of the amended section notes that, by confining the statute's prohibition to "obscene or indecent" speech, its congressional sponsors intended to comply with the Supreme Court's elaboration and application of these concepts in *Miller v. California*[48] and *FCC v. Pacifica Foundation.*[49] In *Miller*, the Court articulated the legal test for obscenity as:

> (a) whether "the average person, applying contemporary community standards" would find that the work, taken as a whole, appeals to the prurient interest, . . .
> (b) whether the work depicts or describes, in a patently offensive way, sexual conduct specifically defined by the applicable state law[,] and
> (c) whether the work, taken as a whole, lacks serious literary, artistic, political, or scientific value.[50]

As described in *Pacifica*, indecent speech must meet only the second branch of the *Miller* test.[51] The indecent speech at issue

---

[44] Enforcement of Prohibitions Against the Use of Common Carriers for the Transmission of Obscene Materials (Second Notice of Proposed Rulemaking), 50 Fed. Reg. 10,510 (1985).

[45] 47 U.S.C.A. § 223(b)(1)(A) (West Supp. 1984).

[46] *Id.* § 223(b)(1)(B).

[47] *Id.* § 223(b)(1), (b)(3)–(5); *see also* 47 U.S.C. § 312(b) (1982) (cease and desist orders authorized).

[48] 413 U.S. 15 (1973), *cited in* 129 CONG. REC. E5966 (daily ed. Dec. 14, 1983) (statement of Rep. Kastenmeier (D-Wis.)).

[49] 438 U.S. 726 (1978), *cited in* 129 CONG. REC. S16,866 (daily ed. Nov. 18, 1983) (statement of Sen. Trible).

[50] *Miller*, 413 U.S. at 24 (quoting Roth v. United States, 354 U.S. 476, 489 (1957)).

[51] *Pacifica*, 438 U.S. at 741.

in *Pacifica* was described as "patently offensive references to excretory and sexual organs and activities."[52] The content of dial-a-porn messages may fall within these definitions of obscenity and indecency and thus within the new statute's prohibitions.

A critical feature of the new law as applied to dial-a-porn is its requirement that a violator "knowingly . . . make[] . . . [a] communication . . . to any person under eighteen years of age."[53] The word *knowingly* in this context may be interpreted either as requiring that a dial-a-porn sponsor know a specific caller is under eighteen, or that the sponsor know that some callers are under eighteen, but not that any specific caller is under eighteen. The statute's legislative history makes clear that the second interpretation of knowingly is the intended interpretation.

The statute uses both "knowingly"[54] and "intentionally"[55] to describe levels of culpability. Because a defendant committing an intentional violation of the statute subjects himself to fines of up to $50,000 per day in addition to the sanctions to which he is exposed for a knowing violation,[56] the intentional level of culpability must exceed the knowing level. A sensible reading of the statute would thus designate as an intentional violation knowledge that a specific caller is underage and would designate as a knowing violation knowledge that only some callers are underage. Moreover, the drafters of this legislation knew how dial-a-porn systems operated,[57] and understood that a sponsor could not be certain of the age of any specific caller. By repeatedly stating their intention that a typical dial-a-porn system would fall within the statute,[58] the drafters made it clear that

[52] *Id.* at 743 (footnote omitted).

[53] 47 U.S.C.A. § 223(b)(1)(A) (West Supp. 1984).

[54] *Id.* § 223(b)(1).

[55] *Id.* § 223(b)(3).

[56] *Id.*

[57] "[O]bscene messages, whether made directly or by recording device, are prohibited without regard to whether the sender of the message initiated the call. The Committee intends that this section will prohibit obscene messages otherwise available over 'Dial-It' services." H.R. REP. NO. 356, *supra* note 26, at 19, *reprinted in* 1984 U.S. CODE CONG. & AD. NEWS 2219, 2235; 129 CONG. REC. H10,209 (daily ed. Nov. 17, 1983) (statement of Rep. Bliley); *id.* at H10,559–61 (daily ed. Nov. 18, 1983) (statement of Rep. Bliley); *id.* at S16,866–67 (statement of Sen. Trible).

[58] *See* H.R. REP. NO. 356, *supra* note 26, at 19, *reprinted in* 1984 U.S. CODE CONG. & AD. NEWS 2219, 2235; 129 CONG. REC. H10,209 (daily ed. Nov. 17, 1983) (statement of Rep. Bliley); *id.* at H10,559–61 (daily ed. Nov. 18, 1983) (statement of Rep. Bliley); *id.* at S16,866–67 (statement of Sen. Trible).

the statutory language should be interpreted to require only that a defendant know that some callers are under eighteen.

Following this interpretation, the dial-a-porn sponsor who learns that children are using his service, and who is unable to screen out such callers from adult callers, has two options for completely avoiding the statute's sanctions: he may sanitize his communications so that they are no longer legally obscene or indecent or he may discontinue his service altogether. A third option might be compliance with FCC screening regulations, which are required to be promulgated by the same statute that amended section 223 of the Communications Act.[59]

III. Constitutionality of the Federal Response

The Supreme Court's application of First Amendment principles to diverse modes of communication has been increasingly flexible and innovative. During the 1950's, the Court was reluctant to modify the constitutional protections accorded communication technologies: "Each method [of communication] tends to present its own peculiar problems. But the basic principles of freedom of speech and the press, like the First Amendment's command, do not vary."[60] As community standards changed and certain technologies became more pervasive, the Court began to tailor its free speech analysis to the specific technology at issue. In *Red Lion Broadcasting Co. v. FCC*,[61] the Court declared that "[a]lthough broadcasting is clearly a medium affected by a First Amendment interest, . . . differences in the characteristics of new media justify differences in the First Amendment standards applied to them."[62] Although the *Red Lion* approach calls for a more penetrating analysis of the varying free speech interests at stake in each case, it does not mitigate the difficulty of developing constitutionally valid regulations for sexually oriented telephone entertainment services.

[59] Compliance with FCC screening regulations will provide a dial-a-porn sponsor with a "defense to a prosecution" regardless of the efficacy of such screening measures. 47 U.S.C.A. § 223 (b)(2) (West Supp. 1984). Such a defense, however, might not apply to the various sanctions in section 223 which do not amount to "prosecution[s]," such as assessments of civil fines or suits by the Attorney General to enjoin violations of section 223(b)(1). *Id.* § 223(b)(4)–(5). If the FCC is unable to develop constitutionally valid screening regulations or if the defense to prosecution provided by section 223(b)(2) is interpreted as not covering both criminal and noncriminal sanctions authorized by the statute, enforcement will force dial-a-porn sponsors to sanitize or to discontinue their communications.

[60] Joseph Burstyn, Inc. v. Wilson, 343 U.S. 495, 503 (1952).

[61] 395 U.S. 367 (1969).

[62] *Id.* at 386 (citing Joseph Burstyn, Inc. v. Wilson, 343 U.S. 495, 503 (1952)).

Cases in this area offer only limited guidance. In *Schad v. Borough of Mount Ephraim*,[63] a case involving allegedly non-obscene nude dancing, Justice White's opinion for the Court noted that "[e]ntertainment, as well as political and ideological speech, . . . fall[s] within the First Amendment guarantee."[64] In addition, various state courts have held that the use of automatic telephone answering equipment to transmit certain controversial messages is constitutionally protected speech.[65] In situations where the telephone is being used for illegal purposes, such as gambling[66] or prostitution,[67] however, courts have upheld terminations of telephone service, even though legal uses involving constitutionally protected speech are also terminated.[68] Dial-a-porn raises the analogous issue of whether the state is empowered to prohibit transmission of sexually explicit communications to children, regardless of the impact that such a prohibition may have on adult free speech rights.

The constitutionality of the federal[69] attempt to preclude child access to dial-a-porn can be evaluated only after the free speech and privacy rights of adults in the area of sexually explicit

---

[63] 452 U.S. 61 (1981).

[64] *Id.* at 65 (citations omitted).

[65] Figari v. N.Y. Tel. Co., 32 A.D.2d 434, 448, 303 N.Y.S.2d 245, 260 (1969) (controversial political message communicated via telephone equipment is protected); Huntley v. Pub. Util. Comm'n, 69 Cal. 2d 67, 78, 442 P.2d 685, 692, 69 Cal. Rptr. 605, 612 (1968) (telephone company tariff, approved by Public Services Commission, requiring that recorded messages include names and addresses of responsible parties violated freedom of speech guarantees); *see also* Anderson v. N.Y. Tel. Co., 42 A.D.2d 151, 160, 345 N.Y.S.2d 740, 749 (1973), *rev'd on other grounds*, 35 N.Y.2d 746, 320 N.E.2d 647, 361 N.Y.S.2d 913 (1974) (constitutional objections to the imposition of any law requiring a telephone company summarily to terminate a subscriber's service because allegedly defamatory messages were being transmitted).

[66] *See* Palma v. Powers, 295 F. Supp. 924 (N.D. Ill. 1969).

[67] *See* Goldin v. Pub. Util. Comm'n, 23 Cal. 3d 638, 592 P.2d 289, 153 Cal. Rptr. 802 (1979).

[68] *See* Palma v. Powers, 295 F. Supp. 924, 941 (N.D. Ill. 1969); Goldin v. Pub. Util. Comm'n, 23 Cal. 3d 638, 657 n.7, 592 P.2d 289, 301 n.7, 153 Cal. Rptr. 802, 814 n.7 (1979). *See generally Recent Developments, Summary Termination of Telephone Service for Suspected Illegal Use*, 20 STAN. L. REV. 136 (1967).

[69] No attempt will be made here to distinguish between state and federal authority in the areas of family law or child protection. *See Developments in the Law—The Constitution and the Family*, 93 HARV. L. REV. 1156, 1159 (1980) (family law traditionally a state rather than federal concern) [hereinafter cited as *Developments—the Family*]. *See generally* W. WADLINGTON, C. WHITEBREAD & S. DAVIS, CASES AND MATERIALS ON CHILDREN IN THE LEGAL SYSTEM 47–48 (1983). Likewise, the limits of congressional power under the Commerce Clause to regulate sexually explicit materials in interstate commerce will not be addressed here. For discussion of the Commerce Clause as a source of national police power, enabling Congress to exclude obscene material from the flow of interstate commerce, see Cushman, *The National Police Power under the Commerce Clause of the Constitution*, 3 MINN. L. REV. 289, 381–412 (1919) (classic statement of Congress's power to exclude from interstate commerce activities deemed to be deleterious to the public interest). *See generally* W. LOCKHART, Y. KAMISAR & J. CHOPER, CONSTITUTIONAL LAW 107–08 (5th ed. 1980).

communications are analyzed and then compared with the corresponding rights of children. Such an analysis of the free speech and privacy rights of children must address and account for the interest of both parent and child in parental supervision, the parent's interest in state support of parental supervision, and the state's "independent interest in the well-being of its youth."[70] After such an analysis and comparison clarifies the differing reaches of adult and child rights, conflicts between the rights of children and their parents and the rights of adults in general must be isolated and resolved.[71]

A. *The Free Speech and Privacy Rights of Adults*

1. Adult Rights to Free Speech

Adults have a First Amendment right to send and to receive any communication constituting "speech," no matter how offensive or sexually explicit. They are presumed to have a "full capacity for individual choice which is the presupposition of First Amendment guarantees."[72]

Modern free speech theorists espouse an array of conflicting and overlapping justifications for the free speech rights of

---

[70] Ginsberg v. New York, 390 U.S. 629, 640 (1968).

[71] A commercial speech analysis of dial-a-porn might be warranted in light of (1) the widespread use of dial-a-porn messages to advertise sexually explicit magazines and "for pay" telephone sex services, (2) the large revenues for dial-a-porn sponsors generated by their tariff arrangements with local telephone companies, and (3) the statute's prohibition of only communications made "for commercial purposes." 47 U.S.C.A. § 223(b)(1)(A) (Supp. 1984). Such an analysis is not undertaken in this Note because the legislative history makes clear that the dial-a-porn statute is intended to address the problem of child access to obscene or indecent telephone communications and is not addressed primarily to the usual commercial speech concerns of false, unprofessional, or otherwise objectionable advertising. *See* 129 CONG. REC. H10,559–61, S16,866–67 (daily ed. Nov. 18, 1983) (statements of Rep. Bliley and Sen. Trible) ("commercial purposes" explained but not emphasized).

More important, the first part of the four-part commercial speech analysis announced in Central Hudson Gas & Elec. Corp. v. Pub. Serv. Comm'n, 447 U.S. 557, 566 (1980), provides: "At the outset, we must determine whether the expression is protected by the First Amendment. For commercial speech to come within that provision, it at least must concern lawful activity and not be misleading." The focus of this Note is on the preliminary step of assessing whether the obscene or indecent communications to children prohibited by the dial-a-porn statute are a lawful activity protected by the First Amendment.

[72] Ginsberg v. New York, 390 U.S 629, 649–50 (Stewart, J., concurring) (footnote omitted).

adults.[73] These justifications are best integrated into a coherent system by Thomas Emerson.[74] Emerson divides these justifications into four "main premises":

> First, freedom of expression is essential as a means of assuring individual self-fulfillment . . . .
> Second, freedom of expression is an essential process for advancing knowledge and discovering truth . . . .
> Third, freedom of expression is essential to provide for participation in decision making by all members of society . . . .
> Finally, freedom of expression is a method of achieving a more adaptable and hence a more stable community, of maintaining the precarious balance between healthy cleavage and necessary consensus.[75]

All four of Emerson's main premises for free speech apply to adults. Together they provide a useful framework for analyzing the free speech rights of adults to send and to receive[76] sexually explicit communications.

Adult access to obscene and indecent materials is most strongly supported by Emerson's first premise, an adult's free speech interest in individual self-fulfillment. Emerson's subsequent premises of advancing knowledge and discovering truth and of participating in decisionmaking are conceivably applicable to some adults.[77] Though potentially genuine in some individual cases,[78] such claims should not obscure the fact that most recipients of these materials choose to receive them for purposes of pleasure and entertainment, purposes supported almost exclusively by the individual self-fulfillment justification for free speech.

---

[73] *See* T. EMERSON, THE SYSTEM OF FREEDOM OF EXPRESSION (1970); Z. CHAFEE, FREE SPEECH IN THE UNITED STATES (1941); A. MEIKLEJOHN, FREE SPEECH AND ITS RELATION TO SELF-GOVERNMENT (1948); Baker, *Scope of the First Amendment Freedom of Speech*, 25 UCLA L. REV. 964 (1978); Bork, *Neutral Principles and Some First Amendment Problems*, 47 IND. L.J. 1 (1971); Redish, *The Value of Free Speech*, 130 U. PA. L. REV. 591 (1982); Scanlon, *A Theory of Freedom of Expression*, 1 PHIL. & PUB. AFF. 204 (1972). *See generally* W. LOCKHART, Y. KAMISAR & J. CHOPER, *supra* note 69, at 684–702; M. NIMMER, NIMMER ON FREEDOM OF SPEECH, A TREATISE ON THE THEORY OF THE FIRST AMENDMENT (1984); L. TRIBE, AMERICAN CONSTITUTIONAL LAW, § 12-1, at 576–79 (1978).

[74] T. EMERSON, *supra* note 73.

[75] *Id.* at 6.

[76] An analysis of the free speech rights of adult *listeners* necessarily accounts for the free speech rights of adult *speakers*, because attempts to regulate the content of communications between them would affect speakers and listeners equally.

[77] T. EMERSON, *supra* note 73, at 6–7.

[78] *See* United States v. 31 Photographs, 156 F. Supp. 350 (S.D.N.Y. 1957) (articles obscene as to the general public, but used solely for bona fide scientific research, held not obscene).

An adult is entitled to exercise his free speech rights based on individual self-fulfillment because he is presumed to have a full capacity for individual choice. Thus, adults enjoy a First Amendment right to send and to receive any sexually explicit material, so long as such material constitutes speech and not conduct.[79]

---

[79] First Amendment theorists disagree on the validity of a speech/conduct distinction. *Compare* T. EMERSON, *supra* note 73, at 18 (distinguishing speech from conduct by judging "whether expression or action is the dominant element") *with* Baker, *supra* note 73, at 1009–29 (arguing that the free speech guarantee protects against "general prohibitions of substantively valued conduct"). *See generally* L. TRIBE, *supra* note 73, § 12-1, at 579 (1978) (criticizing "artificial dichotomy between [protected] speech-related conduct in which 'expression' predominates and [unprotected] conduct in which 'action' is dominant"); W. LOCKHART, Y. KAMISAR & J. CHOPER, *supra* note 69, at 1100–28; Schauer, *Response: Pornography and the First Amendment*, 40 U. PITT. L. REV. 605 (1979). According to Schauer, materials that are designed to be "nothing more than . . . linguistic or pictorial sex aid[s]," without any redeeming social value, should be classified as nonspeech. *Id.* at 609. The state should then be able to regulate them under the police power to the same extent that it can regulate other forms of conduct, such as gambling or prostitution.

Schauer's position parallels the Supreme Court's view that only hard core pornography should be deemed obscene and hence unprotected by the First Amendment. *See* Miller v. California, 413 U.S. 15, 27 (1973); *see also* Jacobellis v. Ohio, 378 U.S. 184, 197 (1964) (Stewart, J., concurring) ("criminal laws in this area are constitutionally limited to hard-core pornography"). The traditional justification for regulation of pornography under the police power has been the protection of public morals. *See, e.g.*, Henkin, *Morals and the Constitution: The Sin of Obscenity*, 63 COLUM. L. REV. 391, 395 (1963) ("[C]ommunities believe, and act on the belief, that obscenity is immoral, is wrong for the individual, and has no place in a decent society . . . . Obscenity is not suppressed primarily for the protection of others. Much of it is suppressed for the purity of the community and for the salvation and welfare of the 'consumer'."). The Supreme Court also has sought to justify regulation of hard core pornography on the basis of public health and safety. *See* Paris Adult Theatre I v. Slaton, 413 U.S. 49, 58 (1973). Regulation under the police power to protect public health, safety, morals, and welfare may be invoked to regulate only conduct and not speech. *See Developments—the Family, supra* note 69, at 1199.

A dial-a-porn service could be characterized as conduct and regulated as such for both adults and children. Congress, however, approached the regulation of dial-a-porn as regulation of speech, *see* 129 CONG. REC. H10,559–61, S16,866–67 (daily ed. Nov. 18, 1983) (statements of Rep. Bliley and Sen. Trible), and challenges to the regulations have been made on that basis. *See* Carlin Communications, Inc. v. FCC, 749 F.2d 113 (2d Cir. 1984). Moreover, dial-a-porn services may be sufficiently erotic to be of concern to the caretakers of children without meeting the Supreme Court's definition of hard core pornography.

If obscenity is defined as conduct and not speech, legislatures should be free to regulate it without amassing irrefutable empirical data. The propriety of regulating obscenity should be judged in relation to the propriety of other forms of morality-based conduct regulations—e.g., prostitution, gambling, homosexuality, drinking alcoholic beverages, or using nonaddictive drugs—and not in relation to state attempts to regulate speech. *See* Schauer, *supra*, at 611, 613.

As applied to adult rights to send and to receive dial-a-porn messages, this position leads to several results. For obscene messages, where the telephone in effect becomes an electronic "sex aid," the participants are engaged in conduct that can be regulated under the police power. For merely indecent messages, which are "speech" for adults, adult free speech rights absolutely protect the communications. Thus, had the new dial-a-porn statute not been confined to those under eighteen years of age, its prohibition of "obscene or indecent communications" would have violated adult free speech rights.

2. Adult Rights to Privacy

Even if dial-a-porn is considered conduct, and thus is unprotected by adult free speech rights, the privacy rights of adults may protect their access to the service. The Supreme Court reconciled adult privacy rights with the state's power to regulate obscenity in *Stanley v. Georgia*.[80] In *Stanley*, the Court upheld an individual's right to possess and to use a private collection of obscene films in his home.[81] Justice Marshall, writing for the Court, reasoned that the individual appellant

> is asserting the right to read or observe what he pleases—the right to satisfy his intellectual and emotional needs in the privacy of his own home. . . . If the First Amendment means anything, it means that a State has no business telling a man, sitting alone in his own house, what books he may read or what films he may watch.[82]

In light of Marshall's description elsewhere in the opinion of a "right to receive information and ideas, regardless of their social worth"[83] and his rejection of state attempts to "control the moral content of a person's thoughts,"[84] this reasoning seems to accord constitutional protection to almost any acquisition or use of obscene materials by adults.[85] While the Supreme Court has not adopted this view nor been receptive to efforts to invalidate controls on the sale and distribution of obscene materials to consenting adults,[86] reconciling *Stanley* with the justifications for regulating adult access to obscenity[87] requires careful scrutiny of the privacy interests at stake in each particular case.

The privacy interests at stake in dial-a-porn communications may be evaluated in light of the three separate strands of the

[80] 394 U.S. 557 (1969).

[81] *Id.* at 559.

[82] *Id.* at 565.

[83] *Id.* at 564.

[84] *Id.* at 565 (footnote omitted).

[85] *See* Paris Adult Theatre I v. Slaton, 413 U.S. 49, 107–08 (Brennan, J., dissenting); *see also* Katz, *Privacy and Pornography:* Stanley v. Georgia, 1969 SUP. CT. REV. 203.

[86] *See Paris Adult Theatre I*, 413 U.S. at 57 (1973) (state may regulate showing of obscene films, even if minors are excluded); United States v. 12 200-Ft Reels of Film, 413 U.S. 123, 128 (1973) ("*Stanley* does not permit one to go abroad and bring such material into the country for private purposes."); United States v. Orito, 413 U.S. 139, 141 (1973) (*Stanley* does not "create[] a correlative right to receive[,] . . . transport[,] . . . or distribute" obscene materials); United States v. Reidel, 402 U.S. 351 (1971) (state regulation of use of mails to distribute obscene materials not forbidden by *Stanley*). *See generally* G. GUNTHER, CASES AND MATERIALS ON CONSTITUTIONAL LAW 1353–55 (10th ed. 1980).

[87] *See* Paris Adult Theatre I, 413 U.S. 49, 57–60 (1973).

Stanley reasoning: (1) the right to private possession and enjoyment of obscenity, (2) the right to receive information and ideas, regardless of social worth, and (3) the right to be free of state control of the content of one's thoughts.

The first strand of *Stanley* protects an adult's right to listen to obscene tape recordings played in the home. It must be extended, however, to apply to telephone use. The Florida Supreme Court undertook such an extension in *State v. Keaton*[88] and applied *Stanley* to the content of telephone conversations between consenting adults. The *Keaton* court struck down on overbreadth grounds a Florida statute prohibiting obscene phone conversations.[89] The statute made criminal the "mak[ing over the telephone of] any comment, request, suggestion, or proposal which is obscene, lewd, lascivious, filthy, or indecent."[90] The court held that such a statute would literally "forbid[] a sexually oriented conversation between lovers."[91] In this situation, the privacy rights of both parties coalesce to protect the conversation, even though the exercise by each party of such rights is not strictly confined to his or her own home.

For calls to the dial-a-porn service, the lack of appreciable privacy interests on the sponsor's end of the calls prevents them from fitting within the *Keaton* analysis. The receiving end of a dial-a-porn telephone call is a tape-playing machine available to thousands of callers per hour.[92] This service appears to be the functional equivalent of a movie theater "open to the public for a fee." The Court has found, however, that movie theaters and private homes do not warrant similar treatment.[93]

Though the first strand of *Stanley*, as elaborated by *Keaton*, may not protect adult access to dial-a-porn, such access may be protected by the second strand of *Stanley*, the right "to receive information and ideas, regardless of their social worth."[94] Determining whether communications contain "information and ideas," however, might violate *Stanley*'s protection

---

[88] 371 So. 2d 86 (Fla. 1979).
[89] *Id.* at 93.
[90] *Id.* at 87 n.1 (quoting FLA. STAT. § 365.16(1)(a) (1977)). The unamended portion of 47 U.S.C. § 223 contains identical language. *See* 47 U.S.C.A. § 223(a)(1)(A) (West Supp. 1984).
[91] *Keaton*, 371 So. 2d at 90.
[92] The same analysis holds true for the "for pay" mode of telephone sex services, *supra* note 8, where the receiving end of the call is a person available as the equivalent of a paid sexual assistant to anyone who calls.
[93] Paris Adult Theatre I, 413 U.S. 49, 65 (1973) (citing *Stanley*, 394 U.S. at 568).
[94] *Stanley*, 394 U.S. at 564.

of the right to private possession and enjoyment of obscenity.[95] In the case of dial-a-porn, no such intrusion is necessary because the content of dial-a-porn messages is open to scrutiny by any caller and may be readily characterized as protected speech or as some form of unprotected telephonic sex aid.[96]

Only the third strand of *Stanley*, the right to be free of governmental control of the contents of one's thoughts,[97] cannot be completely reconciled with governmental regulation of adult access to obscenity. This strand of *Stanley*, taken to an extreme, could prevent all government regulation of speech or actions because such regulation would amount to indirect regulation of a person's thoughts. Justice Marshall's opinion in *Stanley* makes clear that the case was not meant to disrupt a state's "broad power to regulate obscenity"[98] outside the privacy of an individual's home. Subsequent cases have left this power intact.[99]

Thus, the first two strands of privacy rights articulated in *Stanley* with respect to adult access to obscenity do not impede regulation when the obscene nature of the communications in question can be established and their transmission interrupted without inquiring into a person's thoughts or intruding into his home. Although the third strand, the right to be free of governmental thought control, is difficult to reconcile with any degree of obscenity regulation for adults, courts have not permitted it to encroach on the state's "broad power to regulate obscenity."[100]

3. Summary

Adults have a First Amendment right to send and to receive sexually explicit communications, no matter how offensive or indecent, so long as the communication is not conduct, and thus is not the equivalent of a linguistic or pictorial sex aid.[101] Al-

---

[95] *See id.* at 564–65.
[96] *See* Schauer, *supra* note 79, at 609.
[97] *Stanley*, 394 U.S. at 564–66.
[98] *Id.* at 568.
[99] *See, e.g.*, Paris Adult Theatre I, 413 U.S 49, 67 (1973) (Court upheld regulation of movie theatres arguing that "[p]reventing unlimited display or distribution of obscene material . . . is distinct from a control of reason and the intellect"); United States v. Reidel, 402 U.S. 351, 354 (1971) (quoting *Stanley*, 394 U.S. at 568, to uphold the constitutionality of a federal statute prohibiting the mailing of obscene materials, as applied to those who routinely disseminate such material).
[100] *Stanley*, 394 U.S. at 568.
[101] *See* Schauer, *supra* note 79, at 609.

though adults may be entitled to use or to possess hard core pornography in the privacy of their own homes, any extension of such a privacy right to protect the receipt or transmission of sexually explicit communications through a dial-a-porn service appears unjustified.

B. *The Free Speech Rights of Minors and Their Parents*

Children are presumed to be incapable of mature, competent, and responsible choice.[102] Justice Stewart, concurring in *Ginsberg v. New York*,[103] asserted that "a State may permissibly determine that, at least in some precisely delineated areas, a child—like someone in a captive audience—is not possessed of that full capacity for individual choice which is the presupposition of First Amendment guarantees."[104] The impact of a child's limited capacity for choice on the scope of his free speech guarantee is best analyzed using Emerson's four "main premises" justifying free speech in the adult world.[105]

Emerson's four justifications for adult free speech rights initially seem inapplicable to children. Few children, for example, "advanc[e] knowledge" or "participat[e] in decision making."[106] John H. Garvey's "instrumental" theory of children's free speech rights, however, overcomes these limitations.[107] Garvey views free speech rights as instrumental to preparing children for their future roles in adult society.[108] Thus, Emerson's four justifications should be evaluated in terms of their contribution to preparing a child for his eventual exercise of adult free speech rights.

From Garvey's instrumental perspective, a child has a right to free speech to the extent that it "is instrumental in the growth of his ability to participate in self-government"[109] or "plays an instrumental role in advancing the search for knowledge and

---

[102] The potential harshness of this presumption may, in certain circumstances, be mitigated by individual determinations of maturity. *See generally* Tribe, *Childhood, Suspect Classifications, and Conclusive Presumptions: Three Linked Riddles*, 39 LAW & CONTEMP. PROBS. 8, 31–36 (1975).

[103] 390 U.S. 629 (1968).

[104] *Id.* at 649–50 (Stewart, J., concurring) (footnote omitted).

[105] *See supra* text accompanying notes 74–75.

[106] T. EMERSON, *supra* note 73, at 6–7.

[107] Garvey, *Children and the First Amendment*, 57 TEX. L. REV. 321, 350–51 (1979).

[108] *Id.*

[109] *Id.* at 338 (discussing Tinker v. Des Moines School Dist., 393 U.S. 503 (1969)).

truth."[110] In these cases, which parallel Emerson's second, third, and fourth premises,[111] the vital training provided by exercising these rights justifies free speech, despite the child's lack of "moral and rational faculties" equal to those of adults[112] and his likely inability to make a significant contribution to existing knowledge.[113]

In contrast, Garvey restricts children's free speech rights justified solely by individual autonomy or self-realization, goals which correspond to Emerson's "individual self-fulfillment"[114] premise for free speech. For adults, the individual self-fulfillment rationale may be invoked to justify "almost unbounded freedom of expression."[115] But for children, an "undeveloped sense of judgment"[116] and a less than full capacity for choice[117] justify a corresponding diminution in the scope of their free speech rights.[118]

1. Making Choices for Minors: Obscenity

First Amendment protection for sexually explicit entertainment is almost exclusively based on the individual self-fulfillment rationale for free speech.[119] Accordingly, an analysis of the rights of children in this area must first acknowledge a child's less than full capacity for choice[120] and then examine the means and rationales for delegating decisionmaking authority on behalf of a child. Specifically, such authority often must be delegated to the child's parents or to the state.[121]

a. *Parental Choices.* Society looks first to parents for decisionmaking on behalf of their children. A parent's authority over

---

[110] *Id.* at 344 (discussing Meyer v. Nebraska, 262 U.S. 390 (1923)).
[111] T. EMERSON, *supra* note 73, at 6–7.
[112] Garvey, *supra* note 107, at 340.
[113] *Id.* at 344.
[114] T. EMERSON, *supra* note 73, at 6.
[115] Garvey, *supra* note 107, at 345.
[116] *Id.* at 351.
[117] *See id.* at 346–47.
[118] *See id.* at 350.
[119] *See supra* notes 77–79 and accompanying text.
[120] *See* Garvey, *supra* note 107, at 349–50 (in the area of sexuality, "there is something to be said for withholding certain kinds of information from children until the ability to make sound judgments develops with experience." (footnote omitted)).
[121] *See* Schall v. Martin, 104 S. Ct. 2403, 2410 (1984) ("Children, by definition are not assumed to have the capacity to take care of themselves. They are assumed to be subject to the control of their parents, and if parental control falters, the State must play its part as *parens patriae*.").

his or her child is "based upon established cultural preferences for parent-directed family life"[122] and the presumed parental interest in a child's welfare and in childraising.[123] The limits of parental authority are reached when a child is emancipated or becomes mature and are exceeded when the parent has abused or neglected the child or is otherwise unfit to raise the child.[124]

Children themselves also may have an interest in permitting parents to decide which communications best promote their growth and self-fulfillment. Bruce Hafen notes that "preservation of the parental authority that is a prerequisite to meaningful family autonomy is in fact in the interest of children and their most vital long-range rights."[125] Similarly, Garvey argues that a "child has a right to the imposition [by his parents] of certain external preferences,"[126] because "subjection to parental decisions [is] . . . necessary to assure his future ability to choose for himself."[127]

b. *State Support of Parental Supervision.* Where parents lack the power to make their authority over the listening and reading activities of their children effective, they may seek to invoke the power and resources of the state to reinforce their supervision. State support of parental supervision of children is justified by the presumed quality of parental decisions on behalf of children, by a general reluctance to countenance direct state supervision of children,[128] and by the fact that "parental direction is unlikely to be a disguise for governmental direction,"[129] which ordinarily ought not to be allowed to "control . . . what the child reads outside school hours."[130]

The state generally intervenes to support parental authority only with respect to children below a certain age, regardless of

---

[122] Hafen, *Children's Liberation and the New Egalitarianism: Some Reservations About Abandoning Children to Their "Rights"*, 1976 B.Y.U. L. Rev. 605, 619; *see also Developments—the Family, supra* note 69, at 1352.

[123] *Developments—the Family, supra* note 69, at 1353. Professor Hafen suggests that a parental interest in childraising, distinct from other interests in a child's welfare, may be a basis for parental authority. *See* Hafen, *supra* note 122, at 626–29.

[124] *See* Garvey, *supra* note 107, at 332. For example, a parent might abuse a child by indiscriminately exposing the child to hard core pornography, thus exceeding his authority and justifying state intervention. *See* Hafen, *supra* note 122, at 629–30; *Developments—the Family, supra* note 69, at 1218–19.

[125] Hafen, *supra* note 122, at 655.

[126] Garvey, *supra* note 107, at 330.

[127] *Id.*

[128] *See Developments—the Family, supra* note 69, at 1214–16.

[129] Garvey, *supra* note 107, at 332.

[130] *Id.* at 331–32.

any individual child's capacity for choice. The state thus avoids endangering family harmony by encouraging a child to demonstrate his capacity for choice or his independence from his parents in order to be exempted from state regulation.[131]

Supreme Court decisions on children's access to obscene or indecent materials support the use of state power to extend and to reinforce parental authority over children. In *Ginsberg v. New York*,[132] the Court upheld the constitutionality of a New York statute that prohibited the sale to minors of materials defined to be obscene as to them. The Court maintained that "[t]he legislature could properly conclude that parents and others, teachers for example, who have th[e] primary responsibility for children's well-being are entitled to the support of laws designed to aid discharge of that responsibility."[133] Similarly, in *FCC v. Pacifica Foundation*,[134] the Court upheld the FCC's authority to regulate broadcasting of indecent speech in order to protect children and unconsenting adults. Justice Stevens, writing for the majority, asserted that a child's easy access to radio broadcasts and the government's interest in supporting parents' "authority in their own household"[135] supported the FCC's regulatory authority.[136]

The Supreme Court has limited state support of parental supervision by confining the state's regulatory authority to only the most objectionable materials. The state may regulate materials that are "obscene as to minors," such as those at issue in *Ginsberg*.[137] "Obscenity as to minors" is defined by adapting the three parts of the adult obscenity definition—prurient effect, patent offensiveness, and no value—to a hypothetical audience of minors.[138] Because minors are presumed to be more suscep-

---

[131] L. TRIBE, *supra* note 73, § 16-32, at 1097 n.29 (1978). The risk of disturbing family harmony may be worth taking when family harmony is already impaired or when a child invokes privacy rights in the procreation rights area. *See, e.g.*, Bellotti v. Baird, 443 U.S. 622 (1979) (pregnant minor seeking abortion entitled to opportunity for individual determination of maturity); *see also Developments—the Family*, *supra* note 69, at 1375.

[132] 390 U.S. 629 (1968) (affirming conviction of lunch counter operator for selling "girlie" magazines to sixteen year old boy).

[133] *Id.* at 639.

[134] 438 U.S. 726 (1978).

[135] *Id.* at 749 (quoting Ginsberg v. New York, 390 U.S. 629, 639 (1968)).

[136] *Id.* at 758.

[137] *Ginsberg*, 390 U.S. 629, 633 (1968).

[138] *See id.* at 635, 646 (quoting N.Y. PENAL LAW § 484-h (McKinney 1968)). The cited New York statutory provision has been amended and recodified as N.Y. PENAL LAW § 235.20(6) (McKinney 1980). *See generally* Lockhart & McClure, *Censorship of Obscenity: The Developing Constitutional Standards*, 45 MINN. L. REV. 5, 68–88 (1960); Schauer, *The Return of Variable Obscenity?*, 28 HASTINGS L.J. 1275 (1977).

tible than adults to the harms resulting from exposure to sexually explicit materials, "obscenity as to minors" includes materials that are not obscene as to adults.[139]

Also subject to state regulation are "indecent" materials such as those at issue in *Pacifica*. The concept of "indecency" applied in *Pacifica* measures only the offensiveness of materials without requiring a prurient effect.[140] Yet, for Justice Stevens, indecent words "offend for the same reasons that obscenity offends."[141] Similarly, Justice Powell's concurrence asserted that the indecent speech that could be regulated in *Pacifica* was "as potentially degrading and harmful to children as representations of many erotic acts."[142] Although indecency may be analogous to obscenity, it still includes a broader class of materials. By permitting this broadening whenever children are involved, the Court has maintained a disparity between materials that can be regulated as to children and materials that can be regulated as to adults.

The Court also has limited state regulatory authority by requiring accommodation of the rights of parents who wish to forego state assistance and allow their children unfettered access to materials regulated by the state. The Court addressed the rights of these "permissive" parents in *Ginsberg*[143] and *Pacifica*[144] by hypothetical ad hoc assessments of the hardship involved in forcing such parents to receive communications by alternative means and then to relay them to their children. While the burden of requiring parental purchase of a magazine destined for a child, as in *Ginsberg*, is defensible as a surrogate for requiring the child to prove parental consent to a merchant, the burden in *Pacifica* is more formidable. Dissenting in *Pacifica*,

---

[139] *Ginsberg*, 390 U.S. at 634–35.

[140] *Pacifica*, 438 U.S. at 738–41.

[141] *Id.* at 746 (referring to Citizen's Complaint Against Pacifica Foundation Station WBAI (FM), New York, N.Y., 56 F.C.C.2d 94, 98 (1975), *modified*, 59 F.C.C.2d 892 (1976), *rev'd*, 556 F.2d 9 (1977), *rev'd*, 438 U.S. 726 (1978)).

[142] *Pacifica*, 438 U.S. at 758 (Powell, J., concurring in part). Justice Powell's statement in *Pacifica* parallels his conclusion in Erznoznik v. City of Jacksonville, 422 U.S. 205 (1975), that materials obscene as to youth "must be, in some significant way, erotic." *Erznoznik*, 422 U.S. at 213 & n.10 (quoting Cohen v. California, 403 U.S. 15, 20 (1971)).

[143] Justice Brennan noted that the law "does not bar parents who so desire from purchasing the magazines for their children." *Ginsberg*, 390 U.S. at 639; *see also* Estreicher, *Schoolbooks, School Boards, and the Constitution*, 80 COLUM. L. REV. 1092, 1103–04 (1980).

[144] Of the prevailing opinions in *Pacifica*, only Justice Powell's concurrence noted the possibility that some parents might choose to allow their children access to indecent speech. *Pacifica*, 438 U.S. at 758 (Powell, J., concurring in part); *see also Pacifica*, 438 U.S. at 770 (Brennan, J., dissenting).

Justice Brennan noted that, for parents who decide to allow their children access to the speech at issue, a George Carlin monologue, the alternatives of buying tapes or records or listening to late evening "adult" broadcasts or going to theaters or nightclubs[145] "involve the expenditure of money, time, and effort that many of those wishing to hear Mr. Carlin's message may not be able to afford, and [require a change of medium which overlooks] . . . the reality that in many cases the medium may well be the message."[146] These burdens on a permissive parent's choice of free speech activities for his child reflect an accommodation between the rights of permissive and unpermissive parents. Such accommodations, particularly in the broadcasting context, are inevitable when the state tries to support one group of parents without wholly extinguishing the rights of other parents.

c. *State Protection of Children.* The state may directly intervene in the child's or his family's activities to contradict explicit choices made on the child's behalf by his parent or guardian. In *Prince v. Massachusetts*,[147] the Court affirmed a guardian's conviction for violating state child labor laws, despite the guardian's asserted right to decide whether having a child sell religious literature on the streets was an appropriate way to further the child's religious training.[148] Efforts to justify state intervention when the risk is to a child's intellectual or developmental wellbeing, rather than to his physical wellbeing, have been unsuccessful.[149]

When a parental choice on a child's behalf has been made, problems created by a child's incapacity for choice are solved. Direct state intervention probably should be confined to clear risks to a child's welfare resulting from indisputably defective parental choices. Thus, direct state interference with a parent's

---

[145] Justice Stevens' opinion in *Pacifica* suggested these alternatives. *See Pacifica*, 438 U.S. at 750 n.28.

[146] *Pacifica*, 438 U.S. at 774 (Brennan, J., dissenting). "The airways are capable not only of carrying a message, but also of transforming it. A satirist's monologue may be most potent when delivered to a live audience; yet the choice whether this will in fact be the manner in which the message is delivered and received is one the First Amendment prevents the government from making." *Id.* at 775.

[147] 321 U.S. 158 (1944).

[148] *Id.* at 166–71.

[149] *See* Wisconsin v. Yoder, 406 U.S. 205 (1972) (removal of children from school on religious grounds prior to attaining a state-mandated minimum level of education permitted on First Amendment and parental authority grounds).

choice to allow a child access to sexually explicit materials, when no physical or sexual abuse of the child is occurring,[150] probably is not justified.[151] Yet, the state's inability to interfere directly with parental choices for a child in the obscenity area does not imply that a state regulatory scheme may never impose indirect burdens on parents.[152]

State intervention to protect children also may occur when no one has made a choice on the child's behalf. When no parental choice has been made, as when a parent fails to choose[153] or is himself incapable of choice, the problem of the child's incapacity for choice remains. In this situation, the incapacity may be remedied by the imposition of choices from an alternative decisionmaker, such as the state. Independent state authority to protect children is supported by rationales analogous to those supporting parental authority. In *New York v. Ferber*,[154] Justice White argued that "a State's interest in 'safeguarding the physical and psychological well-being of a minor' is 'compelling.'"[155] The state also has an interest in assuring the "growth [of children] into free and independent well-developed men and citizens."[156] Unlike parental authority, however, the state's authority is not supported by such rationales as the bonds of love and kinship in a parent-child relationship, parental knowledge of a particular child's needs, and society's interest in pluralism.[157] Accordingly, the state should be entitled to less authority over a child's free speech activities than the child's parent.

Supreme Court decisions on child access to sexually explicit materials in the absence of parental choice elaborate the state's independent interest in protecting children. In *Ginsberg*, the Court upheld a statutory ban on materials obscene as to minors, based on the legislative finding that such materials were harmful to minors.[158] Even though empirical evidence on the effects of pornography on children was indeterminate, the Court found

---

[150] *See supra* note 124 and accompanying text.
[151] *See supra* notes 122–27 and accompanying text.
[152] *See infra* text accompanying notes 236–40.
[153] Parental failure to choose would include ordinary child neglect and probably should include complete parental indecision. It would not include a conscious parental delegation of authority to a child to select his own communicative materials.
[154] 458 U.S. 747 (1982).
[155] *Id.* at 756–57 (quoting Globe Newspaper Co. v. Superior Court, 457 U.S. 596, 607 (1982)).
[156] Prince v. Massachusetts, 321 U.S. 158, 165 (1944).
[157] *See supra* notes 122–27 and accompanying text.
[158] *Ginsberg*, 390 U.S. at 641–43.

that the legislature "might rationally conclude"[159] that the exposure of children to certain materials would constitute an abuse that might impair their development into "free and independent . . . citizens."[160]

Limits on independent state authority to protect children were articulated in *Erznoznik v. City of Jacksonville*.[161] In *Erznoznik*, the Court invalidated a city ordinance banning theaters from exhibiting various forms of nudity on motion picture screens visible from a public place.[162] Justice Powell, writing for the Court, stated "[s]peech that is neither obscene as to youths nor subject to some other legitimate proscription cannot be suppressed solely to protect the young from ideas or images that a legislative body thinks unsuitable for them."[163] Although he noted that "[t]he First Amendment rights of minors are not 'co-extensive with those of adults,'"[164] Powell emphasized that "[i]n most circumstances, the values protected by the First Amendment are no less applicable when government seeks to control the flow of information to minors."[165]

The Court's recent decision in *Board of Education v. Pico*[166] took Powell's analysis one step further. Five students sued their school board for removing books that the board characterized as "anti-American, anti-Christian, anti-Sem[i]tic, and just plain filthy'" from school library shelves.[167] The board acted pursuant to what it perceived as its "moral obligation[] to protect the children in our schools from this moral danger."[168] In affirming the reversal of a summary judgment granted to the school board by the district court, the plurality articulated a minor's right to receive ideas through access to books in a school library.[169] According to Justice Brennan, author of the plurality opinion, such a right is necessary to protect the rights of communicators to have their ideas received and "is a necessary predicate to the

---

[159] *Id.* at 641.

[160] *Id.* at 640–41 (quoting Prince v. Massachusetts, 321 U.S. 158, 165 (1944)).

[161] 422 U.S. 205 (1975).

[162] *Id.*

[163] *Id.* at 213–14.

[164] *Id.* at 214 n.11 (quoting Tinker v. Des Moines School Dist., 393 U.S. 503, 515 (1969) (Stewart, J., concurring)).

[165] *Id.* at 214 (footnote omitted).

[166] 457 U.S. 853 (1982).

[167] *Id.* at 857 (quoting press release *reprinted in* Pico v. Board of Education, Island Trees Union Free School District, 474 F. Supp. 387, 390 (E.D.N.Y. 1979)).

[168] *Id.* (quoting press release *reprinted* in Pico v. Board of Education, Island Trees Union Free School District, 474 F. Supp. 387, 390 (E.D.N.Y. 1979)).

[169] *Id.* at 868–69.

recipient's meaningful exercise of his own rights of speech, press, and political freedom."[170] While acknowledging that books could be removed for being "pervasively vulgar,"[171] the plurality invoked the right to receive ideas to support its conclusion that school authorities had unconstitutionally removed the books if they "*intended* by their removal decision to deny [the students] access to ideas with which [the school authorities] disagreed, and if this intent [was] the decisive factor in [the school authorities'] decision."[172]

Therefore, the First Amendment offers some protection for the free speech rights of minors against independent efforts by the state to protect children. That First Amendment protection, however, is less than the protection accorded the free speech rights of adults.

2. A Minor's Right to Privacy

The First Amendment protects minors against efforts by the state to curtail their access to materials that are not "obscene as to minors"[173] or "indecent"[174] as defined in *Ginsberg* and *Pacifica.*[175] For materials falling within either of these standards, however, nothing has yet been said about a minor's ability to invoke his right to privacy to ward off state, and possibly even parental, interference.

Recent Supreme Court cases have protected a minor's right to privacy in the area of procreation rights against state attempts to limit the access of minors to abortions[176] and contraceptives.[177] The state's authority to protect children from the harm perceived to flow from exposure to obscenity or pornography seems irreconcilable with the state's lack of authority in the procreation rights area to attempt, however imperfectly, to control sexual activity itself.

---

[170] *Id.* at 867 (emphasis omitted).
[171] *Id.* at 871.
[172] *Id.* (footnote omitted) (emphasis in original).
[173] *See* Ginsberg v. New York, 390 U.S. 629, 634–35 (1968).
[174] FCC v. Pacifica Foundation, 438 U.S. 726, 738–41 (1978).
[175] *See supra* text accompanying notes 138–42.
[176] *See* Bellotti v. Baird, 443 U.S. 622 (1979); Planned Parenthood v. Danforth, 428 U.S. 52 (1976) (state cannot require parental consent for a minor to receive an abortion). *But see* H.L. v. Matheson, 450 U.S. 398 (1981) (state can require physician to give notice to parents before performing abortion on a minor).
[177] *See* Carey v. Population Services Int'l, 431 U.S. 678 (1977).

The analysis, however, should begin with the biological facts governing the procreation rights area, which force minors to make choices whether they have a full capacity for choice or not. Unlike the obscenity area, minors themselves must make their own procreational choices in the first instance, and state or parental sheltering is conceptually and practically much more infeasible. Because these facts have been held to justify a privacy right for minors in the procreation rights area,[178] extending such a right to include a right to receive information necessary to make informed choices about abortion and contraception seems defensible in light of the highly individual nature of these procreational choices.[179]

Obscene and indecent materials, though potentially transmitting relevant factual information, primarily focus on sexual gratification or entertainment. These choices are not as urgent as the procreational choices that minors may confront. Acquisition of obscene or indecent materials thus should not fall within a minor's right to privacy even if this right is enlarged to entitle a minor to receive information bearing on procreational choices.[180]

3. Summary

Minors are presumed to have a diminished capacity for choice and thus cannot be relied upon to exercise responsible judgment in choosing what communications to send or to receive. The authority to choose on behalf of a minor rests initially with his parents, whose decisions are essentially unreviewable for First Amendment purposes. The state may intervene either to support parental choices or, when parental choices are clearly harmful or not discernible, to advance its own interests in a child's well-being. Courts have confined state intervention by (1) establishing a regulated class of sexually explicit materials

---

[178] *See* Bellotti v. Baird, 443 U.S. 622 (1979); Carey v. Population Services Int'l, 431 U.S. 678 (1977); Planned Parenthood v. Danforth, 428 U.S. 52 (1976).

[179] *See* Bolger v. Youngs Drug Products Corp., 103 S. Ct. 2875, 2884 n.30 (1983) (minors' "pressing need" for information about contraception requires "significant measure" of First Amendment protection for such information).

[180] Although Stanley v. Georgia, 394 U.S. 557 (1969), extended the adult right to privacy into the obscenity area, the free speech aspect of this right is based solely on the individual self-fulfillment justification for free speech. Reliance on self-fulfillment makes the right available to children only in the instrumental ways their parents choose for them. *See supra* notes 102–18 and accompanying text. A right contingent upon parental choice will not be analyzed as a minor's right in this Note.

which is larger for children than for adults and (2) requiring the state to permit alternative access to materials for children whose parents choose to forego state assistance in child supervision. Access to materials that the state may otherwise regulate should not be guaranteed on the basis of a minor's right to privacy.

D. *Consequences of Disparate First Amendment Rights of Adults and Children*

1. Overbreadth

When the statutory effort to regulate telephone sex services is not closely tailored to the contours of free speech rights, problems of unconstitutional overbreadth arise.[181] Facial overbreadth in the new statute amending section 223, where the statutory language attempts to regulate both protected and unprotected speech, generally reflects only errant draftsmanship and appears to be easily remedied.

Examples of the new statute's facial overbreadth include its prohibition of obscene or indecent communications to all persons under eighteen years of age, even when they have parental consent,[182] and the potential prohibition of obscene or indecent speech during certain person-to-person telephone conversations "for commercial purposes," such as a merchant's quarrel with a customer under eighteen.[183] If it were clear that minors have a full capacity for choice before their eighteenth birthdays, the statutory age of eighteen might also constitute overbreadth,

---

[181] *See generally* NAACP v. Alabama, 377 U.S. 288, 307 (1964) ("[A] governmental purpose to control or prevent activities constitutionally subject to regulation may not be achieved by means which sweep unnecessarily broadly and thereby invade the area of protected freedoms."); G. GUNTHER, *supra* note 86, at 1185–95; Note, *The First Amendment Overbreadth Doctrine*, 83 HARV. L. REV. 844 (1970) (Many laws "are so broadly drafted that the range of possible applications violating the first amendment . . . [is] substantial.").

[182] *See* 47 U.S.C.A. § 223(b)(1)(A) (West Supp. 1984).

[183] *Id.* The phrase "for commercial purposes" may be unconstitutionally vague because of the ambiguous reach that the drafters of the new legislation intended for it. *See* 129 CONG. REC. H10,559 (daily ed. Nov. 18, 1983) (statement of Rep. Bliley); *id.* at S16,867 (daily ed. Nov. 18, 1983) (statement of Sen. Trible). A statute is unconstitutionally vague if it lacks "sufficiently definite warning as to the proscribed conduct when measured by common understanding and practices." Jordan v. DeGeorge, 341 U.S. 223, 231–32 (1951). *See generally* G. GUNTHER, *supra* note 86, at 1188 n.9; Amsterdam, *The Void-for-Vagueness Doctrine in the Supreme Court*, 109 U. PA. L. REV. 67 (1960).

because older adolescents would be improperly denied access to "indecent"[184] communications that are not "obscene."

In addition to being facially overbroad, the statute is potentially overbroad as applied because the actual operation of its regulatory scheme may unavoidably infringe protected speech. The statute will force any dial-a-porn sponsor who is incapable of screening out calls from children, but who knows some children are using his service, either to cease communicating or to sanitize his communications by deleting anything that might be "obscene or indecent."[185] The dial-a-porn sponsor thus will be forced to censor "indecent" speech that is protected as to adults.[186] The effectiveness of screening, then, becomes the central issue in deciding whether the statute will have the effect of "reduc[ing] the adult population . . . to reading [or receiving] only what is fit for children"[187] in violation of the longstanding rule of *Butler v. Michigan.*[188]

2. Analysis of Screening Methods

The screening methods promulgated by the FCC pursuant to the new statute were struck down by the United States Court of Appeals for the Second Circuit as unconstitutional.[189] At present, these and other screening methods appear to suffer from constitutional or practical problems.

a. *Time Channeling.* The Second Circuit gave several reasons for striking down the FCC's attempt to screen children from dial-a-porn by confining the service to adult hours.[190] Finding the FCC's regulations "both overinclusive and underinclusive,"[191] the court noted that adults would be prevented from reaching dial-a-porn during the day, while children could "easily

---

[184] The balance of this discussion will use the term "indecent," and not "obscene as to minors" or "pervasively vulgar," to describe the nonobscene materials that the state may regulate as to children but not as to adults.

[185] See *supra* note 59 and accompanying text.

[186] See *supra* note 79 and accompanying text.

[187] Butler v. Michigan, 352 U.S. 380, 383 (1957).

[188] 352 U.S. 380, 383 (1957); *see infra* notes 213–27.

[189] *See supra* notes 37–43 and accompanying text.

[190] Carlin Communications, Inc. v. FCC, 749 F.2d 113 (2d Cir. 1984); *see supra* text accompanying notes 39–43; *see also* Report and Order, *supra* note 8, at 25,001–02 (difficulty of establishing simultaneous "adult hours" in all eight time zones of the United States and its possessions).

[191] Carlin Communications, Inc. v. FCC, 749 F.2d 113, 121 (2d Cir. 1984).

pick up a private or public telephone" and reach dial-a-porn during the evening.[192] Dial-a-porn sponsors could enhance child patronage in the adult hours by using the daytime hours to play advertisements for the evening service.[193] The court also found the FCC's day/night distinction unsupported by the record in light of the fact that children are often better supervised during the day hours at school than they are during the evening hours at home.[194]

b. *Screening and Blocking*. The FCC decided not to promulgate regulations using screening and blocking techniques.[195] The FCC concluded that

> a blocking or screening scheme, whether implemented from the central office, subscriber premises or coin-operated telephones, would require time to develop and could entail costs which would outweigh the benefits to be obtained. From an economic, technical, as well as practical standpoint, blocking or screening schemes do not, at this time, represent viable regulatory options.[196]

The Second Circuit criticized the FCC's consideration of blocking and screening methods. Specifically, the court criticized the FCC's failure to consider enabling telephone customers to block all calls to the 976 exchange from their premises[197] and questioned the FCC's conclusion that a blocking capability in customer premises equipment would not be possible.[198] The court noted that the costs of these methods could be placed upon the dial-a-porn sponsor,[199] though the court later acknowledged that "any regulation that drives [a dial-a-porn sponsor] out of business would seem to fall under *Butler v. Michigan*."[200]

---

[192] *Id.*
[193] *Id.*
[194] *Id.*
[195] Blocking and screening are used by the FCC to refer to using the physical apparatus of the telephone network, including customer premises equipment, to sort wanted from unwanted calls. Screening by customer-initiated blocking would be analogous to the postal service's screening option contained in 39 U.S.C. § 3008 (1982). Report and Order, *supra* note 8, at 24,998–99.
[196] *Id.*
[197] Carlin Communications, Inc. v. FCC, 749 F.2d 113, 122 (2d Cir. 1984). The 976 exchange is the exchange assigned for all dial-it services.
[198] *Id.* The court noted that such blocks appear to be feasible for large telephone systems. *Id.*; *see* N.Y. Times, Dec. 15, 1983, at A25, col. 6 (block installed at Defense Intelligence Agency).
[199] Carlin Communications, Inc. v. FCC, 749 F.2d 113, 122 n.16 (2d Cir. 1984).
[200] *Id.* at 123 n.19.

The Second Circuit's analysis of screening and blocking methods did not address the overbreadth issues raised by its proposal to enable telephone customers to block all calls to the 976 exchange.[201] In addition, blocking methods raise the practical issue of the monitoring and updating burdens that would be placed on customers who are able to install number-specific blocking equipment on their home telephones.[202] Finally, in its discussion of some of the options, the court's analysis overlooks the statutory language, which reflects a congressional judgment as to who should be responsible for implementing screening measures. The regulations described in section 223(b)(2) of the statute expressly are required to be procedures by which "the defendant" will restrict child access to dial-a-porn.[203] Dial-a-porn sponsors, rather than parents, are responsible for screening child access to dial-a-porn under the statute.

c. *Access and Identification Codes.* Child callers could be screened by requiring intervention by either live operators or a computer, which would then obtain an access or identification code from each caller. The FCC rejected this scheme because it "would place substantial economic and administrative burdens on recorded service providers."[204] The Second Circuit thought that a scheme whereby dial-a-porn callers could apply to be included on a dial-a-porn sponsor's approved caller list, and thereby allow their ages to be verified, might be feasible, and it therefore recommended that the FCC give it further scrutiny.[205] The court recognized that inconveniences attending this scheme might diminish dial-a-porn patronage and that adults lacking identification or access codes would be disadvantaged.[206] It failed to recognize the substantial privacy interests implicated when adults seeking to receive dial-a-porn messages are forced to apply in writing, have their names included on a dial-a-porn master list, and submit to age verification by dial-a-porn sponsors.[207]

---

[201] *Id.* at 122 n.14.

[202] For this reason, options such as requiring parents to put locks on their telephones, wholly aside from the safety issues raised whenever a child might need to use the telephone in an emergency, are also not contemplated by section 223(b)(2).

[203] 47 U.S.C.A. § 223(b)(2) (West Supp. 1984).

[204] Report and Order, *supra* note 8, at 25,000.

[205] Carlin Communications, Inc. v. FCC, 749 F.2d 113, 123 n.16 (2d Cir. 1984).

[206] *Id.*

[207] *See supra* notes 80–99 and accompanying text.

d. *Miscellaneous*. Miscellaneous additional screening methods considered by the FCC, but not by the Second Circuit, include restrictions on the advertisement of dial-a-porn numbers, disclaimers preceding dial-a-porn messages, and an approach by which screening recommendations would be solicited from dial-a-porn sponsors.[208] The FCC rejected all of these schemes because it either lacked jurisdiction to implement them or found them ineffective.[209]

While screening adult callers from child callers may someday be possible, and while some current proposals show promise, such screening seems currently infeasible.[210] If screening is infeasible, the new statute will force dial-a-porn sponsors to censor speech that is protected as to adults, a result prohibited by *Butler v. Michigan*,[211] or face criminal, civil, administrative, or injunctive sanctions.[212] Unless modified or flexibly applied, therefore, *Butler* would seem to render the new statute unconstitutional.

IV. THE *Butler v. Michigan* DOCTRINE REEXAMINED

A. *Origins and Evolution*

In *Butler v. Michigan*,[213] the Supreme Court reversed the conviction of a merchant who sold a book to a police officer in violation of a Michigan statute that criminalized the sale or distribution of "obscene, immoral, lewd, or lascivious [materials] . . . manifestly tending to the corruption of the morals of youth."[214] Justice Frankfurter's opinion for the Court noted that

> appellant was convicted because Michigan . . . made it an offense for him to make available for the general reading public . . . a book that the trial judge found to have a potentially deleterious influence upon youth. The State insists that, by thus quarantining the general reading public against

---

[208] *See* Report and Order, *supra* note 8, at 25,000.
[209] *Id.*
[210] *See generally* Further Notice of Inquiry and Notice of Proposed Rulemaking, *supra* note 12, at 2125 n.4 ("[A]lthough we intend to make every effort to fashion rules that serve their intended purpose while being both practicable and constitutional, we note that adoption of specific regulations may not be feasible.").
[211] 352 U.S. 380 (1957).
[212] 47 U.S.C.A. § 223(b)(1), (3)–(5) (West Supp. 1984).
[213] 352 U.S. 380 (1957).
[214] *Id.* at 381 (quoting MICH. COMP. LAWS § 750.343 (Supp. 1954)).

> books not too rugged for grown men and women in order to shield juvenile innocence, it is exercising its power to promote the general welfare. Surely, this is to burn the house to roast the pig.[215]

Frankfurter concluded that "[t]he incidence of this enactment is to reduce the adult population of Michigan to reading only what is fit for children."[216] The statute therefore violated the First Amendment because it was "not reasonably restricted to the evil"[217] of child exposure to obscenity.

The Supreme Court's decision in *Butler* was its first decision "squarely fac[ing] the . . . problem of the constitutionality of official censorship of obscenity."[218] The decision was the Court's first step toward resolving a longstanding debate in obscenity law over what audience should be used to gauge the erotic impact of reading materials available to the general public.[219] In rejecting Michigan's attempt to use children as the appropriate audience, the Court impliedly rejected all tests focusing on the impact of obscenity on "the young and vulnerable."[220] *Butler* thus set the stage for the landmark case of *Roth v. United States*,[221] in which the Court explicitly rejected tests based on "particularly susceptible persons"[222] and adopted a test based on "the average person, applying contemporary community standards."[223]

Although applied infrequently during the Court's ensuing struggle with "the intractable obscenity problem"[224] the principle of *Butler*—that the state may not regulate obscenity in a way that "reduce[s] the adult population . . . to reading only what is fit for children"[225]—has proven durable. It has been invoked to invalidate a state-sponsored system of informal censorship of

---

[215] *Id.* at 382–83.
[216] *Id.* at 383.
[217] *Id.*
[218] Lockhart & McClure, *supra* note 138, at 5 (footnote omitted).
[219] Case Comment, *Constitutional Law—Obscenity Statutes—Freedom of Speech*, 11 MIAMI L.Q. 523, 523–24 (1957); *see* Roth v. United States, 354 U.S. 476, 488–89 nn.24–26 (1957); *see also* Lockhart & McClure, *supra* note 138, at 70–88.
[220] Kalven, *The Metaphysics of the Law of Obscenity*, 1960 SUP. CT. REV. 1, 6, 7. The "young and vulnerable" test had an antecedent in the "particularly susceptible person" test of Regina v. Hicklin, 3 L.R.-Q.B. 360 (1868). *See* Kalven, *supra*, at 3.
[221] 354 U.S. 476 (1957).
[222] *Id.* at 489 (citing Regina v. Hicklin, 3 L.R.-Q.B. 360 (1868)) ("The *Hicklin* test . . . judg[ed] obscenity by the effect of isolated passages upon the most susceptible persons . . . .").
[223] *Id.* at 489.
[224] Interstate Circuit, Inc. v. Dallas, 390 U.S. 676, 704 (1968) (Harlan, J., dissenting).
[225] *Butler*, 352 U.S. at 383.

materials unfit for youths[226] and a statutory ban on mailing unsolicited advertisements for contraceptives.[227] The durability of the *Butler* rule came into question, however, in the unique circumstances of the *Pacifica* case.[228]

The issue in *Pacifica* was the FCC's authority to regulate indecent speech[229] in radio broadcasts during a time of day when the audience might contain substantial numbers of children.[230] The impossibility of screening adult from child listeners in the broadcasting medium[231] left the Court with a choice between upholding the FCC's regulatory authority over indecent speech, thereby infringing some adult free speech rights, or denying the FCC regulatory authority, but fully protecting adult free speech rights. As pointed out by two of the three separate opinions in the court of appeals[232] and by Justice Brennan's dissent in the Supreme Court,[233] the *Butler* rule seemingly required the Court to deny the FCC regulatory authority because, in the context of radio broadcasting, exercising such authority would indeed reduce adults to receiving only radio communications fit for children during the daytime hours in question.

---

[226] Bantam Books, Inc. v. Sullivan, 372 U.S. 58, 71 (1963) (practice of notifying distributors that certain materials had been found objectionable for sale, distribution, or display to youths under 18 years of age and advising them that Commission had duty to recommend prosecution found unconstitutional).

[227] Bolger v. Youngs Drug Products Corp., 103 S. Ct. 2875, 2884 (1983) (statute prohibiting the mailing of unsolicited advertisements for contraceptives invalidated for violating First Amendment right to commercial speech).

[228] FCC v. Pacifica Foundation, 438 U.S. 726 (1978).

[229] Indecent speech such as that at issue in *Pacifica* is protected by adult free speech rights but unprotected by the free speech rights of minors without parental consent. *See supra* notes 79, 137–42, 184 and accompanying text. Because a child with parental permission is actually exercising his parent's adult-level free speech rights, such a child should be treated as an adult for these purposes. *See supra* notes 72–79, 132–42 and accompanying text. Although such children will be treated as adults in the instant discussion, the special problems created with respect to them by modifying the *Butler* doctrine are addressed *infra* at text accompanying notes 277–80.

[230] The Supreme Court did not address the problem of the number of affected children that would be required to justify regulation under *Pacifica*. The question of how many children must be exposed to sexually explicit communications for a different constitutional analysis to apply is outside the scope of this Note. *See generally* Pacifica Foundation v. FCC, 556 F.2d 9, 36 (Leventhal, C.J., dissenting), *rev'd*, 438 U.S. 726 (1978) (considering the possibility that restrictions on broadcast content might be relaxed during periods when "the great preponderance of children are subject to parental control"); Tribe, *supra* note 102, at 31–36 (possibility of childhood as a suspect classification).

[231] *Pacifica*, 438 U.S. at 750.

[232] Pacifica Foundation v. FCC, 556 F.2d 9, 17 (Tamm, J.); *id.* at 27 (Bazelon, C.J., concurring), *rev'd*, 438 U.S. 726 (1978).

[233] *Pacifica*, 438 U.S. at 766, 768–69 (Brennan, J., dissenting).

Justice Stevens's opinion,[234] however, expanded the *Butler* analysis by focusing on whether regulation of a message in one mode of communication would unduly restrict adult access to the same message in all other modes of communication.[235] Stevens noted that adults could receive the same message, George Carlin's "Filthy Words" monologue, by "purchas[ing] tapes and records or go[ing] to theaters and nightclubs to hear these words," and that late evening broadcasts might be permissible.[236] Justice Stevens therefore ruled that adults would not be "reduce[d] . . . to hearing only what is fit for children,"[237] and that the *Butler* rule would not be violated. For Stevens, this array of alternatives for adults prevented FCC regulation of daytime radio broadcasting from unconstitutionally restricting adults to only child-worthy communications.[238]

Though a majority of the *Pacifica* Court supported Justice Stevens's conclusion that *Butler v. Michigan* did not bar the FCC from regulating indecent broadcasts, two members of this majority offered a different analysis of the *Butler* issue. Justice Powell's concurrence, joined by Justice Blackmun, reasoned that the *Butler* argument that the Court's ruling would reduce adults to hearing only child-worthy material "is not without force. The [Federal Communications] Commission certainly should consider it as it develops standards in this area. But *it is not sufficiently strong* to leave the Commission powerless to act in circumstances such as those in this case."[239] Powell then recited the alternative methods of adult access[240] and concluded that where a radio station inflicts a "verbal shock treatment" on its audience, an FCC order regulating its broadcasts would not violate the station's First Amendment rights.[241]

Thus, according to Justice Stevens's view in *Pacifica*, the *Butler* analysis should consider the full array of communication media through which adults may receive certain communica-

---

[234] Chief Justice Burger and Justices Rehnquist, Powell, and Blackmun joined in the portion of Justice Stevens's opinion described in the text accompanying notes 235–38. *Id.* at 727. This part of Stevens's opinion therefore was endorsed by a majority of the Court.

[235] *See id.* at 750 n.28.

[236] *Id.*

[237] *Id.* at 760 (citing Butler v. Michigan, 352 U.S. 380, 383 (1957)).

[238] *Id.*

[239] *Id.* at 760 (Powell, J., concurring in part) (emphasis added).

[240] *Id.*

[241] *Id.* at 761.

tions. In Justice Powell's view, however, the *Butler* doctrine would not be "sufficiently strong,"[242] in certain circumstances, to bar regulation.

Commentators have criticized the *Pacifica* case for authorizing content-based regulation of nonobscene speech[243] and have argued that at best it should be limited to its unique facts.[244] Professor Laurence Tribe has argued:

> the result the Court reached in *Pacifica* seems a poor basis for doctrinal extrapolation . . . . It would be regrettable . . . if the opinions of the majority were allowed to leave any enduring marks on First Amendment jurisprudence: *Pacifica* should be confined to its facts, and eventually discarded as a "derelict in the stream of the law."[245]

A crucial fact in the *Pacifica* case was the impossibility of screening out children from the radio audience during a broadcast intended for adults. Had a method existed for separating the adult and child members of the broadcasting audience so that only suitable material could be broadcast to each audience, a complaint might not have been filed with the FCC.[246] Confining *Pacifica* to its facts, at the very least, means confining it to situations where the audience cannot be segregated into adults and children. Rather than discarding the fact-specific *Pacifica* decision, this Note attempts to harmonize *Pacifica* with *Butler* and to generalize *Pacifica* into a workable First Amendment treatment of attempts to regulate communications to nonsegregable audiences of adults and children.

B. *Modifying the* Butler *Analysis*

1. Changing the Focus of the *Butler* Analysis

The *Butler* doctrine focuses on protecting the individual First Amendment rights of adults engaged in communication.[247] While

---

[242] *Id.*

[243] *See* G. GUNTHER, *supra* note 86, at 1258 n.4 (quoting Gunther, *The Highest Court, The Toughest Issues*, STANFORD MAGAZINE, Fall–Winter 1978, at 34); L. TRIBE, *supra* note 73, at 61–68 (Supp. 1979).

[244] L. TRIBE, *supra* note 73, at 67–68 (Supp. 1979); *The Supreme Court, 1977 Term*, 92 HARV. L. REV. 1, 161–63 (1978).

[245] L. TRIBE, *supra* note 73, at 67–68 (Supp. 1979) (quoting North Dakota State Bd. of Pharmacy v. Snyder's Drug Store, Inc., 414 U.S. 156, 167 (1973) (Douglas, J.)).

[246] *See Pacifica*, 438 U.S. at 749. The case started with a father's complaint to the FCC about an indecent radio broadcast that he heard on his radio while driving with his young son. *Id.* at 729–30.

[247] The "communication" in *Butler* was the purchase of pornographic reading materials. *See Butler*, 352 U.S. at 381.

this focus may be justified when a bilateral communication, such as a personal conversation or the sale of a book, is being analyzed, it is unrealistic in a situation where a single speaker simultaneously can reach thousands of listeners, including children. While such a speaker and his willing adult listeners are entitled to exercise their individual rights, the fullest enjoyment of these rights will eventually collide with the individual rights held by others in the speaker's audience to avoid offense or to shield children. In order to reconcile these conflicting individual rights,[248] the audience within which any such reconciliation would take place must first be isolated and described.

Justice Jackson remarked in *Kovacs v. Cooper*[249] that each mode of communication "is a law unto itself."[250] A logical first step in developing the "law" of a new communication technology is to identify any nonsegregable audiences created by the technology. Such audiences would be those portions of the entire potential audience that could not be further subdivided by either the communicator or his listeners—that is, the communicator could not reach one member of a nonsegregable audience without reaching all members of that audience. In any particular case, the exact size and composition of a nonsegregable audience will depend on the means available for channeling communications to only willing listeners, on the means available for protecting unwilling listeners, and on the ability and maturity of potential audience members to protect themselves.[251] As an analytical tool, the nonsegregable audience better reflects the practical impact of mass media than does *Butler*'s focus on the

---

[248] These conflicting individual rights are considered to be of widely disparate constitutional statures. *See, e.g.*, Erznoznik v. City of Jacksonville, 422 U.S. 205, 210 (1975) ("[I]n our pluralistic society . . . [m]uch that we encounter offends our esthetic, if not our political and moral sensibilities. Nevertheless, the Constitution does not permit government to decide which types of otherwise protected speech are sufficiently offensive to require protection for the unwilling listener or viewer."). *But see* Note, *Content Regulation and the Dimensions of Free Expression*, 96 HARV. L. REV. 1854, 1861–62 (1983) ("The basic problem in content regulation cases is not, as modern first amendment theory assumes, that of reconciling free expression with other interests. Instead, the problem is better understood as one of reconciling competing interests in expression.").

[249] 336 U.S. 77 (1949) (upholding constitutionality of reasonable regulation of sound amplification in streets and public places).

[250] *Id.* at 97 (Jackson, J., concurring).

[251] For example, in person-to-person conversations or ordinary telephone conversations, the available means for listeners to screen themselves and for speakers to select only willing listeners make the nonsegregable audience one person. However, in daytime radio broadcasting, as noted in *Pacifica*, 438 U.S. at 748–50, *see also id.* at 758–59 (Powell, J., concurring in part), the nonsegregable audience contains adults and children as well as willing and unwilling listeners.

individual rights of each member of an audience as if that member and the mass communicator were engaged in a private conversation.

2. Changing the Substance of the *Butler* Analysis

a. *Content-Based Regulation in General.* Accommodation between the free speech and privacy interests of some listeners and the nonspeech interests of other listeners in a nonsegregable audience can take place only if content-based regulation of the speech directed at such an audience can be justified.[252] The rationales for content-based regulation of indecent speech offered by the *Pacifica* Court,[253] though perhaps intuitively appealing, are somewhat incomplete.[254] Justice Stevens found content-based regulation of patently offensive words justified by analogizing them to obscenity, which is beyond First Amendment protection.[255] Justice Brennan's dissent also acknowledges that some content-based regulation of "offensive language on broadcasts directed specifically at younger children" might be permissible.[256]

A more complete justification for content-based regulation of communications directed at a nonsegregable audience must address the conflict between the rights of individuals and the rights of communities.[257] This conflict can be analyzed and resolved by utilizing certain aspects of the "community standards" doc-

---

[252] *See* FCC v. Pacifica Foundation, 438 U.S. 726, 744–48 (1978) (plurality opinion); *see also* Note, *Content Regulation and the Dimensions of Free Expression*, *supra* note 248, at 1859 ("Consideration of the content of expression is inescapable in first amendment adjudication. The real question . . . is not whether, but for what purposes and in what ways, judges should consider content in determining whether speech is protected.").

[253] *See infra* text accompanying notes 255–56 (rationales for content-based regulation in *Pacifica*).

[254] *See Pacifica*, 438 U.S. at 773 (Brennan, J., dissenting).

[255] "These words offend for the same reasons that obscenity offends." *Pacifica*, 438 U.S. at 746 (footnote omitted); *see* Roth v. United States, 354 U.S. 476 (1957); *see also supra* text accompanying notes 140–42.

[256] *Id.* at 768 n.3 (Brennan, J., dissenting) (quoting Ginsberg v. New York, 390 U.S. 629, 650 (1968) (Stewart, J., concurring)).

[257] *See* J. ELY, DEMOCRACY AND DISTRUST: A THEORY OF JUDICIAL REVIEW 105–06 (1980) (use of judicial review to prevent "any inhibition of expression that is unnecessary to the promotion of a government interest" by weighing the governmental interest against the individual's First Amendment rights).

trine of obscenity law enunciated in *Roth*[258] and *Miller*.[259] This doctrine, which mitigates extreme individualism in deference to community values or sensibilities,[260] requires that the prurient or erotic effect of allegedly obscene materials be judged by "the average person, applying contemporary community standards."[261] Though often mired in debate over what the relevant community is,[262] the doctrine in effect provides that a community can require restraints on the free speech rights of some of its members for the good of the community as a whole.

The community standards rule technically infringes the free speech right of the most "rugged"[263] community members to receive materials that are not prurient, and are thus "speech," as to them. If such infringements were forbidden, the community would have to accord free speech protection to materials that *are* prurient to a majority of the community. This degree of protection would burden the community with unrestricted quantities of obscene materials merely to accommodate the free speech rights of those few individuals as to whom the materials are not obscene. This result would be unacceptable for the same reason that the "particularly susceptible person" test,[264] which some courts had employed to justify restrictions on a vast array of materials prior to *Roth*,[265] was unacceptable. It would grant excessive deference to the unique "ruggedness" or "prudishness" of certain individual community members in disregard of the community-wide interest, recognized in *Roth*, in permitting

---

[258] Roth v. United States, 354 U.S. 476 (1957) (obscene materials denied First Amendment protection because their content is offensive to "contemporary community standards").

[259] Miller v. California, 413 U.S. 15, 24 (1973) (Court rejected the "utterly without redeeming social value" test and adopted the less restrictive "lacks serious artistic, political, or social value" test). *See supra* text accompanying note 50.

[260] See Note, *Content Regulation and the Dimensions of Free Expression*, *supra* note 248, at 1865–66.

[261] Roth v. United States, 354 U.S. 476, 489 (1957).

[262] For a discussion of proper "community" standards, see United States v. Langford, 688 F.2d 1088, 1091–95 (7th Cir. 1982) (jury properly instructed to apply "contemporary community standards" of state where defendant was convicted and from which obscene materials depicting minors were mailed). *See also* Waples & White, *Choice of Community Standards in Federal Obscenity Proceedings: The Role of the Constitution and the Common Law*, 64 VA. L. REV. 399 (1978); *see generally* W. LOCKHART, Y. KAMISAR & J. CHOPER, *supra* note 69, at 901–02.

[263] Butler v. Michigan, 352 U.S. 380, 383 (1957) ("rugged" used to distinguish "grown men and women" from "juvenile innocence").

[264] Roth v. United States, 354 U.S. 476, 489 (1957).

[265] *See, e.g.*, U.S. v. Kennerly, 209 F. 119 (S.D.N.Y. 1913); MacFadden v U.S., 165 F. 51 (3d Cir. 1909); U.S. v. Bennett, 24 F. Cas. 1093 (S.D.N.Y. 1879).

regulation of materials viewed as obscene by the "average person, applying contemporary community standards."[266]

A nonsegregable audience is analogous to the "community" that is the focus of the community standards doctrine and should be treated similarly. Content-based restraints on the exercise of some individual free speech rights should be allowed to advance audience-wide interests in avoiding offensive communications and protecting children. In accordance with this view, Justice Powell's concurrence in *Pacifica* came close to acknowledging that the nonsegregable nature of an audience might alone justify some level of content-based regulation.[267] While arguing that the Court should not engage in content-based determinations of "value" in regulating protected speech,[268] Powell concurred, as Justice Stevens aptly noted, "in a judgment that could not otherwise stand."[269] Moreover, Powell addressed the nonsegregability problem directly:

> In most instances, the dissemination of this kind of speech to children may be limited without also limiting willing adults' access to it . . . . The difficulty is that . . . a physical separation of the audience [possible in other media] cannot be accomplished in the broadcast media. During most of the broadcast hours, both adults and unsupervised children are likely to be in the broadcast audience, and the broadcaster cannot reach willing adults without also reaching children. This . . . is one of the distinctions between the broadcast and other media to which we often have adverted as justifying a different treatment of the broadcast media for First Amendment purposes.[270]

Given that Powell voted for content-based regulation in *Pacifica*, such regulation must be one of the "different treatment[s]"[271] of broadcasting justified by the nonsegregability of the audience.

If some level of content-based regulation of speech directed to a nonsegregable audience is permissible, the corollary proposition that members of such an audience may be required to shift to alternative modes of communication if they wish to continue receiving the same speech is also established. These

---

[266] *Roth*, 354 U.S. at 489.
[267] *Pacifica*, 438 U.S. at 758–59 (Powell, J., concurring in part).
[268] *Id.* at 761.
[269] *Id.* at 745 n.20 (opinion of the Court).
[270] *Id.* at 758–59 (Powell, J., concurring in part) (citations omitted).
[271] *Id.* at 759.

shifts are never truly costless.[272] The degree of hardship involved in shifting to these alternatives is a critical factor in judging the degree to which the state should be permitted to undertake content-based regulation. A court called upon to make such a judgment must make a number of difficult determinations: the degree of hardship imposed on those whose expressive activities have been regulated, the amount of free expression that may actually be chilled, and perhaps even the value[273] of any speech that is chilled.

Thus, some degree of content-based regulation of communications to nonsegregable audiences can be justified by analogy to the community standards doctrine, which attempts to balance both individual and community interests in free speech. Justifications for restricting adult free speech rights to protect children—the central concern of the original *Butler* doctrine—still must be articulated. The rights of children whose parents choose to permit them access to sexually explicit materials must also be addressed.[274]

b. *Regulating Adult Free Speech to Protect Children.* The state's strong interests in protecting children and in supporting parental supervision of children[275] weigh heavily in favor of not invalidating attempts to regulate sexually explicit materials available to nonsegregable audiences. Nevertheless, the resulting infringement of the free speech rights of adults, previously protected in full by the *Butler* doctrine, must be independently justified. The intergenerational relationship between adults and children serves as one justification for imposing some obligation on adults in a nonsegregable audience to refrain from exercising

---

[272] *See* FCC v. Pacifica Foundation, 438 U.S. 726, 774 (1978) (Brennan, J., dissenting).

[273] *See* New York v. Ferber, 458 U.S. 747, 762–63 (1982) (child pornography can be banned if adequately defined by applicable state law, even if material falls short of *Miller* obscenity standard); *see also* FCC v. Pacifica Foundation, 438 U.S. 726, 744–48 (1978); Young v. American Mini Theatres, Inc., 427 U.S. 50, 70 (1976) (opinion of Stevens, J.) (zoning regulation for adult movie theatre upheld; although First Amendment protects erotic material with some artistic value from total suppression, society's interest in protecting this type of expression is of a different and lesser magnitude than its interest in "untrammeled political debate"). While a given form of expression might have an ascertainable "value" in furthering the societal rationales for free speech, such as advancing knowledge or facilitating political participation, the proposition that it has an ascertainable "value" in furthering individual rationales for free speech, such as individual autonomy and self-fulfillment, seems more dubious. *See* T. EMERSON, *supra* note 74, at 6–7; *see also* Redish, *supra* note 73, at 594–95, 635–40.

[274] FCC v. Pacifica Foundation, 438 U.S. 726, 768–70 (1978) (Brennan, J., dissenting).

[275] *See supra* text accompanying notes 128–30.

their free speech rights in a way that tends to expose children to sexually explicit materials.

The First Amendment character of this intergenerational relationship is based on the fact that the robust exercise of adult First Amendment freedoms is made possible only by years of careful nurturing during childhood.[276] Adults, who are the presumed beneficiaries of this process, should defer to the nurturing process and seek alternative modes for receiving communications likely to damage or to disrupt this process.

The long-term durability of free speech thus justifies minor infringements of the free speech rights of adults in nonsegregable audiences in order to support parental authority in nurturing children who are also part of the nonsegregable audience. To minimize the restraints placed on adults' rights, the intergenerational obligation should be confined to nonsegregable audiences containing substantial numbers of children. To prevent this obligation from becoming a pretense for censorship of an unduly broad array of influences on children, it should apply only to obscene or extremely offensive sexually explicit materials, which are uniquely capable of disrupting the nurturing process.

c. *Impairing Parental Choices for Children.* The impact of modifications to the *Butler* doctrine on children whose parents choose to permit them access to sexually explicit materials should be analyzed separately. Although these children have a First Amendment right to such materials, derived from their parents' free speech rights,[277] the children's derivative free speech rights should be diminished to the same degree that their parents' rights are diminished due to the parents' intergenerational obligation.[278] The children who are to benefit from the imposition of this obligation are affected as significantly by exposure to communications directed toward adults as by those directed toward children of permissive parents.

Restrictions on the free speech rights of children of permissive parents might be construed as an indirect imposition of the state's notions of what materials are fit for children, notwithstanding express parental choices to the contrary.[279] State inter-

---

[276] *See* Garvey, *supra* note 107, at 328–33; Hafen, *supra* note 122, at 651–58.
[277] *See supra* text accompanying notes 143–46.
[278] *See supra* notes 275–76 and accompanying text.
[279] *Cf.* Wisconsin v. Yoder, 406 U.S. 205 (1972) (upholding right of Amish parents to withdraw children from public school prior to reaching state-mandated minimum levels of education).

ference with parental choices in this situation may be attributed directly to the fact that the audience is nonsegregable. Such interference is an unavoidable side-effect of the state's effort to support choices made by other parents. Because this effort to protect children from obscenity by audience-wide regulation is not directed at ferreting out and punishing parents who deviate from the state's conception of proper childraising,[280] regulations of general application should be justifiable whether or not the state shows that there is a clear risk of harm to a child.

C. *Applying a Modified* Butler *Analysis*

In the special circumstances of a nonsegregable audience, the *Butler* doctrine should be modified to permit content-based regulation of sexually explicit materials that ordinarily may be regulated only as to children,[281] whenever adults have access, with reasonable effort, to alternative modes of communication.

The result in *Pacifica* is fully consistent with this modified version of the *Butler* doctrine. Daytime radio broadcasts in *Pacifica* reached a nonsegregable audience of adults and children while late evening broadcasts arguably did not.[282] The FCC was thus authorized to regulate daytime broadcasting of "indecent" speech, even though such regulation would subject adults during that time to receiving only communications fit for children,[283] in part because the reasonable alternatives of late evening broadcasts, records, tapes, and live performances were available to adults.[284] This array of alternatives offered the same communicative content to adults at a wide range of added expense and inconvenience. Hence, the Court suggested that adults could avail themselves of these various alternatives.[285]

Though consistent with the modified *Butler* doctrine, the *Pacifica* case does not govern the constitutionality of the federal

---

[280] Prince v. Massachusetts, 321 U.S. 158 (1944) (parent convicted of violating child labor laws).

[281] *See supra* text accompanying notes 138–39.

[282] *See* Citizen's Complaint Against Pacifica Foundation Station WBAI (FM) New York, N.Y., 56 F.C.C.2d 94, 98 (1975), *modified*, 59 F.C.C.2d 892 (1976), *rev'd*, 556 F.2d 9 (1977), *rev'd*, 438 U.S. 726 (1978) ("When the number of children in the audience is reduced to a minimum, for example during the late evening hours, a different standard might conceivably be used.").

[283] *See* FCC v. Pacifica Foundation, 438 U.S. 726, 750 (1978).

[284] *Id.* at 750 n.28.

[285] *See id.* Justice Brennan in dissent argued that the cost of some alternatives was too high. *See id.* at 774–75 (Brennan, J., dissenting).

response to the dial-a-porn problem because the alternatives that this response leaves open to adults appear more onerous than those suggested in *Pacifica*. Time-of-day regulation of dial-a-porn content results in both inconvenient and few "adult" hours per day.[286] The virtual impossibility of screening callers by age, while still operating such a system with automatic answering equipment, will effectively ban free dial-a-porn, because it would be economically infeasible to staff a free system with live performers or even with live operators playing taped performances.[287] Thus, unlike the late evening option in *Pacifica*, there appears to be no zero-cost alternative for adults.

Other alternatives to dial-a-porn include the $15 to $30 "for pay" services,[288] which offer live performances. The price of the "for pay" system probably could be reduced substantially, though not to zero, by converting it to taped rather than live performances. An adult could also buy or rent tapes or records, though an attempt to replicate exactly the content of dial-a-porn in this fashion could prove exorbitant because dial-a-porn tapes are changed frequently.[289] Taken together, these possible alternatives are more expensive and inconvenient than those suggested in *Pacifica*. The new statute therefore presents problems beyond the facts of *Pacifica*.[290]

Finally, the applicability of *Pacifica* is further undermined by the lack of unwilling listeners in the dial-a-porn situation. Protecting unwilling listeners and children were the twin justifications for upholding regulation of "indecent" broadcasting in *Pacifica*.[291] In the dial-a-porn context, callers will generally not be unwilling listeners.

As proposed and developed, the modified *Butler* doctrine offers a rationale for not confining *Pacifica* to its facts and generalizes the treatment of new communication media that reach nonsegregable audiences. The first step in applying this doctrine to the dial-a-porn problem is isolating dial-a-porn's nonsegre-

---

[286] The United States and its possessions cover eight time zones. Uniform Time Act of 1966, 15 U.S.C. § 261 (1982). Simultaneous "adult hours" across all eight zones would therefore be confined to the very early hours of the morning in the eastern United States.

[287] Further Notice of Inquiry and Notice of Proposed Rulemaking, *supra* note 12, at 2125.

[288] *See supra* note 8.

[289] Notice of Inquiry, *supra* note 9, at 43,349 (messages changed "at least once daily").

[290] *See supra* text accompanying notes 229–31, 243–45.

[291] *See Pacifica*, 438 U.S. at 748–50.

gable audience. Because dial-a-porn uses the telephone system, its audience includes everyone in the country who has access to a telephone and a dial-a-porn number. This audience is completely nonsegregable in terms of a caller's age and geographic location.[292] The audience will also contain substantial numbers of children and adolescents.[293]

Thus characterized, a nonsegregable audience merits content-based regulation of sexually explicit materials transmitted to it whenever such materials may be regulated as to children. The statute conforms to this standard by banning only "obscene or indecent" communications, which may be fully regulated as to children.[294]

The regulatory scheme enacted by the statute also leaves adults with an array of reasonable alternatives for access to dial-a-porn messages. These alternatives appear to protect substantially the free speech interests of adults in receiving sexually explicit communications over the telephone from an automatic tape machine.

VI. Conclusion

Assuming that its facial overbreadth defects[295] can be cured by minor redrafting, the recent an.endment to section 223 of the Communications Act of 1934, designed to protect children from sexually oriented telephone entertainment services, should not be rendered unconstitutional by its effect of reducing adults, in certain circumstances, to receiving only communications fit for children, in apparent violation of *Butler v. Michigan.* When a new communication technology inseparably unites adults and children in the same audience, and when adults in that audience can readily receive the same communications from alternative sources, the *Butler* rule should be modified to allow some content-based regulation of sexually explicit communications directed at such an audience.

---

[292] *See supra* text accompanying notes 199–210.

[293] *See* N.Y. Times, Feb. 19, 1984, § 1, at 80, col. 5 (sampling of public comments responding to FCC Further Notice of Inquiry and Notice of Proposed Rulemaking, *supra* note 12).

[294] 47 U.S.C.A. § 223(b)(1)(A) (West Supp. 1984); *see supra* text accompanying notes 137–45.

[295] *See supra* text accompanying notees 181–83.

APPENDIX

TEXT OF SECTION 223 OF THE COMMUNICATIONS ACT OF 1934, 47 U.S.C. § 223, AS AMENDED.

NOTE: Section 223(b) and the bracketed portions of section 223(a) were added by section 8 of the Federal Communications Commission Authorization Act of 1983, Pub. L. No. 98-214, 1984 U.S. CODE CONG. & AD. NEWS (97 Stat.) 1467 (effective Dec. 8, 1983).

Section 223.

Obscene or harassing telephone calls in the District of Columbia or in interstate or foreign communications.

[(a)] **Whoever—**

(1) in the District of Columbia or in interstate or foreign communication by means of telephone—

(A) makes any comment, request, suggestion or proposal which is obscene, lewd, lascivious, filthy, or indecent;

(B) makes a telephone call, whether or not conversation ensues, without disclosing his identity and with intent to annoy, abuse, threaten, or harass any person at the called number;

(C) makes or causes the telephone of another repeatedly or continuously to ring, with intent to harass any person at the called number; or

(D) makes repeated telephone calls, during which conversation ensues, solely to harass any person at the called number; or

(2) knowingly permits any telephone [facility] under his control to be used for any purpose prohibited by this section,

shall be fined not more than [$50,000] or imprisoned not more than six months, or both.

(b)

(1) Whoever knowingly—

(A) in the District of Columbia or in interstate or foreign communication, by means of telephone, makes (directly or by recording device) any obscene or indecent communication for commercial purposes to any person under eighteen years of age or to any other person without that person's consent, regardless of whether the maker of such communication placed the call; or

(B) permits any telephone facility under such person's control to be used for an activity prohibited by subparagraph (A), shall be fined not more than $50,000 or imprisoned not more than six months, or both.

(2) It is a defense to a prosecution under this subsection that the defendant restricted access to the prohibited communication to persons eighteen years of age or older in accordance with procedures which the [Federal Communications] Commission shall prescribe by regulation.

(3) In addition to the penalties under paragraph (1), whoever, in the District of Columbia or in interstate or foreign communication, intentionally violates paragraph (1)(A) or (1)(B) shall be subject to a fine of not more than $50,000 for each violation. For purposes of this paragraph, each day of violation shall constitute a separate violation.

(4)

(A) In addition to the penalties under paragraphs (1) and (3), whoever, in the District of Columbia or in interstate or foreign communication, violates paragraph (1)(A) or (1)(B) shall be subject to a civil fine of not more than $50,000 for each violation. For purposes of this paragraph, each day of violation shall constitute a separate violation.

(B) A fine under this paragraph may be assessed either—

(*i*) by a court, pursuant to a civil action by the Commission or any attorney employed by the Commission who is designated by the Commission for such purposes, or

(*ii*) by the Commission after appropriate administrative proceedings.

(5) The Attorney General may bring a suit in the appropriate district court of the United States to enjoin any act or practice which violates paragraph (1)(A) or (1)(B). An injunction may be granted in accordance with the Federal Rules of Civil Procedure.

COMMENT

UNION RESTRICTIONS ON THE RIGHT TO RESIGN: A PROPOSAL FOR A NEW REASONABLENESS TEST

DENNIS C. SHEA*

In an attempt to prevent reduction of their ranks during strikes, unions have incorporated provisions into their constitutions that restrict the right of their members to resign.[1] Ag-

* A.B., Harvard University, 1983; A.M., Harvard University, 1983; member, Class of 1986, Harvard Law School.

[1] The International Association of Machinists and Aerospace Workers, AFL-CIO, has adopted the following rule in its constitution:

> *Improper Conduct of a Member*: . . . Accepting employment in any capacity in an establishment where a strike or lockout exists as recognized under this Constitution, without permission. Resignation shall not relieve a member of his obligation to refrain from accepting employment at the establishment for the duration of the strike or lockout or within 14 days preceding its commencement . . . .

INTERNATIONAL ASSOCIATION OF MACHINISTS AND AEROSPACE WORKERS CONST., *quoted in* Machinists Local 1327 v. NLRB (Dalmo Victor), 725 F.2d 1212, 1214 (9th Cir. 1984).

The constitution of the Pattern Makers' League of North America, AFL-CIO, includes a similar provision. This provision, entitled League Law 13, states: "[n]o resignation or withdrawal from an Association, or from the League, shall be accepted during a strike or lockout, or at a time when a strike or lockout appears imminent." PATTERN MAKERS' LEAGUE CONST., Law 13, *quoted in* Pattern Makers' League of North America v. NLRB, 724 F.2d 57, 58 (7th Cir. 1983), *cert. granted* 105 S. Ct. 79 (1984) (No. 83-1894), *argued* Feb. 27, 1985, 53 U.S.L.W. 3632 (U.S. Feb. 12, 1985), *reargument ordered*, 53 U.S.L.W. 3686 (U.S. Mar. 26, 1985).

The United Auto Workers ("UAW") has permitted resignations only if tendered within the ten days at the end of the union's fiscal year. This provision reads:

> A member may resign or terminate his membership only if he is in good standing, is not in arrears or delinquent in the payment of any dues or other financial obligation to the International Union or to his Local Union and there are no charges filed and pending against him. Such resignation or termination shall be effective only if by written communication, signed by the member, and sent by registered or certified mail, return receipt requested, to the Financial Secretary of the Local Union within the ten (10) day period prior to the end of the fiscal year of the Local Union as fixed by this Constitution, whereupon it shall become effective sixty (60) days after the end of such fiscal year . . . .

UNITED AUTO WORKERS UNION CONST. art. 6, § 17 (1977). Although the UAW has argued that this provision is necessary as a way of maintaining solidarity during strike periods, the NLRB has consistently rejected this and other arguments offered by the UAW in defense of the provision. *See, e.g.*, UAW Local 647 (General Electric Corp.), 197 N.L.R.B. 608, 609 (1972) (ruling that the escape route offered by the constitutional provision is inadequate); Local 1384, United Auto Workers (Ex-Cell-O Corp.), 219 N.L.R.B. 729, 730 (1975) (affirming the administrative law judge's finding that the UAW's constitutional obstacle to resignation could not bar the resignations of members). For a discussion of the litigation arising from the UAW constitutional provision, see Gould, *Solidarity Forever—or Hardly Ever: Union Discipline, Taft-Hartley, and the Right of Union Members to Resign*, 66 CORN. L. REV. 74, 87–91 (1980).

grieved union members, through the National Labor Relations Board ("NLRB") and the courts, have recently challenged the legality of these restrictive provisions[2] under the National Labor Relations Act.[3] Opponents of the provisions contend[4] that they violate section 8(b)(1)(A) of the Act by "restraining or coercing employees in the exercise of the rights guaranteed in [section 7]."[5] Although section 7 of the Act grants employees "the right to self-organization, to form, join, or assist labor organizations, to bargain collectively through representatives of their own choosing, and to engage in other concerted activities for the purpose of collective bargaining or other mutual aid or protection," it also grants employees the right "to refrain from any or all such activities . . . ."[6] By charging that the right to resign

---

[2] In litigation involving the Pattern Makers' League of North America, the union fined eleven employees who had resigned their memberships and had returned to work during a strike. These employees subsequently filed a complaint before the NLRB challenging the union's authority to restrict their resignations. Both the NLRB and the Court of Appeals for the Seventh Circuit held that the union had violated section 8(b)(1)(A) by imposing fines on the employees pursuant to the union's constitutional provision prohibiting resignations during a strike. Pattern Makers' League of North America v. NLRB, 724 F.2d 57, 59–60 (7th Cir. 1983), *cert. granted* 105 S. Ct. 79 (1984) (No. 83-1894), *argued* Feb. 27, 1985, 53 U.S.L.W. 3632 (U.S. Feb. 12, 1985), *reargument ordered*, 53 U.S.L.W. 3686 (U.S. Mar. 26, 1985). In litigation involving the International Association of Machinist and Aerospace Workers, however, the Court of Appeals for the Ninth Circuit held that a restriction on resignation during a strike is a reasonable rule protected by the proviso to section 8(b)(1)(A). Machinists Local 1327 v. NLRB (Dalmo Victor), 725 F.2d 1212, 1217–18 (9th Cir. 1984). This statutory proviso permits a union to "prescribe its own rules with respect to the acquisition or retention of [union] membership." 29 U.S.C. § 158(b)(1)(A) (1982).

The Supreme Court granted certiorari in October of 1984 to resolve this apparent conflict between the two circuits. *Pattern Makers*, 105 S. Ct. 79. At the time of this writing, the Court had not yet announced its decision.

[3] National Labor Relations Act, §§ 1–9, 29 U.S.C. §§ 151–59 (1982).

[4] *See* Brief for Respondent at 12–21, Pattern Makers' League of North America v. NLRB, No. 83-1894 (U.S. filed May 18, 1984), *cert. granted* 105 S. Ct. 79 (1984), *argued* Feb. 27, 1985, 53 U.S.L.W. 3632 (U.S. Feb. 12, 1985), *reargument ordered*, 53 U.S.L.W. 3686 (U.S. Mar. 26, 1985).

[5] 29 U.S.C. § 158(b)(1)(A) (1982). The full text of section 8(b)(1)(A) reads as follows:

> It shall be an unfair labor practice for a labor organization or its agents (1) to restrain or coerce (A) employees in the exercise of the rights guaranteed in Section 7: *Provided*, That this paragraph shall not impair the right of a labor organization to prescribe its own rules with respect to the acquisition or retention of membership therein . . .

[6] *Id.* § 157. The "right to refrain" language was first incorporated in the Taft-Hartley Amendments of 1947. Labor-Management Relations Act, ch. 114, § 7, 61 Stat. 140. Concerned by certain widespread "bad practices," Congress adopted these amendments partly to protect employees from union coercion. The provision "is designed to protect members of those unions that . . . treat their members as pawns and exploit them . . . . [I]t is incumbent upon us . . . to assure to the employees whom we subject to union control some voice in the union's affairs." H.R. REP. No. 245, 80th Cong., 1st Sess. 28 (1947); *see also* A. COX, D. BOK & R. GORMAN, CASES AND MATERIALS ON LABOR LAW 83–85 (9th ed. 1981).

during a strike is a necessary component of the general right "to refrain from union activities,"[7] opponents of these constitutional provisions argue that the provisions impermissibly abridge employees' section 7 rights.

This Comment will argue that the decision to uphold a constitutional provision restricting resignations during a strike should depend upon the presence of a union security clause[8] in the collective bargaining agreement between the union and the employer.[9] Courts should allow unions to restrict member resignations during a strike, or during the period immediately preceding a strike,[10] only if there is no such security clause. Conversely, if a security clause is present in the agreement, courts should strike down such provisions as abridgments of employees' section 7 rights.

As this Comment will demonstrate, the presence of a union security clause can quite properly influence the balance struck between an employee's section 7 right to refrain from union activities and a union's institutional interests in maintaining the solidarity of its members and the resiliency of its treasury during a strike. An analysis of the arguments for and against union restrictions on resignations highlights the delicacy of this balance. Thus, Part I of the Comment will outline the NLRB's current opposition to these restrictive provisions. Part II will

---

[7] *See* Brief for Respondent at 28–29, Pattern Makers' League of North America v. NLRB, No. 83-1894 (U.S. filed May 18, 1984), *cert. granted* 105 S. Ct. 79 (1984), *argued* Feb. 27, 1985, 53 U.S.L.W. 3632 (U.S. Feb. 12, 1985), *reargument ordered*, 53 U.S.L.W. 3686 (U.S. Mar. 26, 1985). *But see* Brief for the Petitioners at 22–27, Pattern Makers' League of North America v. NLRB, No. 83-1894 (U.S. filed May 18, 1984), *cert. granted* 105 S. Ct. 79 (1984), *argued* Feb. 27, 1985, 53 U.S.L.W. 3632 (U.S. Feb. 12, 1985), *reargument ordered*, 53 U.S.L.W. 3686 (U.S. Mar. 26, 1985).

[8] A union security clause is a provision in a collective bargaining agreement "whereby the employer agrees to require his employees, as a condition of their employment, to affiliate with the union in some way." T. HAGGARD, COMPULSORY UNIONISM, THE NLRB, AND THE COURTS 4 (1977) (Labor Relations and Public Policy Series Report No. 15). There are a number of different types of union security provisions. They include the closed shop agreement, the union shop agreement, the agency shop agreement, the maintenance of membership agreement, and the representation fee agreement. These provisions will be defined in Part III of this Comment. *See infra* notes 95–99 and accompanying text.

[9] Under § 9(a) of the Act, a union enjoys the exclusive right to represent all the employees—both union members and nonmembers—in a single bargaining unit. 29 U.S.C. § 159(a) (1982). As the exclusive representative of the unit, the union is obligated to bargain collectively with the employer. 29 U.S.C. § 158(d) (1982).

[10] This Comment will not attempt to define the period preceding a strike during which a restriction on resignations would be reasonable. One commentator, however, would permit union restrictions on resignations "ten to fifteen days after negotiations have commenced if those negotiations are initiated approximately sixty days before the contract expires." Gould, *supra* note 1, at 100.

describe the thirty-day reasonable restriction rule formerly offered by the NLRB and rejected by the Ninth Circuit. Finally, Part III will offer a new analytic framework for evaluating the reasonableness of union rules restricting resignations.

I. THE CURRENT POSITION OF THE NLRB: NO RESTRICTIONS ON RESIGNATIONS

In two recent decisions—*International Association of Machinists and Aerospace Workers, Local Lodge 1414 v. Neufeld Porsche-Audi, Inc.*,[11] and *Pattern Makers' League of North America v. NLRB*[12]—the NLRB has declared that the Act permits no restriction on the right of union members to resign. In making this declaration, the NLRB has adopted the position expressed in the concurring opinion of an earlier NLRB decision, *Machinists Local 1327 (Dalmo Victor)*.[13]

All three decisions describe similar scenarios: after participating in a strike, a small group of employees decided to resign from their unions and return to work.[14] The unions responded to these defections by fining the members who had tendered their resignations.[15] The unions justified these fines by pointing

[11] 270 N.L.R.B. No. 209, 1983–1984 NLRB Dec. (CCH) ¶ 16,436 (1984).

[12] 265 N.L.R.B. 1332 (1982), *enforced*, 724 F.2d 57 (7th Cir. 1983), *cert. granted*, 105 S. Ct. 79 (1984) (No. 83-1894).

[13] 263 N.L.R.B. 984, 988 (1982) (Van de Water & Hunter, concurring), *enforcement denied*, 725 F.2d 1212 (9th Cir. 1984). This decision held that the constitutional provision was an unreasonable restriction on employees' section 7 rights because employees could resign only during nonstrike periods. The majority would have found that a union could impose a general 30-day restriction on the right to resign. 263 N.L.R.B. at 987. The concurring opinion by Van de Water and Hunter, however, would have disallowed any restrictions on resigning. This opinion was expressly adopted by the NLRB in *Neufeld Porsche-Audi*. 270 N.L.R.B. No. 209, 1983–84 NLRB Dec. (CCH) ¶ 16,436 (1984), at 28,094. By adopting the concurring opinion, the NLRB has thereby overruled its decision in *Dalmo Victor*, although for reasons different from those offered by the Ninth Circuit in denying enforcement of that decision. *See* Machinists Local 1327 v. NLRB (Dalmo Victor), 725 F.2d 1212, 1217–18 (9th Cir. 1984).

[14] In *Neufeld Porsche-Audi*, for example, only one employee resigned during the strike. 270 N.L.R.B. No. 209, 1983–84 NLRB Dec. (CCH) ¶ 16,436 (1984), at 28,092. In *Dalmo Victor*, however, three employees had resigned their memberships and returned to work. 263 N.L.R.B. at 984, 725 F.2d at 1214. The number of resigning employees was even greater in *Pattern Makers*; at least eleven employees had tendered their resignations to the union during the strike. 265 N.L.R.B. at 1332 n.4, 724 F.2d at 58.

[15] These fines varied in their severity. In *Neufeld Porsche-Audi*, for example, the fine imposed by the International Association of Machinists and Aerospace Workers on the individual employee amounted to a substantial $2,250. 270 N.L.R.B. No. 209, 1983–84 NLRB Dec. (CCH) ¶ 16,436 (1984), at 28,093. The fines imposed by the Pattern Makers' League of North America on the disgruntled employees were "in an amount roughly equal to their earnings during the strike." Brief for Respondent at 3–4, Pattern Makers' League of North America v. NLRB, No. 83-1894 (U.S. filed May 18, 1984), *cert. granted* 105 S. Ct. 79 (1984), *argued* Feb. 27, 1985, 53 U.S.L.W. 3632 (U.S. Feb. 12, 1985), *reargument ordered*, 53 U.S.L.W. 3686 (U.S. Mar. 26, 1985).

to provisions in their constitutions that prohibited resignations during a strike. In both *Neufeld Porsche-Audi* and *Dalmo Victor*, the contested constitutional provisions also restricted resignations during the fourteen days preceding the commencement of the strike.[16] In *Pattern Makers* the constitutional provision is less specific, prohibiting resignations during the strike and during the days "when a strike . . . appears imminent."[17]

The NLRB's determination that these provisions violate section 8(b)(1)(A) by impairing the section 7 right of an employee to refrain from concerted activities is a result of its consideration of three complementary sources of analysis. These sources are (1) statutory interpretation of the place of section 7 rights in the scheme created by the Act; (2) judicial doctrines concerning the boundaries between a permissible regulation of an internal union affair and an impermissible regulation of an affair external to union concerns; and (3) the three-part test laid down by the Supreme Court in *Scofield v. NLRB*.[18] Although the NLRB has not explicitly organized its decisions in *Neufeld Porsche-Audi*, *Pattern Makers*, and the *Dalmo Victor* concurrence in the systematic fashion described below, the division of these decisions into three separate lines of analysis presents a clearer picture of the reasons underlying the NLRB's position.

A. *The Section 7 Right to Refrain and the Union Interest in Maintaining Solidarity*

The first line of analysis argues that the right to resign is a part of the section 7 right of employees "to refrain from engaging

---

[16] The same International Association of Machinists and Aerospace Workers constitutional provision quoted in note 1, *supra*, was contested in *Neufeld Porsche-Audi*. 270 N.L.R.B. No. 209, 1983–84 NLRB Dec. (CCH) ¶ 16,436 (1984), at 3. Initially, the NLRB did not consider this provision to be a restriction on the member's right to resign. *See* Machinists Local 1327 v. NLRB (Dalmo Victor), 231 N.L.R.B. 719 (1977), *enforcement denied*, 608 F.2d 219 (9th Cir. 1979). Instead, it viewed the provision as "an unlawful attempt . . . to restrict the post-resignation conduct of former members." 231 N.L.R.B. at 984. When the Ninth Circuit first reviewed the NLRB's decision on appeal, it disagreed with the NLRB's construction of the provision and remanded to the NLRB to consider the validity of the provision as a restriction on resignations. 608 F.2d at 1222.

[17] For the text of the union constitutional provision challenged in *Pattern Makers*, see *supra* note 1.

[18] 394 U.S. 423 (1969).

in union activities."[19] Although conceding that it may be necessary to balance the section 7 right against "a corresponding *right* of relatively equal import and legal significance,"[20] the current NLRB insists that it is inappropriate to balance a statutory right—the section 7 right—against a union's institutional goal of maintaining solidarity.[21] Because this goal is not explicitly protected in the statute, the NLRB treats it not as a *right*, but only as an *interest*.

According to this view, a hierarchy of values exists in which explicitly listed statutory rights hold a position superior to that of so-called interests that do not enjoy a similar congressionally-endowed status. In addition, the respective positions of a right and an interest can never overlap. To balance a statutory right with a nonstatutory interest would constitute, in the NLRB's view, an act of legislation beyond the competence of the NLRB itself.[22] Consequently, the first line of analysis reduces to a straightforward maxim: statutory rights outweigh mere institutional interests, even though advancement of these interests may be central to the union's effectiveness as the exclusive bargaining agent.

B. *The Internal-External Distinction in* Allis-Chalmers *and its Progeny*

The second line of analysis applies a cluster of Supreme Court decisions that have examined the authority of a union to discipline its members for breaking certain union regulations.[23] In

[19] *See* International Association of Machinists and Aerospace Workers, Local Lodge 1414 v. Neufeld Porsche-Audi, Inc., 270 N.L.R.B. No. 209, 1983–84 NLRB Dec. (CCH) ¶ 16,436 (1984), at 28,096; Pattern Makers' League of North America v. NLRB, 265 N.L.R.B. 1332, 1333 (1982), *enforced*, 724 F.2d 57 (7th Cir. 1983), *cert. granted*, 105 S. Ct. 79 (1984) (No. 83-1894); Machinists Local 1327 (Dalmo Victor), 263 N.L.R.B. 984, 986 (1982), *enforcement denied*, 725 F.2d 1212 (9th Cir. 1984).

[20] Machinists Local 1327 (Dalmo Victor), 263 N.L.R.B. 984, 990–91 (1982) (Van de Water & Hunter, concurring), *enforcement denied*, 725 F.2d 1212 (9th Cir. 1984).

[21] 263 N.L.R.B. at 991 (Van de Water & Hunter, concurring); *see also* International Association of Machinists and Aerospace Workers, Local Lodge 1414 v. Neufeld Porsche-Audi, Inc., 270 N.L.R.B. No. 209, 1983–84 NLRB Dec. (CCH) ¶ 16,436 (1984), at 28,097 (citing with approval the concurring opinion in *Dalmo Victor*).

[22] Machinists Local 1327 (Dalmo Victor), 263 N.L.R.B. 984, 991 (1982) (Van de Water & Hunter, concurring), *enforcement denied*, 725 F.2d 1212 (9th Cir. 1984).

[23] This cluster of Supreme Court decisions includes NLRB v. Allis-Chalmers Mfg. Co., 388 U.S. 175 (1967), *reh'g denied*, 389 U.S. 892 (1967); Scofield v. NLRB, 394 U.S. 423 (1969); NLRB v. Granite State Joint Board, Textile Workers Union, 409 U.S. 213 (1972); Booster Lodge No. 405, International Association of Machinists v. NLRB, 412 U.S. 84 (1973). For a discussion and analysis of these decisions, see Gould, *supra*

NLRB v. Allis-Chalmers Mfg. Co.,[24] the progenitor of these decisions, the Court established that the exercise of union disciplinary power over employees who maintained full membership[25] in the union did not constitute "restraint" or "coercion" of a section 7 right within the meaning of section 8(b)(1)(A).[26] Distinguishing between the internal and the external enforcement of union rules, the Court comprehensively reviewed the legislative history of section 8(b)(1)(A)[27] and concluded that "Congress did not propose any limitations with respect to the *internal* affairs of unions . . . ."[28] Consequently, the Court permitted the United Auto Workers ("UAW"), the union whose disciplinary rule was challenged in *Allis-Chalmers*, to impose fines on its members who had crossed a picket line and returned to work during a strike.[29] Because these members enjoyed full union benefits and "had fully participated in the proceedings leading to the strike,"[30] the Court reasoned that fining them was a legitimate exercise of the union's power to regulate internal matters. In an attempt to clarify the distinction between an internal and an external regulation, the Court identified a category of external actions that were proscribed by section 8(b)(1). These actions either interfered with an employee's em-

---

note 1, at 76–86; Note, *Restrictions on the Right to Resign: Can a Member's Freedom to "Escape the Union Rule" Be Overcome by Union Boilerplate?*, 42 GEO. WASH. L. REV. 397, 400–05 (1974) [hereinafter cited as *Restrictions*]; Note, *Union Disciplinary Fines and the Right to Resign*, 30 WASH. & LEE L. REV. 664, 667–81 (1973); Note, *The Inherent Conflict Between Sections 7 and 8(b)(1)(A) of the National Labor Relations Act—Union Attempts to Discipline Resigning Strikebreakers*, 1978 WIS. L. REV. 859, 859–69 [hereinafter cited as *Conflict*].

[24] 388 U.S. 175 (1967), *reh'g denied*, 389 U.S. 892 (1967).

[25] In *Allis-Chalmers* the Court distinguished full union membership from membership requiring only the payment of monthly union dues. 388 U.S. at 196–97. The implications of this distinction for union discipline have been pointed out by several commentators. *See, e.g.*, Gould, *supra* note 1, at 82 ("[The *Allis-Chalmers* Court] implied that a worker whose union 'membership' was limited to paying dues and initiation fees should be less vulnerable to sanctions than a 'full' union member"); *Restrictions*, *supra* note 23, at 410 ("The distinction between a full member and a so-called financial core member raises the crucial question whether the financial core member is bound by a strike vote of the union membership") (citation omitted).

[26] *Allis-Chalmers*, 388 U.S. at 195; *see id.* at 178–95.

[27] *Id.* at 184–95. The Court also briefly discussed the legislative history of section 8(b)(2), which prohibits a union from compelling an employer to discharge a former union member, except for "failure of an 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(b)(2) (1982).

[28] *Allis-Chalmers*, 388 U.S. at 195 (emphasis added).

[29] *Id.* at 177; *see id.* at 195–96.

[30] *Id.* at 196.

ployment status,[31] or interfered with the rights of nonmembers or employees outside the bargaining unit.[32]

Two years later, in *Scofield v. NLRB*,[33] the Supreme Court was asked to determine whether a union rule imposing a ceiling on the amount of production for which an employee could receive compensation violated section 8(b)(1)(A).[34] Finding that the rule was lawful, the Court rejected, as it had in *Allis-Chalmers*, a claim by the aggrieved union members that the rule infringed on their section 7 right to refrain from concerted activities. Although the Court in *Scofield* admitted that the ceiling on production had certain external effects,[35] it nonetheless classified the contested union rule as an internal regulation designed to "strengthen the union's hand in bargaining."[36]

The distinction between legitimate internal regulations and illegitimate external regulations was clarified in *NLRB v. Granite State Joint Board, Textile Workers Union*[37] and in *Booster Lodge No. 405, International Association of Machinists v. NLRB*.[38] In both cases, the Supreme Court held that the unions had violated section 8(b)(1)(A) by imposing fines, pursuant to rules prohibiting strikebreaking by members, on employees who had resigned union membership and returned to work during a strike. In *Granite State*, the union membership voted shortly after the strike began to institute a rule prohibiting any member from "aiding or abetting the employer" during the remaining days of the strike.[39] Notwithstanding this rule, thirty-one union

---

[31] *Id.* at 195; *see also* Machinists Local 1327 (Dalmo Victor), 263 N.L.R.B. 984, 988 (1982) (Van de Water & Hunter, concurring) (citing interpretation of section 8(b)(1) by *Allis-Chalmers* Court), *enforcement denied*, 725 F.2d 1212 (9th Cir. 1984).

[32] *Allis-Chalmers*, 388 U.S. at 189 & n.25.

[33] 394 U.S. 423 (1969).

[34] *Id.* at 426–27. In *Scofield*, half of the production employees of the Wisconsin Motor Corporation received their wages on a piecework basis. The union that had organized the corporation's plant adopted a rule permitting each union member to produce as much each day as that member wanted. Nonetheless, under the rule the daily wages of each member could not exceed a certain rate. Wages for production that exceeded the ceiling were "retained by the company and paid out to the employee for days on which the production ceiling ha[d] not been reached because of machine breakdown or for some other reason." *Id.* at 424–25. The rule permitted the union to fine those members who demanded payment above the ceiling rate. A union member who failed to pay this fine, or a larger fine for repeated violations of the union rule, also risked expulsion from the union. *Id.* at 425.

[35] *Id.* at 431–32. The Court admitted, for example, that the imposition of union discipline here may manifest itself externally "in the [union] member's refusal to accept work offered by the employer." *Id.* at 436.

[36] *Id.* at 435.

[37] 409 U.S. 213 (1972).

[38] 412 U.S. 84 (1973).

[39] *Granite State*, 409 U.S. at 214.

members resigned from the union and subsequently returned to work.[40] In *Booster Lodge*, on the other hand, the union's constitution prohibited members from strikebreaking.[41] When several members resigned from the union and resumed employment during the strike, the union immediately fined them and claimed that its constitutional prohibition was enforceable regardless of the employees' resignations.[42]

Although recognizing that a union may lawfully regulate its own internal affairs, in both cases the Court emphasized that when a member resigns from the union, the union's power over the member ends. In other words, the enforcement of post-resignation fines was *external* to the union's legitimate interests.[43] As the Court in *Granite State* declared, "when there is a lawful dissolution of a union-member relation [that is, a lawful resignation], the union has no more control over the former member than it has over the man in the street."[44]

Although the Court in *Allis-Chalmers*, *Scofield*, *Granite State*, and *Booster Lodge* did not specifically address the issue of union constitutional provisions restricting resignations,[45] the NLRB has nonetheless drawn upon the language of these decisions in striking down restrictions on resignations. The NLRB has viewed these restrictions as "a unilateral reordering of the basic employee-union relationship that directly and fundamentally redraws the line between internal and external actions."[46] Because

---

[40] *Id.* A union member violating this rule was subject to a $2,000 fine. *Id.*

[41] *Booster Lodge*, 412 U.S. at 85–86. The constitution of the International Association of Machinists and Aerospace Workers did not explicitly prohibit resignations.

[42] *Id.* at 86–89.

[43] *See Granite State*, 409 U.S. at 217; *see also Booster Lodge*, 412 U.S. at 89–90 (applying the reasoning of *Granite State* to determine the enforceability against union resignees of the constitutional prohibition against strikebreaking).

[44] *Granite State*, 409 U.S. at 217. In response to these words, one prominent labor scholar has suggested that *Granite State* should be overruled, because it "obviously undermines the solidarity interest of *Allis-Chalmers*." Gould, *supra* note 1, at 92–93.

[45] In *Granite State*, for example, the Court explicitly reserved decision on the union's authority to restrict the resignations of its members: "We do not now decide to what extent the contractual relationship between union and member may curtail the freedom to resign." 409 U.S. at 217. Because of this reservation, some commentators have correctly predicted that both the NLRB and the courts would eventually confront the issue of union rules prohibiting resignations. *See, e.g.*, Wellington, *Union Fines and Workers' Rights*, 85 YALE L.J. 1022, 1044 (1976) ("The courts will undoubtedly be faced with [union constitutional provisions] barring resignations entirely, barring it during strikes, or permitting it only at certain inconvenient times."). Some commentators have even predicted how the NLRB is likely to resolve this issue. *See, e.g.*, *Conflict*, *supra* note 23, at 874 ("the results of the great majority of [NLRB] decisions after *Textile Workers* and *Booster Lodge* indicate that a majority of the NLRB may well be unwilling to uphold any union restrictions on resignations").

[46] *Neufeld Porsche-Audi*, 270 N.L.R.B. No. 209, 1983–84 NLRB Dec. (CCH) ¶ 16,436 (1984), at 28,096 n.16.

both *Granite State* and *Booster Lodge* prohibit union regulation of a former union member's post-resignation conduct, the NLRB considers a restriction on resignation as a way of circumventing this prohibition. As it observed in *Neufeld Porsche-Audi*: "By unilaterally extending an employee's membership obligation through restriction on resignations a union artificially expands the definition of internal action and can thus continue to regulate conduct over which it would otherwise have no control."[47]

C. *The* Scofield *test*

The third line of analysis involves the application of a test elaborated in *Scofield* to determine whether a union rule violates section 8(b)(1). According to this test, a union may enforce a rule if the rule (1) reflects a legitimate union interest; (2) impairs no congressional policy underlying the federal labor laws; and (3) is "reasonably enforced against union members who are free to leave the union and escape the rule."[48] Although the NLRB has readily admitted that a restriction on resignations serves a union's legitimate interest in maintaining strike solidarity,[49] it still has insisted that such a rule fails the *Scofield* test for two reasons. First, adhering to its absolutist interpretation of section 7 rights, the NLRB has insisted that the only statutory limitations on this right are contained in section 8(b)(2)[50] and in the second proviso to section 8(a)(3)[51] of the National Labor Rela-

---

[47] *Id.* at 28,096.

[48] Scofield v. NLRB, 394 U.S. 423, 430 (1969).

[49] *See* International Association of Machinists and Aerospace Workers, Local Lodge 1414 v. Neufeld Porsche-Audi, Inc., 270 N.L.R.B. No. 209, 1983–84 NLRB Dec. (CCH) ¶ 16,436 (1984), at 28,096; Machinists Local 1327 (Dalmo Victor), 263 N.L.R.B. 984, 985 (1982), *enforcement denied*, 725 F.2d 1212 (9th Cir. 1984).

[50] Section 8(b)(2) provides that:

> It shall be an unfair labor practice for a labor organization or its agents . . . (2) to cause or attempt to cause an employer to discriminate against an employee in violation of Section [8](a)(3) or to discriminate against an employee with respect to whom membership in such organization has been denied or terminated on some ground other than his failure to tender the periodic dues and the initiation fees uniformly required as a condition of acquiring or retaining membership.

29 U.S.C. § 158(b)(2) (1982).

[51] Section 8(a)(3) protects the right of a union to require union membership as a condition of employment. The text of section 8(a)(3) reads:

> It shall be an unfair labor practice for an employer . . . (3) by discrimination in regard to hire or tenure of employment to encourage or discourage membership in any labor organization: *Provided*, That nothing in this Act, or in any

tions Act.[52] Because even these provisions do not compel full union membership, the NLRB has argued that when a union restricts a member's resignation, it directly impairs his or her section 7 rights, thereby violating the congressional policy that union membership be fully voluntary.[53] Second, the NLRB has concluded that a restriction on resignations is clearly invalid under the third part of the *Scofield* test requiring that an employee be "free to leave the union and escape the rule."[54]

The NLRB's current approach is only one possible interpretation of the Act. The coexistence of section 7 employee rights and section 8(b)(1)(A) union rulemaking powers presents an inherent possibility of conflict. Because Congress did not explicitly authorize reading the latter as a possible limit on the former, the NLRB feels bound to give primacy to section 7 when addressing this conflict.[55] This narrow, technical approach could frustrate broader policies underlying the Act by failing to consider the effects of employee action on the union's legitimate

---

> 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 on or after the thirtieth day following the beginning of such employment or the effective date of such agreement, whichever is the later, . . . *Provided further*, That no employer shall justify any discrimination against an employee for nonmembership in a labor organization . . . 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).

[52] International Association of Machinists and Aerospace Workers, Local Lodge 1414 v. Neufeld Porsche-Audi, Inc., 270 N.L.R.B. No. 209, 1983–84 NLRB Dec. (CCH) ¶ 16,436 (1984), at 28,096. In the concurring opinion in *Dalmo Victor*, subsequently adopted by the NLRB majority in *Neufeld Porsche-Audi*, Chairman Van de Water and Member Hunter did concede that in certain instances nonstatutory property rights may outweigh the individual worker's section 7 rights. Machinists Local 1327 (Dalmo Victor), 263 N.L.R.B. 984, 991 n.43 (1982) (Van de Water & Hunter, concurring), *enforcement denied*, 725 F.2d 1212 (9th Cir. 1984).

[53] *See* International Association of Machinists and Aerospace Workers, Local Lodge 1414 v. Neufeld Porsche-Audi, Inc., 270 N.L.R.B. No. 209, 1983–84 NLRB Dec. (CCH) ¶ 16,436 (1984), at 28,096.

[54] *Id.* at 28,097; Pattern Makers' League of North America v. NLRB, 265 N.L.R.B. 1332, 1333 (1982), *enforced*, 724 F.2d 57 (7th Cir. 1983), *cert. granted*, 105 S. Ct. 79 (1984) (No. 83-1894); Machinists Local 1327 (Dalmo Victor), 263 N.L.R.B. 984, 986 (1982), *enforcement denied*, 725 F.2d 1212 (9th Cir. 1984).

[55] According to the concurring opinion in *Dalmo Victor*, now expressly adopted by the NLRB, "by equating institutional 'interests' with statutory 'rights' and utilizing the existence of 'conflict' between disputants to justify reduction of the Act's protections, we respectfully submit that our colleagues are engaging in legislating rather than interpreting our Act's intent and objectives. This they may not do." 263 N.L.R.B. at 991 (Van de Water & Hunter, concurring).

interests, no matter how serious those effects may be.[56] Alternative approaches are available.

II. Two Views of Reasonableness: The NLRB and the Ninth Circuit in *Dalmo Victor*

One such approach would involve a balancing of the importance of the union's interest in restricting resignations with the degree of infringement of the employee's rights. The NLRB itself formerly used such an analysis. In its decision in *Dalmo Victor*, a plurality[57] of the NLRB concluded that a union constitutional provision prohibiting resignations during a strike or within fourteen days preceding the strike's commencement unreasonably interfered with the union members' section 7 rights and that an attempt to collect fines under such a provision violated section 8(b)(1)(A).[58] Nonetheless, in sharp contrast with its concurring opinion and with the NLRB majority in the subsequent *Neufeld Porsche-Audi* case, the plurality, in reaching its conclusion, did concede that the section 7 right of an employee to refrain from collective activity could be balanced with a union's interest in maintaining solidarity among its members.[59] After performing some balancing acrobatics, the plurality concluded that the contested constitutional provision was unreasonable because it allowed a union member to resign only during nonstrike periods.[60] Recognizing that some restrictions on resignations were permissible, the plurality then constructed its own rule, which would limit a union member's right to resign

---

[56] The notion that the construction of the Act by the NLRB and the courts may violate the Act's underlying spirit is not novel. *See, e.g.*, Weiler, *Promises to Keep: Securing Workers' Rights to Self-Organization under the NLRA*, 96 HARV. L. REV. 1769, 1777 (1983) (arguing that "the current certification procedure does not effectively insulate employees from the kinds of coercive antiunion employee tactics that the [Act] was supposed to eliminate").

[57] Of the five members who sat on the NLRB, only two constituted the plurality in *Dalmo Victor*. Two other members—Van de Water and Hunter—agreed with the plurality in its decision to invalidate the contested union constitutional provision but nonetheless rejected the plurality's formulation of a 30-day standard. *Dalmo Victor*, 263 N.L.R.B. at 987 (Van de Water & Hunter, concurring). The fifth member of the NLRB—Jenkins—dissented from the plurality's decision. *Dalmo Victor*, 263 N.L.R.B. at 993 (Jenkins, dissenting). On appeal, the Ninth Circuit questioned whether it was appropriate to describe the plurality's 30-day standard as a rule, because only two members had supported it and three had opposed it. Machinists Local 1327 v. NLRB (Dalmo Victor), 725 F.2d 1212, 1215 n.3 (9th Cir. 1984).

[58] *Dalmo Victor*, 263 N.L.R.B. at 984.

[59] *Id.* at 986.

[60] *Id.*

for a period not more than thirty days after the member had tendered his resignation, regardless of whether a strike was in progress.[61]

The Court of Appeals for the Ninth Circuit struck down the NLRB plurality's thirty-day rule[62] and upheld the original union constitutional provision.[63] The court emphasized that the challenged provision did not absolutely prohibit all resignations from the union, but only placed "some obstacles in the way of resignation."[64] Like the plurality below, the court also emphasized the economic superiority of a collective agency working on behalf of the individual worker and pointed out that each worker did not have an absolute right to order his relations with management.[65]

The truly distinctive feature of the court's decision, however, is its formulation of the *Scofield* test. According to the Ninth Circuit, *Scofield* permits a union to enforce a properly adopted rule that (1) reflects a legitimate union interest; (2) impairs no policy that Congress has embedded in the labor laws; and (3) is reasonably enforced against union members.[66] By failing to require that a member be "free to leave the union and escape the rule" in part three, the court has made the *Scofield* test substantially easier to pass.

Because the test is unambiguously stated in *Scofield*,[67] the Ninth Circuit cannot defend its interpretation of the test by arguing that the Supreme Court's test is unclear. Instead, the Ninth Circuit apparently decided to read the third part of the test in a way that would allow the union's interests to be considered. By not insisting that a union member be "free to leave the union and escape the rule," the court could then give more

---

[61] *Id.* at 987. The NLRB subsequently rejected the thirty-day rule in *Neufeld Porsche-Audi.* International Association of Machinists and Aerospace Workers, Local Lodge 1414 v. Neufeld Porsche-Audi, Inc., 270 N.L.R.B. No. 209, 1983–84 NLRB Dec. (CCH) ¶ 16,436 (1984), at 28,094. For a general discussion of the NLRB's decision in *Dalmo Victor,* see Note, *Union Security and Union Members' Freedom to Resign: The NLRB's Thirty-day Rule in* Dalmo Victor, 14 TEX. TECH. L. REV. 593 (1983).

[62] Machinists Local 1327 v. NLRB (Dalmo Victor), 725 F.2d 1212, 1215 (9th Cir. 1984).

[63] *Id.* at 1214, 1218.

[64] *Id.* at 1217. Minimizing the restrictive effects of the constitutional provision is misleading because when "a member contemplating both strikebreaking and resignation is threatened with union sanctions after he resigns, it is unlikely he will utilize the resignation escape route contemplated by *Granite State.*" Gould, *supra* note 1, at 104.

[65] Machinists Local 1327 v. NLRB (Dalmo Victor), 725 F.2d 1212, 1215–16 (9th Cir. 1984).

[66] *Id.* at 1216.

[67] Scofield v. NLRB, 394 U.S. 423, 430 (1969).

weight to the argument that "[t]here is little point in taking a strike vote if the people who disagree with the outcome are free to resign anytime and escape its effects."[68] In this way, a restriction on resignations became a reasonable means of preserving union interests.[69] The Supreme Court in *Scofield* also conceded that the union rule at issue, which imposed a production ceiling, served the union's interest and could validly be enforced against union members.[70] The Court nonetheless took the position that, in spite of the rule, a union member "may leave the union and obtain whatever benefits in job advancement and extra pay [that] may result from extra work, [while] at the same time enjoying . . . the job security which compliance with the union rule by union members tends to promote."[71] The reasonableness of the rule could not negate the union member's right to be free to escape the rule. As these two examples illustrate, different formulations of the *Scofield* test have led to radically different judicial determinations.

After applying its revised version of the *Scofield* test to the challenged union rule, the Ninth Circuit concluded that the rule served a legitimate union interest, impaired no national labor policy, and was a reasonable means of enforcing union discipline.[72] In reaching this conclusion, the court explicitly rejected the NLRB plurality's attempt to balance the employee's right to resign and the union's interest in enforcing its own disciplinary rules.[73] Pointing out that "*both* the employee's right and the union's interest are policies that have been 'embedded' in the labor laws for over 35 years," the court argued that neither can outweigh the other under *Scofield*, and that "[t]hey must—and do—coexist."[74]

By rejecting the NLRB's balancing approach and by insisting that the section 7 and section 8(b)(1)(A) rights do not conflict

---

[68] Machinists Local 1327 v. NLRB (Dalmo Victor), 725 F.2d 1212, 1217 (9th Cir. 1984). *See also* Justice Blackmun's dissent in NLRB v. Granite State Joint Board, Textile Workers Union, 409 U.S. 213, 221 (1972) (Blackmun, J., dissenting) (arguing to uphold union rule against strikebreaking because members who resigned had participated in the strike vote).

[69] Machinists Local 1327 v. NLRB (Dalmo Victor), 725 F.2d 1212, 1218 (9th Cir. 1984).

[70] Scofield v. NLRB, 394 U.S. 423, 432–36 (1969).

[71] *Id.* at 435.

[72] Machinists Local 1327 v. NLRB (Dalmo Victor), 725 F.2d 1212, 1217–18 (9th Cir. 1984).

[73] *Id.* at 1216.

[74] *Id.* at 1217.

but peacefully coexist, the Ninth Circuit embraces an untenable view of this nation's labor laws. Contrary to the court's view, an employee's right to refrain from concerted activities and a union's right to create and to enforce its own legitimate regulations often cannot peacefully coexist in any meaningful sense. For example, when a union tells one of its members that he or she can resign his or her membership only during a nonstrike period, the union is certainly impairing the employee's section 7 right. Similarly, when a union member resigns and crosses a picket line in violation of a union rule, the employee challenges the union's section 8(b)(1)(A) right to regulate the acquisition and retention of its membership.[75]

This inherent conflict suggests that in determining whether the section 7 rights or the 8(b)(1)(A) rights should prevail in a particular situation, the decisionmaker's most natural role is to balance the competing rights and interests of both the union and the individual employee. The attempts of the current NLRB and the Ninth Circuit to avoid this balance result only in awkward straining. The NLRB suggests that whenever there is a conflict, the employee's section 7 right is supreme. This approach seriously hampers the union's ability to pursue its interests. The Ninth Circuit suggests that there is no conflict between sections 7 and 8(b)(1)(A), treats both as absolute, but gives the union side an artificial advantage: a union restriction on resignations will be reasonable and valid if it passes the Circuit's relaxed version of the *Scofield* test.[76] This view is no more faithful to the Act than the NLRB's position.

III. Union Security and the Right to Resign: A New View of Reasonableness

This Part offers an analytic framework with which to evaluate the reasonableness of union rules restricting resignations. The foundations of this framework can be summarized as follows. First, the section 7 right of an employee to refrain from union activities is not absolute.[77] In certain instances a union may

[75] *See* NLRB v. Allis-Chalmers Mfg. Co., 388 U.S. 175, 176–78 (1967), *reh'g denied*, 389 U.S. 892 (1967).

[76] Machinists Local 1327 v. NLRB (Dalmo Victor), 725 F.2d 1212, 1218 (9th Cir. 1984).

[77] *See* International Association of Machinists and Aerospace Workers, Local Lodge 1414 v. Neufeld Porsche-Audi, Inc., 270 N.L.R.B. No. 209, 1983–84 NLRB Dec. (CCH) ¶ 16,436 (1984), at 28,099–100 (Zimmerman, dissenting).

need to restrict the resignation rights of its members in order to represent effectively the entire membership during collective bargaining. Because the underlying policy expressed in the Act is "to eliminate the causes of certain substantial obstructions to the free flow of commerce . . . by encouraging the practice and procedure of collective bargaining"[78] and because the Act declares that "the advancement of the general welfare of the Nation can most satisfactorily be secured by the settlement of issues between employers and employees through the process of conference and collective bargaining,"[79] the union's interest should not always be sacrificed to enforcement of section 7 rights.

Second, union security clauses offer a union a guarantee of continued financial support from those employees whom it represents as their exclusive bargaining agent. Under most security clauses, even if an employee resigns from the union during a strike and subsequently returns to work, he or she must still continue to contribute union dues.[80] Often a portion of these dues is directly deposited in a union strike fund.[81] Thus, if a security clause is in force, there is less reason to restrict the employee's resignation.

Third, it seems likely that it is generally the weaker union, or the union in its first years after certification, that is unable to obtain a security clause in its collective bargaining agreement.[82] Allowing unrestricted resignations might leave such a union with no real bargaining ability.

---

[78] 29 U.S.C. § 151 (1982).

[79] *Id.* § 171.

[80] *See* Marlin Rockwell Corporation, 114 N.L.R.B. 553, 561 (1955) (an employee may resign from a union, protected by a security clause in its collective bargaining agreement, only if the employee continues to tender union dues); *see also Restrictions, supra* note 23, at 410 (pointing out that the NLRB in *Marlin Rockwell* recognized an employee's obligation to continue to pay union dues after his resignation from the union).

[81] The United Auto Workers Union, for example, diverts 30% of its union security payments (that is, members' dues and agency shop fees) to its strike insurance fund. Cantor, *Uses and Abuses of the Agency Shop*, 59 NOTRE DAME LAW. 61, 93 n.154 (1983) (citing Kolinske v. Lubbers, 516 F. Supp. 1171, 1173 (D.D.C. 1981), *rev'd*, 712 F.2d 471, 473 (D.C. Cir. 1983)).

[82] Because a security clause helps to protect the union's financial stability, *see infra* notes 95–125 and accompanying text, it is unlikely that a union would not bargain for such a clause. Security clauses are accepted by employers in the great majority of collective bargaining agreements, *see infra* text accompanying note 131. The absence of such a clause, therefore, seems to indicate some weakness in the union's bargaining position.

A. *The Limited Section 7 Right*

Although the legislative history of section 7 is susceptible to different interpretations,[83] Congress probably intended that the right to refrain from union activities would include the right to resign from the union.[84] This intention, however, does not necessarily mean that these rights are absolute. The current NLRB, nevertheless, has interpreted section 7 broadly: "[T]he only statutory limitation on [the] section 7 rights is contained in section 8(b)(2) and in the second proviso to section 8(a)(3) . . . ."[85] This interpretation is a somewhat remarkable position for the NLRB to adopt, especially because section 7 itself provides at least one exception to the rights that it guarantees. This exception permits the inclusion of a union security clause in a collective bargaining agreement.[86]

Similarly, in allowing unions to prescribe their own internal rules, Congress must have anticipated that this section 8(b)(1)(A) right might occasionally conflict with the individual worker's section 7 right to refrain from union activities. Unless Congress intended section 8(b)(1)(A) to be an empty provision, the individual employee's rights cannot always be supreme. The Supreme Court has recognized that section 8(b)(1)(A) can

---

[83] *See, e.g.*, Brief for Respondent at 24–29, Pattern Makers' League of North America v. NLRB, No. 83-1894 (U.S. filed May 18, 1984) (arguing that Congress's failure to enact a specific statutory provision granting the right to resign does not indicate that section 7 does not encompass this right), *cert. granted* 105 S. Ct. 79 (1984), *argued* Feb. 27, 1985, 53 U.S.L.W. 3632 (U.S. Feb. 12, 1985), *reargument ordered*, 53 U.S.L.W. 3686 (U.S. Mar. 26, 1985); Brief for the Petitioners at 24–32, Pattern Makers' League of North America v. NLRB, No. 83-1894 (U.S. filed May 18, 1984) (arguing that it is unclear whether "the addition of the 'right to refrain' to section 7 is intended to grant union members a right to resign at will in contravention of the union's rules limiting resignations"), *cert. granted* 105 S. Ct. 79 (1984), *argued* Feb. 27, 1985, 53 U.S.L.W. 3632 (U.S. Feb. 12, 1985), *reargument ordered*, 53 U.S.L.W. 3686 (U.S. Mar. 26, 1985).

[84] In *Pattern Makers*, for example, the Seventh Circuit declared that "[a]n employee's right to resign not only is guaranteed by Section 7, but also is supported by an employee's own strong interests." Pattern Makers' League of North America v. NLRB, 724 F.2d 57, 60 (7th Cir. 1983), *cert. granted* 105 S. Ct. 79 (1984) (No. 83-1894), *argued* Feb. 27, 1985, 53 U.S.L.W. 3632 (U.S. Feb. 12, 1985), *reargument ordered*, 53 U.S.L.W. 3686 (U.S. Mar. 26, 1985).

[85] International Association of Machinists and Aerospace Workers, Local Lodge 1414 v. Neufeld Porsche-Audi, Inc., 270 N.L.R.B. No. 209, 1983–84 NLRB Dec. (CCH) ¶ 16,436 (1984), at 28,096.

[86] Section 7 allows an employee to refrain from union activities "*except* to the extent that such right may be affected by an agreement requiring membership in an organization as a condition of employment as authorized in Section 8(a)(3)." 29 U.S.C. § 157 (1982) (emphasis added). In this way, both § 7 and § 8(a)(3) authorize the inclusion of union security clauses in collective bargaining agreements.

eclipse section 7. In *Allis-Chalmers*, for example, the punished employee sought to abstain from a strike, but the Court permitted the union to exercise its authority under section 8(b)(1)(A) and to fine the employee for his strikebreaking activity.[87]

Thus, the courts have acknowledged that competing statutory rights may qualify the section 7 right. Determining the scope of such qualifications is the task that courts must undertake.

At first glance, it may appear that the rules of contract law could resolve many of the conflicts arising between a union and its members.[88] In *Allis-Chalmers*, for example, the Court defended the enforcement of union discipline by pointing to the pervasiveness of the contract theory of the union-member relationship at the time of the Taft-Hartley amendments.[89] The Court's rationale in both *Granite State* and *Booster Lodge* also suggests that contract analysis is at least relevant in defining the union-member relationship.[90] In describing the impact of *Sco-*

---

[87] *See supra* text accompanying notes 23–32.

[88] *See, e.g., Restrictions, supra* note 23, at 405 ("[t]he decisions [of the 1972-73 Term] . . . indicate that contractual analysis is relevant, at least to the extent that only express as opposed to implied waivers of statutory rights may be effective"); Brief for Petitioners at 34–38, Pattern Makers' League of North America v. NLRB, No. 83-1894 (U.S. filed May 18, 1984) (arguing that a contractual restriction on a union member's right to resign should be judged under the common law of associations), *cert. granted* 105 S. Ct. 79 (1984), *argued* Feb. 27, 1985, 53 U.S.L.W. 3632 (U.S. Feb. 12, 1985), *reargument ordered*, 53 U.S.L.W. 3686 (U.S. Mar. 26, 1985).

Some commentators have argued that union constitutions—the instruments defining the union-member relationship—are in fact contracts of adhesion, because the individual employee is generally unable to bargain with the union over the terms of the constitution. Consequently, they suggest that it is inappropriate to judge the validity of a union rule based simply on the union-member contractual relationship. The parties' additional rights under the Act must be considered. *See* Gould, *supra* note 1, at 97; *see also Restrictions, supra* note 23, at 413; Brief for Respondent at 19 n.9, Pattern Makers' League of North America v. NLRB, No. 83-1894 (U.S. filed May 18, 1984), *cert. granted* 105 S. Ct. 79 (1984), *argued* Feb. 27, 1985, 53 U.S.L.W. 3632 (U.S. Feb. 12, 1985), *reargument ordered*, 53 U.S.L.W. 3686 (U.S. Mar. 26, 1985).

In its recent decision in United Association of Journeymen & Apprentices of the Plumbing & Pipefitting Industry v. Local 334, 452 U.S. 615 (1981), the Supreme Court declared that a union constitution is a contract between the local union and its international parent for the purposes of filing suit under section 301(a) of the Labor Management Relations Act. *Id.* Section 301(a) allows the federal district courts to resolve contract disputes between labor organizations. 29 U.S.C. § 301(a) (1982). For a critical analysis of the Court's decision in *United Association*, see Note, *Bringing Union Constitutions Within the Sweep of Section 301:* United Association of Journeymen v. Local 334, 24 B.C.L. REV. 145, 150–51 (1982).

[89] *Allis-Chalmers*, 388 U.S. at 192.

[90] Although finding against the unions in both cases, the Court emphasized that there were no provisions in the union-member contracts restricting resignations or indicating employee knowledge of, or consent to, limitations on the right to resign. The Court admitted that the existence of such clauses in the union-member contract would have made the cases more difficult to decide. *Booster Lodge*, 412 U.S. at 88; *Granite State*, 409 U.S. at 216.

field on union discipline, for example, the Court in *Granite State* acknowledged that the union-member contract can empower the union to restrict some of the activities of its members.[91] Although conceding that boilerplate provisions in a union constitution are poor indicia of a member's obligations to his union, Justice Blackmun's dissenting opinion in *Granite State* also implied that the disgruntled employees were *contractually* obliged to adhere to the union's rules, because they had participated in the strike vote and had joined in the decision to impose fines on strikebreakers.[92]

Nonetheless, invoking contract law cannot adequately resolve a union-member conflict touching the fundamental protections under the Act. In *Scofield*, for example, the Court emphasized that a union rule is unenforceable if it "invades or frustrates an overriding policy of the labor laws."[93] Thus, even if a union member has explicitly consented to a union rule, this rule is invalid if it violates one of the fundamental protections of the Act.

But if union interests are to be evaluated against the purposes of the labor laws, employee interests must also be so evaluated. Congress did not intend to limit the unions' "powers necessary to the discharge of their role as exclusive statutory bargaining agents by impairing the usefulness of labor's cherished strike weapon."[94] In evaluating union restrictions on resignation, the decisionmaker should be able to weigh the possibility that exercise of the employee's right to resign may threaten these necessary powers to such an extent that the restriction becomes reasonable.

B. *Union Security and the Free Rider*

Sections 7 and 8(a)(3) of the Act allow unions to incorporate security clauses into their collective bargaining agreements. Al-

---

[91] *See Granite State*, 409 U.S. at 217.

[92] *Id.* at 220 (Blackmun, J., dissenting). In the recent oral arguments before the Supreme Court in *Pattern Makers*, Justice Brennan asked whether union members were aware of the resignation restriction when they joined the union. 53 U.S.L.W. 3646, 3647 (U.S. Mar. 12, 1985). By asking this question and indicating that "there was nothing in the record to show [that the members] were affirmatively informed of [the restriction]," *id.*, Brennan seems to be suggesting that the element of consent, or contract analysis, is relevant to the resolution of the issue presented by the case.

[93] *Scofield*, 394 U.S. at 429.

[94] NLRB v. Allis-Chalmers Mfg. Co., 388 U.S. 175, 183 (1967), *reh'g denied*, 389 U.S. 892 (1967).

though Congress has prohibited the closed shop,[95] the most coercive form of union security, it has still permitted unions to negotiate the union shop,[96] the agency shop,[97] the maintenance of membership agreement,[98] and the representation fee agreement.[99] These forms of union security all vary in the degree to which they compel employees to participate in the activities of the union.

In defense of the union security provision, some commentators have argued that there is a "need for coercion implicit in attempts to provide collective goods to large groups."[100] The labor union engages in collective bargaining for a large number

---

[95] A closed shop agreement requires that "[a]n individual . . . be a member of the union in order to be eligible for hire and must retain this membership as a condition of continued employment with the contracting employer." T. HAGGARD, *supra* note 8, at 4. Concerned by the broad powers granted to a union under a closed shop agreement, Congress outlawed these agreements by enacting the Taft-Hartley Amendments. Labor-Management Relations Act, ch. 114, § 8(a)(3), 61 Stat. 140–41 (1947) (codified at 29 U.S.C. § 158(a)(3)).

> Why should a union be able to say to an employee "If you do not join this union we will see that you cannot work in this plant."? They have said to them "Sooner or later we are going to organize this plant with a closed shop and you will be out." It seems to me perfectly clear that this is a reprehensible practice.

93 CONG. REC. 4142 (1947) (statement of Sen. Taft (R-Ohio)); *see also* T. HAGGARD, *supra* note 8, at 36; A. COX, D. BOK & R. GORMAN, *supra* note 6, at 84.

[96] Under a union shop agreement, "[a]n individual who is not a member of the union may be hired but within a specified time after hire must become and remain a member as a condition of continued employment with the contracting employer." T. HAGGARD, *supra* note 8, at 4. Some union shop agreements do not require workers who were hired before the date of the collective bargaining agreement to become union members. BUREAU OF LABOR STATISTICS, U.S. DEP'T OF LABOR, BULL. 1425-21, MAJOR COLLECTIVE BARGAINING AGREEMENTS: UNION SECURITY AND DUES CHECKOFF PROVISIONS 5 (1982) [hereinafter cited as BLS BULLETIN].

[97] Under an agency shop agreement, "[a]n individual who is not a member of the union may be hired and retained in employment without the necessity of becoming a member of the union, but he is required to tender the equivalent of initiation fees and periodic dues to the union as a condition of continuing employment with the contracting employer." T. HAGGARD, *supra* note 8, at 4; *see* BLS BULLETIN, *supra* note 96, at 10–12.

[98] A maintenance of membership agreement requires "[a]n employee who is a member of the union at the beginning of the contract or who becomes a member during the term of the contract . . . to remain a member until the termination of the contract." T. HAGGARD, *supra* note 8, at 6. Maintenance of membership agreements are a relatively weak form of union security, because they do not compel nonmembers to contribute union dues and fees. BLS BULLETIN, *supra* note 96, at 8; *see infra* note 150 and accompanying text.

[99] Under a representation fee agreement, "[a]n individual who is not a member of the union may be hired and retained in employment without the necessity of becoming a member of the union, but as a condition of employment he is required to tender to the union his pro rata share of the costs incurred by the union in performing its statutory function as the exclusive bargaining representative." T. HAGGARD, *supra* note 8, at 4–5.

[100] M. OLSEN, THE LOGIC OF COLLECTIVE ACTION 71 (1965).

of employees. Proponents of the security clause argue that these clauses must be coercive if they are to eliminate the "free rider," the employee who reaps the benefits of a collective bargaining agreement without supporting the union.[101] Thus, while recognizing the coercive nature of the security clause, these proponents claim that to denounce union security as a restriction of individual freedom requires one also to denounce all coercion used to support the provision of collective services, including the provision of services by the government.[102]

Opponents of union security emphasize that security clauses require employees to pay for services that the employees would rather forego.[103] They claim that such a requirement is an impermissible burden on the individual worker's fundamental constitutional rights of association, speech, and even religion.[104] Opponents of union security also challenge the very assumption underlying the free rider argument. In their view, it is wrong to assume that unions actually benefit all whom they represent, particularly in light of those union-represented workers who are now unemployed allegedly as a result of the excessive demands of union officials.[105]

---

[101] *See id.* at 94–96. For some commentators, eliminating the free rider "appeals to the canons of fairness." Eissinger, *The Right-to-Work Imbroglio*, 51 N.D.L. REV. 571, 584 (1975). This appeal is based upon the following set of facts:

> A majority of workers through the democratic process have selected a bargaining agent. This agent incurs expenses in carrying out his statutory obligations which may include the salary of the employees' negotiator, the costs of negotiating and administering the agreement and the cost of operating the grievance and arbitration machinery.

Id. at 584. *But see* Freed, Polsby & Spitzer, *Unions, Fairness, and the Conundrums of Collective Choice*, 56 S. CAL. L. REV. 461, 468–69 n.15 (1983) (". . . any attempt to generate a broad principle of *fairness* and apply it to labor union conduct produces inconsistent and otherwise unacceptable results") (emphasis added).

[102] In his discussion of the agency shop, Professor Cantor makes the same point:

> So long as the organization does in fact perform a useful function for the fees payors, and so long as the organization is legally bound to use the funds to promote the related functions and goals of the organization, then the disgruntled fee payor cannot complain any more than the taxpayer whose funds are used by the government for programs ideologically offensive to the taxpayer.

Cantor, *supra* note 81, at 70–71.

[103] For a discussion of the standard arguments made against union security agreements, see T. HAGGARD, *supra* note 8, at 278–84.

[104] *Id.* at 278–79.

[105] *See* National Right to Work Committee, Exclusive Representation—The Foundation of Compulsory Unionism 3–5 (Issue Briefing Paper Sept. 2, 1980) (available from National Right to Work Committee, 8001 Braddock Road, Springfield, Va. 22160); *see also* Bailey & Heldman, *The Right to Work Imbroglio: Another View*, 53 N.D.L. REV. 163, 166 (1976) ("the assumption that the employees in a collective bargaining unit enjoy even a net economic benefit as a result of collective bargaining ignores the wage losses and other tangible and intangible suffering which accompany long strikes").

Unpersuaded by these right-to-work arguments, Congress has given its support to the proponents of the union security clause in sections 7 and 8(a)(3) of the Act, which authorize the adoption of security clauses. The courts have consistently pointed out that "Congress' essential justification for authorizing the union shop [and other types of union security] was the desire to eliminate free riders . . . ."[106]

Although section 14(b) of the Act also allows the individual states to prohibit the adoption of security clauses,[107] these clauses continue to pervade those collective bargaining agreements made in jurisdictions outside the right-to-work states. The important characteristics of the two most common forms of union security—the union shop and the agency shop—are discussed below.

The union shop is the most common type of union security clause[108] and appears to be the most coercive in its requirements.

---

[106] Ellis v. Brotherhood of Railway, Airline and Steamship Clerks, 104 S. Ct. 1883, 1892 (1984), *aff'g in part and rev'g in part* 685 F.2d 1065 (9th Cir. 1982).

[107] Section 14(b) provides: "Nothing in this subchapter 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). Twenty-one states have now adopted right-to-work laws. ALA. CODE §§ 25-7-30 to -36 (1975); ARIZ. CONST. amend. 34; ARIZ. REV. STAT. ANN. § 23-1301 to -1307 (1983); ARK. CONST. amend. 34; ARK. STAT. ANN. §§ 81-201 to -205 (1976); FLA. CONST. art. I, § 6; FLA. STAT. ANN. § 447.09(11) (West 1981); GA. CODE ANN. §§ 34-6-23 to -28 (1982); Idaho Act of Jan. 31, 1985, H.B.2, 3 EMPL. REL. WKLY. (BNA) 134; IOWA CODE ANN. §§ 731.1–.8 (West 1979); KAN. CONST. art. 15, § 12; LA. REV. STAT. ANN. §§ 23:981–87 (West Supp. 1985); MISS. CONST. art. 7, § 198-A; MISS. CODE ANN. § 71-1-47 (1972); NEB. CONST. art. 15, §§ 13–15; NEB. REV. STAT. § 48-217 (1978); NEV. REV. STAT. §§ 613.230–.300 (1979); N.D. CENT. CODE §§ 34-01-14, 34-08-02 (1980); N.C. GEN. STAT. §§ 95-78-.83 (1981); S.C. CODE ANN. §§ 41-7-10 to 7-90 (1976); S.D. CONST. art. VI, § 2; S.D. CODIFIED LAWS ANN. §§ 60-8-3 to -8 (1978); TENN. CODE ANN. §§ 50-1-201 to -204 (1983); TEX. REV. CIV. STAT. ANN. arts. 5154a, 5154g, 5207a (Vernon 1971); UTAH CODE ANN. §§ 34-34-1 to -17 (1974); VA. CODE §§ 40.1-58 to -69 (1981); WYO. STAT. §§ 27-7-108 to -115 (1977). All of these state laws and constitutional provisions prohibit employers from requiring union membership as a condition of employment. Some of these laws also prohibit a union from requiring the payment of agency or service fees. *E.g.*, IOWA CODE ANN. §§ 731.4–.5 (West 1979); VA. CODE § 40.1-62 (1981); WYO. STAT. § 27-7-111 (1977); *see also* Note, *Union Security Agreements in the Public Sector Since* Abood, 33 S.C.L. REV. 521, 529 (1982).

This Comment argues that the NLRB and the courts should adopt a rule that would prohibit unions protected by valid security clauses from restricting the resignations of their members. In those states with right-to-work laws, such a rule would have little relevance, because the enforcement of a security clause by a union would be illegal in the first instance. Determining the extent to which these state laws may be preempted by the NLRB's application of federal law is beyond the scope of this Comment.

[108] In a recent study conducted by the Bureau of Labor Statistics of the U.S. Department of Labor, 72% of all the collective bargaining agreements surveyed by the study contained a union shop clause. BLS BULLETIN, *supra* note 96, at 5.

A typical union shop agreement, for example, would not require union membership as an immediate condition of employment, but would insist that after a specified time the recently hired worker become and remain a union member.[109] These agreements often permit a thirty-day "grace period" after which employees must join the union.[110]

A pure agency shop agreement, on the other hand, does not require workers to join the union but insists that all workers—both member and nonmember—pay for the union's services.[111] These payments typically take the form of union dues, although they may be collected from the nonmember as his pro rata share of the union's collective bargaining expenses.[112]

The differences between the requirements of membership under a union shop agreement and those requirements of nonmembership under an agency shop agreement are not as significant as they may at first appear. In fact, although these differences are not negligible, collective bargaining agreements requiring union shops today effectively require only agency shops.[113] Because the courts have declared that "membership" in a labor organization includes only the duty to pay dues and initiation fees,[114] an employee bound by a union shop agreement is not required to become a formal member of the union by taking an

---

[109] The following clause establishes a typical union shop:

> Employees who are now members of the union shall, as a condition of employment, remain members of the union. All other employees within the bargaining unit and all new employees employed within the bargaining unit shall, as a condition of employment, become members on or after 30 calendar days of the execution of this agreement or their date of employment, whichever is later.

Id at 6.

[110] *Id.* A provision that shortens this "grace period" to less than thirty days is unenforceable. 29 U.S.C. § 158(a)(3) (1982). *See also* Zipp, *Rights and Responsibilities of Parties to a Union-Security Agreement*, 33 LAB. L.J. 203, 207 (1982).

[111] *See* BLS BULLETIN, *supra* note 96, at 10.

[112] In these situations, an agency shop agreement and a representation fee agreement are indistinguishable. Although a nonmember's pro rata share of the union's representation costs may equal monthly union dues, on some occasions the service fee is actually less than these dues. Eissinger, *supra* note 101, at 587.

[113] Gould, *supra* note 1, at 78; Cantor, *supra* note 81, at 61 n.2; *see also* Haggard, *Right-to Work Laws in the Southern States*, 59 N.C.L. REV. 29, 33 n.20 (1980) (arguing that state right-to-work laws may prohibit agency shop fees "because federal law prohibits anything more stringent [than the agency shop]"). *But see* Eissinger, *supra* note 101, at 579 ("There are still very important and viable distinctions between the union shop allowed under the federal law and the agency shop allowed under a state right-to-work statute.").

[114] *See* NLRB v. General Motors Corp., 373 U.S. 734, 742 (1963); United Stanford Employees, Local 680, Service Employees International Union v. NLRB, 601 F.2d 980 (9th Cir. 1979); NLRB v. Hershey Foods Corp., 513 F.2d 1083, 1087 (9th Cir. 1975).

oath of allegiance. Instead, his *entire* obligation to the union is satisfied by paying the equivalent of union dues and fees. As the Supreme Court announced in *NLRB v. General Motors Corp.*,[115] an employee's membership obligation to a union under a union shop agreement has been "whittled down to its financial core."[116]

Congress and the courts have imposed similar financial obligations on employees who have chosen not to join the union in an agency shop. This union is permitted to exact "regular" and "periodic" fees from nonmembers[117] in order to spread the costs of collective representation among everyone within the bargaining unit. In *Abood v. Detroit Board of Education*,[118] however, the Supreme Court also announced that the use of agency fees for "ideological activity unrelated to collective bargaining"[119] is impermissible, if the employees in the agency shop object to the use of the fees for these purposes.[120]

More recently, the Court of Appeals for the District of Columbia concluded that employees in an agency shop can be required to contribute to a strike insurance fund, even though the money in this fund is available only to those participating in the strike.[121] Although it is unclear whether a union may lawfully compel nonmembers in an agency shop to pay an irregular "strike assessment,"[122] the D.C. Circuit ruled that a union may divert regular agency shop fees contributed by nonmembers to a strike insurance fund. If we characterize strikes "as natural,

---

[115] 373 U.S. 734 (1963).

[116] *Id.* at 742.

[117] *See* International Harvester Co., 95 N.L.R.B. 730, 732–33 (1951); Electric Auto-Lite Co., 92 N.L.R.B. 1073, 1078 (1950), *aff'd*, 196 F.2d 500 (6th Cir. 1952). For a discussion of the "regularity" and "periodicity" requirements that a union must meet before it may exact union dues, see Cantor, *supra* note 81, at 63–65; Zipp, *supra* note 110, at 209–11.

[118] 431 U.S. 209 (1977).

[119] *Id.* at 235–36.

[120] *Id.* at 238.

[121] Kolinske v. Lubbers, 712 F.2d 471, 480 (D.C. Cir. 1983); *see* Cantor, *supra* note 81, at 94. In *Kolinske* the court rejected the argument that the decision to incorporate a security provision in the union's collective bargaining agreement with the employer was "state action" subject to constitutional scrutiny. 712 F.2d at 479–80.

[122] An irregular strike assessment may not satisfy the requirements of "periodicity" and "regularity." *See supra* note 117 and accompanying text; *see also* Cantor, *supra* note 81, at 63 ("The National Labor Relations Board . . . and reviewing courts have consistently ruled that the periodic dues and initiation fees collectible through union security provisions do not include 'assessments'."); Zipp, *supra* note 110, at 210–11 (arguing that an assessment "does not come within the gambit of the proviso [to section 8(a)(3)] if it is temporary in duration, of a special purpose nature not essentially related to the union's exclusive bargaining agent role, and is so regarded by the membership and union officials").

necessary, and accepted components of a collective bargaining system,"[123] a strike insurance fund surely fits into the rubric established by *Abood* for the legitimate uses of agency shop fees. Thus, when an employee resigns his or her membership in a union protected by a union security clause, even as a nonmember this employee must still continue to tender union fees.[124] If the employee resigns his or her membership during a strike, these fees will still help to finance the union's strike activities.[125]

This discussion of the similarities between the union shop and the agency shop illustrates that the courts have been motivated more by a desire to make the free rider pay his or her fair share than by an eagerness to enforce the notion of union solidarity against the wills of individual workers. By insisting, however, that nonmembers pay their share of collective bargaining expenses, the courts have implicitly acknowledged that the union's interest in preserving its financial security may outweigh an individual employee's right to refuse to associate with the union.

C. *Union Security and Reasonable Restrictions on Resignations*

This section will offer one view on when a union restriction on resignations during a strike or during the period immediately preceding the strike is reasonable. This view is based on the recognition that a union, protected by a security clause, may compel both members and nonmembers to contribute to a strike insurance fund.

One labor economist, Wallace Atherton, has suggested that the protection of a security clause has a moderating effect on a union's leadership during the collective bargaining process.[126] According to his model, the typical union's negotiators "will seek to maximize the organization's net revenue" when choos-

---

[123] T. KOCHAN, COLLECTIVE BARGAINING AND INDUSTRIAL RELATIONS 237 (1980); *see also* A. COX, D. BOK & R. GORMAN, *supra* note 6, at 484.

[124] *See* Marlin Rockwell Corp., 114 N.L.R.B. 553, 561 (1955).

[125] *See supra* note 81 and accompanying text.

[126] W. ATHERTON, THEORY OF UNION BARGAINING GOALS 38-39 (1973); *see also* T. HAGGARD, *supra* note 8, at 275 ("In the absence of a union security agreement, the employees who join and become active in the union will tend to be the more militant, aggressive, and perhaps even irresponsible employees who will select leaders to represent their views, even though a majority of the workforce is much more moderate in its outlook.").

[127] W. ATHERTON, *supra* note 126, at 38.

ing which bargaining objectives to pursue.[127] Because the expenses of a protracted strike will generally consume a large portion of a union's treasury, these negotiators will try to reach an earlier settlement in order to reduce expenditures and to maximize revenues.[128] This settlement may often fall short of satisfying the average union member's preferences.[129] On the other hand, Atherton suggests that when a union is unprotected by a security clause and the payment of union dues is entirely voluntary, the union's leadership will vigorously pursue ambitious bargaining goals in an attempt to increase the union's popularity and to widen its membership.[130] This pursuit may ultimately lead to a strike. In this way, the union leadership hopes to raise the union's level of income, despite the high costs of a possible strike.

Thus, a union not enjoying the protections of a security clause may often be forced to adopt a reckless "go-for-broke" bargaining strategy in order to increase membership and to maximize union revenues. Because the employees in the union's bargaining unit are not required to pay union fees under these circumstances, the union cannot depend upon regular fee payments to support this strategy. Although a union may arguably be satisfying the preferences of the employees in its bargaining unit by pushing hard for a variety of ambitious bargaining goals, this ambition may often lead to financial hardship and even to the extinction of the union itself.

It is no easy task to identify a list of objective factors by which to determine the economic strength and member solidarity of a union. It seems probable, however, that a union unable during collective bargaining to win any type of security provisions for its agreement is comparatively weak. One recent study conducted by the Bureau of Labor Statistics of the United States Department of Labor found that "union security provisions . . . were negotiated in 1,100 (83 percent) of the 1,327 major agreements covered by [the] study."[131] Furthermore, state right-to-work laws were responsible for the absence of security provisions in some of the remaining 227 agreements.[132]

---

[127] W. ATHERTON, *supra* note 126, at 38.

[128] *Id.*

[129] *Id.*

[130] *Id.*

[131] BLS BULLETIN, *supra* note 96, at 5. These agreements covered 90%, or 5.5 million, of the 6.1 million workers included in the study. *Id.*

[132] *Id.* The study did show, however, that many collective bargaining agreements

As these statistics indicate, only a small but significant minority of unions are not protected by some form of union security. It seems likely that in most cases these are unions that either have just been formed or have just recently been certified as the exclusive bargaining agent for a particular unit of employees.[133]

In *Allis-Chalmers* Justice Brennan noted the comparative strengths of different unions when discussing the proviso to section 8(b)(1)(A) and a union's power to expel one of its members for breaking a union rule:

> Where the union is strong and membership therefore valuable, to require expulsion of the member visits a far more severe penalty upon the member than a reasonable fine. Where the union is weak, and membership therefore of little value, the union faced with further depletion of its ranks may have no real choice except to condone the member's disobedience. Yet it is just such weak unions for which the power to execute union decisions taken for the benefit of all employees is most critical to effective discharge of its statutory function.[134]

By distinguishing weak unions in this way, Brennan highlighted the need for these unions to be able to enforce their rules effectively. If the courts did not allow such enforcement, Brennan implied, weak unions could not satisfactorily "discharge . . . their role as exclusive statutory bargaining agents . . . ."[135]

Although critical of the coerciveness of the union and agency shops and their deleterious effects on participation by union members in union affairs, Professor George Brooks of the Cornell School of Industrial and Labor Relations has nonetheless attempted to assist these weak unions by suggesting that a "compulsory union provision be permitted in all new bargaining relationships for seven years, after which it would be automatically cancelled."[136] Citing the recent re-emergence of union-

made in right-to-work states did contain some type of union security provision. Attempting to explain this surprising phenomenon, the study suggested that "[m]any of the union shop clauses might have been negotiated to take effect if future changes in state laws [were] to allow such provisions." *Id.*

[133] *See supra* note 82.

[134] *Allis-Chalmers*, 388 U.S. at 183–84 (citation omitted).

[135] *Id.* at 183.

[136] Brooks, *The Strengths and Weaknesses of Compulsory Unionism*, 11 N.Y.U. REV. L. & SOC. CHANGE 29, 38 (1982–83). Under this plan, assuming three-year collective bargaining agreements, a union shop provision would cover the first three negotiations, expiring after the first year of the third agreement. This plan assumes that union-management "relationships which do not 'mature' in seven years are very rare." *Id.*

busting activity and the growing antagonism between employers and unions, Brooks supports his suggestion by stating that "[i]t would be too much of a burden upon a newly-certified union to have to cope with that antagonism without a union shop."[137] This statement clearly evinces a concern that a union in its first years after certification is in a peculiarly vulnerable position vis-à-vis the employer.

Union rules restricting resignations can help reduce this vulnerability. First, proponents argue that a union rule that restricts resignations promotes the union's unity of purpose.[138] If union members were allowed to resign without hindrance during a strike, for example, the union's solidarity would be undermined. With a declining membership, a union could not present itself as a united and indivisible organization capable of challenging an employer during the heated days of a strike. Second, the proponents argue that a union rule that prevents resignations during a strike also prevents an employee from free riding.[139] By resigning membership during a strike and then returning to work, an employee bears none of the burdens of his striking colleagues, yet nonetheless enjoys any benefits won by the union after the strike ends.

The Supreme Court has expressed approval of the union solidarity goal in language contained in *Allis-Chalmers*:

> National labor policy has been built on the premise that by pooling their economic strength and acting through a labor organization freely chosen by the majority, the employees of an appropriate unit have the most effective means of bargaining for improvements in wages, hours, and working conditions. The policy therefore extinguishes the individual employee's power to order his own relations with his employer and creates a power vested in the chosen representative to act in the interests of all employees. Congress has seen fit to clothe the bargaining representative with powers comparable to those possessed by a legislative body both to create and restrict the rights of those whom it represents . . . .[140]

---

[137] *Id.*

[138] *See* T. HAGGARD, *supra* note 8, at 276.

[139] *See id.* at 272–73; BLS BULLETIN, *supra* note 96, at 4. In addition, "if strikers perceive that some workers are not carrying their fair share of the burden of the strike, their willingness to participate in the strike is adversely affected." Kolinske v. Lubbers, 712 F.2d 471, 473 (D.C. Cir. 1983).

[140] 388 U.S. at 180 (citations omitted). *See also* Emporium Capwell Co. v. Western Addition Community Organization, 420 U.S. 50, 61–64 (1975) (citing *Allis-Chalmers* to explain the policies underlying the principle of majority rule); Machinists Local 1327 v.

In subsequent cases, however, the Supreme Court has emphasized the individual union member's associational rights at the expense of the union's solidarity interest. The *Scofield* Court insisted that a union member be free to resign his membership and to "escape the [union] rule."[141] In *Granite State*, similarly, the Court dealt a sharp blow to legal claims based on the union solidarity interest by declaring that "the [union] member be free to refrain in November from the actions he endorsed in May and that his section 7 rights are not lost by a union's plea for solidarity or by its pressures for conformity and submission to its regime."[142] As this sampling of Supreme Court language demonstrates, the solidarity goal has fallen from its former position of prestige and privilege in the eyes of the Court itself.

The second goal—that of eliminating the free rider—has received more favorable treatment from the Supreme Court.[143] In *Oil, Chemical & Atomic Workers International Union v. Mobil Oil Corp.*,[144] for example, the Court reviewed the legislative history of section 8(a)(3) and applauded Congress's decision "to provide that there be no employees . . . getting the benefits of

---

NLRB (Dalmo Victor), 725 F.2d 1212, 1217 (9th Cir. 1984) (emphasizing that union members who participate in a strike vote mutually rely on each other to adhere to the requirements of the vote).

[141] *See supra* text accompanying notes 48–54.

[142] *Granite State*, 409 U.S. at 217–18; *see* Gould, *supra* note 1, at 91; *see also* International Association of Machinists and Aerospace Workers, Local Lodge 1414 v. Neufeld Porsche-Audi, Inc., 270 N.L.R.B. No. 209, 1983–84 NLRB Dec. (CCH) ¶ 16,436 (1984), at 28,095 n.12 (arguing that the *Granite State* Court rejected the "mutual subscription" theory requiring the preservation of solidarity during a strike); Pattern Makers' League of North America v. NLRB, 724 F.2d 57, 60 (7th Cir. 1983) (declaring that an employee's right to resign "cannot be overridden by union interests in 'group solidarity and mutual reliance . . .'"), *cert. granted* 105 S. Ct. 79 (1984) (No. 83-1894), *argued* Feb. 27, 1985, 53 U.S.L.W. 3632 (U.S. Feb. 12, 1985), *reargument ordered*, 53 U.S.L.W. 3686 (U.S. Mar. 26, 1985).

[143] Even opponents of union rules restricting resignations have conceded that "[t]he only solidarity which the union is entitled to *enforce* is that which comes from its position as exclusive bargaining agent—*i.e.*, every employee's obligation to pay the costs reasonably related to the union's duties as exclusive bargaining agent." Brief for Respondent at 9–10, Pattern Makers' League of North America v. NLRB, No. 83-1894 (U.S. filed May 18, 1984), *cert. granted* 105 S. Ct. 79 (1984), *argued* Feb. 27, 1985, 53 U.S.L.W. 3632 (U.S. Feb. 12, 1985), *reargument ordered*, 53 U.S.L.W. 3686 (U.S. Mar. 26, 1985). Although this statement appears to sanction union attempts to curb the free rider, it nonetheless contradicts another statement made later in the NLRB's brief: "Certainly an employee should have no *less* freedom to avoid union discipline when the contract contains no union security clause." *Id.* at 24 n.16. These two statements are irreconcilable. If a union is an exclusive bargaining agent, and if it incurs expenses in the performance of its duties, then the union is unable to enforce "the only solidarity" to which it "is entitled" only if the collective bargaining agreement contains no security clause, so the absence of such a clause does justify restricting the employees' freedom.

[144] 426 U.S. 407 (1976).

union representation without paying for them."[145] One year later, in *Abood*, the Supreme Court upheld the constitutionality of an agency shop agreement between a municipal employer and a union representing local municipal employees because "the desirability of labor peace is no less important in the public sector, nor is the risk of 'free riders' any smaller."[146] Finally, in a recent case[147] brought under section 2, Eleventh, of the Railway Labor Act,[148] the Court detailed those union activities in which the free rider problem is particularly acute and acknowledged that "[t]he very nature of the free rider problem and the government interest in overcoming it require that the union have a certain flexibility in its use of compelled funds."[149]

Although no union wants its members to resign during a strike, these resignations are particularly harmful to a union unprotected by a security clause and therefore susceptible to the free rider. Unable to compel its former members to pay union fees, this union faces the prospect of being left with a strike insurance fund precariously supported by a shrinking treasury. In addition, a union whose only form of security is a maintenance of membership clause faces a similar prospect, because some forms of these clauses allow union members to withdraw from the union and often do not require nonmembers to pay union fees.[150] In these instances, describing the maintenance of membership clause as a type of union security is somewhat inaccurate.

It is therefore appropriate that the courts adjudge the legitimacy of a union constitutional provision restricting resignations on the basis of the presence or absence of a valid security clause in the union's collective bargaining agreement with the employer. By allowing only those unions unprotected by a security

---

[145] *Id.* at 416.

[146] *Abood*, 431 U.S. at 224.

[147] Ellis v. Brotherhood of Railway, Airline and Steamship Clerks, 104 S. Ct. 1883 (1984), *aff'g in part and rev'g in part*, 685 F.2d 1065 (9th Cir. 1982).

[148] 45 U.S.C. § 152, Eleventh (1982). Section 2, Eleventh permits union security clauses in the railroad industry, including union shop clauses.

[149] Ellis v. Brotherhood of Railway, Airline and Steamship Clerks, 104 S. Ct. 1883, 1896 (1984), *aff'g in part and rev'g in part*, 685 F.2d 1065 (9th Cir. 1982).

[150] *See* BLS BULLETIN, *supra* note 96, at 8. A typical maintenance of membership clause provides:

> Subject to applicable law, all employees who, as of the date of this agreement are members of the union in good standing in accordance with the constitution and by-laws of the union or who become members of the union following the effective date of this agreement shall, as a condition of employment, remain members of the union in good standing insofar as the payment of periodic dues and initiation fees, uniformly required, is concerned.

Id.

clause to restrict resignations during a strike, or during those tense days before a strike, the courts would reasonably balance the associational rights of the individual worker and the legitimate institutional needs of a struggling union.

IV. Conclusion

Restricting the resignation of members whose union does not enjoy a security provision may appear at first to fail the *Scofield* test requiring that a union member be "free to leave the union and escape the rule."[151] But, as the Court itself noted in *Scofield*, the affected union was protected by an agency shop agreement that compelled nonmembers in the bargaining unit to pay a substantial service fee.[152] Thus, the union in *Scofield* was shielded from the burdens of the free rider. Although the Court ruled that the disgruntled employees must be able to resign from the union and to violate the production quota, the financial contributions of those employees continued to enhance the union's treasury, even after they had tendered their resignations.[153]

This distinction makes a great difference. Weaker unions without a security clause are more in need of preserving their financial base, yet are also more vulnerable to defections and free riders, making financial stability difficult. To create conditions that make effective union representation possible, a restriction on resignations during a strike becomes reasonable and even necessary, while it would be outweighed by the employee's section 7 rights if the union were stronger. Thus, the proposed framework focuses primarily on the strength or weakness of individual unions rather than on union and employee rights considered in the abstract. It looks to the presence of a security

---

[151] As a way to satisfy the requirements of the *Scofield* test and to preserve union solidarity at the same time, one commentator has suggested that all employees wishing to resign during a strike be required to make a "special showing of hardship." *Conflict*, *supra* note 23, at 880.

[152] *Scofield*, 394 U.S. at 424 n.1.

[153] In addition, one commentator has suggested that the third part of the *Scofield* test should not be applied in determining the validity of resignation restrictions. Because the rule at issue in *Scofield* was a production ceiling, testing its validity in light of the affected employee's ability "to leave the union and escape the rule" is appropriate. When a union restricts the right to resign, however, application of the third part of the *Scofield* test would automatically invalidate the restriction. Application of the *Scofield* test in this way would render its first and second parts superfluous. Recent Development, *Protecting a Union Member's Right to Resign—Resolution of the Conflict Between* Dalmo Victor *and* Rockford-Beloit, 38 Vand. L. Rev. 201, 233–34 (1985).

clause as an objective test of strength. It looks to overriding purposes of the Act rather than to a narrow application of specific sections: it seeks to preserve "the usefulness of labor's cherished strike weapon."[154] Allowing weaker unions without a security clause to restrict member resignations during a strike, or during the period immediately preceding a strike, is a reasonable means of promoting the Act's underlying policy of protecting both employers and organized labor.[155]

This Comment suggests that section 8(b)(1)(a) should be read as a limit on section 7.[156] The test proposed by the Comment, however, does not constitute an attempt to place an employee's right to refrain from associational activities and a union's interest in enforcing its rules on an equal footing. Under the proposed test, the section 7 right would remain unlimited unless and until its exercise threatened not just the union's solidarity interest, but also the union's ability to maintain itself financially during the course of collective bargaining. Because this ability, and the threat to it, is not susceptible of direct measurement, the Comment suggests using the easily ascertainable existence or nonexistence of a security clause as a substitute measure.

Only a small minority of unions are not protected by some security clause.[157] In addition, even if the union could enforce a rule restricting resignation and requiring continued membership during a strike, the obligations of membership have been "whittled down to [their] financial core."[158] Thus, in practice the proposed test would probably not, and is not intended to, equalize the union's interests and the employee's section 7 rights. Unlike the approach of both the current NLRB and the Ninth Circuit in *Dalmo Victor*, the proposed test would require a balancing of competing rights and interests to resolve the union-employee conflict. This balancing, however, respects the importance of the individual employee's right to resign.

---

[154] NLRB v. Allis-Chalmers Mfg. Co., 388 U.S. 175, 183 (1967), *reh'g denied*, 389 U.S. 892 (1967).

[155] "The purpose of federal labor legislation is to reconcile and, insofar as possible, equalize the power of competing economic forces within the society in order to encourage the making of voluntary agreements governing labor-management relations and prevent industrial strife." Pittsburgh Plate Glass Co. v. NLRB, 427 F.2d 936, 946 (6th Cir. 1970), *aff'd* 404 U.S. 157 (1971).

[156] *See supra* notes 83–94 and accompanying text.

[157] *See supra* note 131 and accompanying text.

[158] NLRB v. General Motors Corp., 373 U.S. 734, 742 (1963); *see supra* text accompanying notes 114–16.

COMMENT

STATE DISCRIMINATORY ACTION AGAINST NONRESIDENTS: USING THE ORIGINAL POSITION THEORY AS A FRAMEWORK FOR ANALYSIS

ADAM B. SCHIFF*

States seeking to hoard their resources in times of shortage or attempting to bolster their economy during periods of recession have found legislation which discriminates against nonresidents to be an attractive method of achieving their protectionist goals. State passage of discriminatory legislation is limited by the federal government's power to regulate interstate commerce[1] and by the entitlement of the citizens of each state to "all Privileges and Immunities of citizens in the several states."[2] If the Supreme Court is to deal effectively with states' attempts to cordon themselves off from the nation's problems, and yet preserve the states' individual identities, it must possess a coherent and articulate view of federalism. The following analysis of the Court's recent opinions confronting the problem of state discriminatory laws illustrates that the Court has yet to articulate such a view.

The Court has analyzed cases of state discriminatory legislation under the Commerce Clause of Article I and the Privileges and Immunities Clause of Article IV of the Constitution.[3] Because the two clauses have a "relationship that stems from their common origin in the Fourth Article of the Articles of Confederation and their shared view of federalism,"[4] the Court has naturally tended to draw on the doctrine of one clause to aid analysis of discriminatory legislation under the other clause. The Court, however, has recently adopted divergent conceptual frameworks of analysis under the two clauses, highlighting their differences instead of their common goals; it has also failed to

---

* A.B., Stanford University, 1982; Member, Class of 1985, Harvard Law School.

[1] U.S. CONST. Art. I, § 8.

[2] U.S. CONST. Art. IV, § 2.

[3] U.S. CONST. Art. I, § 8; *id.* Art. IV, § 2. *See generally*, L. TRIBE, CONSTITUTIONAL LAW § 6-2, at 320–21 (1978).

[4] Hicklin v. Orbeck, 437 U.S. 518, 531–32 (1978) (footnote omitted).

articulate a clear vision of federalism to inform either system of rules.

The absence of a coherent view of federalism is most apparent in the so-called dormant Commerce Clause cases. The Constitution gives Congress the power to regulate commerce "among the several states."[5] It remains silent, however, on the question of whether states may exercise a similar power in the area of interstate commerce and, if so, to what extent this power may be exercised. Although the text of the Constitution is silent, Supreme Court decisions have developed the dormant Commerce Clause doctrine to define the states' power to act in this area.[6] The doctrine defines the parameters of the states' power by interpreting the Commerce Clause as implicitly blocking states from acting in a way that burdens interstate commerce. Within the last decade, however, the Court has reverted to an historical distinction between state proprietary action and state regulatory action, which allows states to avoid Commerce Clause scrutiny and to favor their own citizens when the states enter the market as participants. The Court, acting largely by intuition, has had great difficulty providing a coherent framework for this potentially all-encompassing doctrine.

The Court's privileges and immunities analysis is also based on an historical anachronism, for the Court has revived a "fundamental" rights basis for the application of the Clause, which dates back as far as 1825[7] and was thought to have been put to rest in 1948.[8] The Privileges and Immunities Clause provides that citizens of each state are entitled to "all the Privileges and Immunities of Citizens in the several states," and thus secures the ability of nonresidents to enjoy the same privileges which residents of a state receive.[9] The Court, however, has recently found that this protection extends only to privileges that may be deemed fundamental.[10] The resurrected fundamental rights analysis threatens to widen the discontinuities in the application of the two clauses, preventing achievement of their shared goals and rendering the Court's view of federalism incoherent.

---

[5] U.S. CONST. Art. I, § 8.

[6] For early decisions concerning the scope of the Commerce Clause see, for example, Gibbons v. Ogden, 22 U.S. (9 Wheat.) 1 (1824); Cooley v. Board of Wardens, 53 U.S. (12 How.) 299 (1851).

[7] Corfield v. Coryell, 6 Fed. Cas. 546 (C.C.E.D. Pa. 1825) (No. 3230).

[8] Toomer v. Witsell, 334 U.S. 385, 395–407 (1948).

[9] U.S. CONST. Art. IV. § 2.

[10] Baldwin v. Fish and Game Commission of Montana, 436 U.S. 371, 388 (1978).

This Comment argues that the Court's intuition in deciding cases of state discriminatory legislation reflects a vision of federalism that can be developed by analogy to the "original position" concept utilized by John Rawls in his *A Theory of Justice*.[11] Rawls's original position is a hypothetical environment where principles chosen by individuals are fair. "It is a state of affairs in which the parties are equally represented as moral persons and the outcome is not conditioned by arbitrary contingencies or the relative balance of social forces."[12] This Comment argues that the application of the original position theory to problems of federalism provides a coherent framework for Commerce Clause and Privileges and Immunities Clause analyses. The essential idea is to put states in the original position and to deduce what principles of federalism would consequently evolve. With respect to the Commerce Clause, the view of federalism that emerges from this theory will show that the Court's distinctions are not wholly without merit. In fact, the original position theory of federalism will show that the Court's distinctions, if applied properly and not mechanically, can be useful in indicating which state actions are proper. In the area of the Privileges and Immunities Clause, the original position theory of federalism will show why the Court has intuitively resorted to a revival of the fundamental rights analysis, and it will demonstrate why this approach is misguided.

I. The Dormant Commerce Clause

The dormant Commerce Clause cases rely on a few mechanistically applied distinctions. These distinctions are based on: (1) whether the state has imposed regulations to prohibit some form of commerce or has participated in the market; (2) whether the good in question is a natural resource or is the product of state "foresight, risk and industry";[13] and (3) whether the good is state owned or privately owned.

Like all mechanistically applied distinctions, these three tend to break down because the policy behind them is not clearly articulated. The difficulty with the dormant Commerce Clause

[11] J. Rawls, A Theory of Justice (1971).

[12] *Id.* at 12.

[13] Reeves, Inc. v. Stake, 447 U.S. 429, 446 (1980) (cement from state-owned plant is the result of state foresight, risk, and industry).

analysis is that the Court gives undue weight to some of these distinctions and yet has no clear conception of why the distinctions should be given weight at all. A survey of the cases illustrates the problem.

The first recent dormant Commerce Clause case, *Hughes v. Alexandria Scrap Corp.*,[14] pronounced what would soon become the most critical of the dormant Commerce Clause distinctions. The Court in that case emphasized the difference between the state as market regulator and the state as market participant. The Court precluded the state from regulating interstate commerce in a discriminatory fashion but allowed the state to participate in the market and exercise its right thereby to favor its citizens.[15] In *Alexandria Scrap*, Maryland had provided a subsidy for the processing of old automobile hulks that favored in-state processors over out-of-state processors. The Court found that Maryland had not interfered with the functioning of the interstate market through burdensome regulation. Instead, its actions—essentially bidding up the price of the hulks[16]—were those of a market participant and therefore unsuited for Commerce Clause scrutiny.[17] The Court argued that the distinction between market regulator and market participant can be inferred from the history of the Commerce Clause, which was traditionally targeted at regulatory and taxing actions that impeded free private trade.[18] The Court broadly stated that "[n]othing in the purposes animating the Commerce Clause prohibits a State, in the absence of congressional action, from participating in the market and exercising the right to favor its own citizens over others."[19]

The facts of *Alexandria Scrap* did not provide an opportunity for the Court to demonstrate clearly the scope of the participatory action that would be upheld under the dormant Commerce Clause. It is not difficult to realize that if all state actions characterized as market participation are taken out of the strictures of the Commerce Clause, then only time prevents the states from structuring many, if not most, discriminatory actions in the form of market participation. More importantly, however,

---

[14] 426 U.S. 794 (1976).
[15] *Id.* at 806, 810.
[16] *Id.* at 806.
[17] *Id.* at 806.
[18] *Id.* at 807–08; *see* L. TRIBE, *supra* note 3, § 6-10, at 336.
[19] *Id.* at 810 (footnote omitted).

there is no reason why state discriminatory actions that take the form of market participation should inherently be immune from scrutiny under the Commerce Clause.

The Court's historical justification for the participant/regulator distinction is insufficient. It may be that the Framers primarily intended the Commerce Clause to create national authority to regulate interstate commerce, and sought to prohibit state regulation that inhibited the flow of such commerce, because they believed state regulation to be the greatest threat to free trade and interstate harmony.[20] But even though the Framers may not have anticipated the enormous growth of state government and the potentially disruptive effect of state participation in the market, that does not render the Commerce Clause incapable of addressing the problem.

In other areas, the Commerce Clause has not been so narrowly interpreted.[21] That the Framers may not have anticipated the large scale nature of modern interstate business has not prevented the Court from upholding regulations when Congress has applied its regulatory power to local evils that have interstate effects in cases in which such regulations are necessary to effectuate the purposes of the Commerce Clause.[22] The means of the threat may have been unanticipated, but the purpose of the Commerce Clause—to promote interstate harmony and to remove "injurious impediments to intercourse between different [states]"[23]—was clear from the start. The Court's naked assertion in *Alexandria Scrap* that "[n]othing . . . prohibits a State, in the absence of Congressional action, from participating in the market and exercising the right to favor its own citizens over others"[24] is conclusory and unjustified.

---

[20] For a discussion of the origins and development of the Commerce Clause and the state proprietary exception, see Blumoff, *The State Proprietary Exception to the Dormant Commerce Clause: A Persistent Nineteenth Century Anomaly*, 1 S. ILL. L.J. 73 (1984).

[21] For a recent example, see Garcia v. San Antonio Metropolitan Transit Authority, 53 U.S.L.W. 4135 (U.S. Feb. 19, 1985) (holding that the Commerce Clause permits Congress to apply the wage and hour provisions of the Fair Labor Standards Act to employees of a state mass-transit authority). Note also the Court's rejection of the historical approach to state immunity from taxation, recognizing that such an approach "prevents a court from accomodating changes in the historical functions of states." *Id.* at 4139.

[22] *See, e.g.*, Hodel v. Virginia Surface Mining and Reclamation Ass'n, 452 U.S. 264, 276–77 (1981); Heart of Atlanta Motel v. United States, 379 U.S. 241, 258 (1964); Wickard v. Filburn, 317 U.S. 111, 125 (1942).

[23] THE FEDERALIST NO. 22, at 144–45 (A. Hamilton) (McClean ed. 1961).

[24] *Alexandria Scrap*, 426 U.S. at 794, 810 (footnotes omitted).

The decision in *Reeves, Inc. v. Stake*,[25] one of the Court's next major opportunities to confront state discriminatory legislation, turned on the distinction between natural resources and produced goods and the distinction between state owned resources and privately owned resources. In *Reeves* the Court faced a South Dakota policy dictating that in times of shortage cement produced at a state owned plant could be sold only to state residents.[26] The Court relied on *Alexandria Scrap* in finding that South Dakota was acting as a market participant,[27] arguing that "South Dakota, as a seller of cement, unquestionably fits the 'market participant' label more comfortably than a State acting to subsidize local scrap processors."[28]

Dismissing an argument that upholding a policy in favor of residents would allow states to hoard natural resources, the Court introduced the distinction between a natural resource "like coal, timber, wild game, or minerals . . ." and an "end product of a complex process whereby a costly physical plant and human labor act on raw materials."[29] Natural resources were to be distinguished from the products of a state's "foresight, risk and industry."[30] A state could not claim exclusive rights to the natural resources found fortuitously within its borders although it could, however, maintain control over the results of its investments when it developed those resources. The Court in *Reeves* also distinguished between resources that were state owned and those that were privately owned. The Court emphasized that the cement production was accomplished in a state plant, pursuant to a state program, and funded by state residents.[31] The decision implied that if the plant were privately owned, or the cement produced by the state were reduced to private possession, and the state tried to administer a program favoring residents, that program would not survive scrutiny under the Commerce Clause.[32]

The *Reeves* Court gave no reasons for either of its distinctions. The Court failed to express a coherent view of federalism

[25] 447 U.S. 429 (1980).
[26] *Id.* at 402 (Brennan, J., dissenting).
[27] *Id.* at 440.
[28] *Id.*
[29] *Id.* at 443–44.
[30] *Id.* at 446.
[31] *Id.* at 442.
[32] If the state required privately owned cement to be sold on a residents-first basis, then that regulation would also be subject to attack as an impermissable market regulation because there is no state proprietary interest in the private industries. *See id.*

because it explained neither how natural resources are to be distinguished from the fruits of state efforts[33] nor why the distinction should be made at all.[34]

In *White v. Massachusetts Council of Construction Employers*,[35] the Court reaffirmed the principle of *Alexandria Scrap* and *Reeves* that "when a state or local government enters the market as a participant it is not subject to the restraints of the Commerce Clause."[36] The Court upheld an executive order of the Mayor of Boston requiring that all construction projects funded in whole or in part by city funds employ work crews, at least half of which were made up of city residents.[37] Justices Blackmun and White, in a concurring and dissenting opinion, saw the order as going beyond the choice of parties with whom the city would deal, and as restricting the choice of whom private employers may hire.[38] The majority of the Court, however, seemed to find the case uncomplicated, stating that the sole inquiry was whether the state directly participated in the market.[39] Impact on out-of-state residents would be a consideration in Commerce Clause analysis only if it were decided that the city regulated the market.[40]

In *South-Central Timber Development, Inc. v. Wunnicke*,[41] the Court struggled to define the limits of the market participant/market regulator distinction. *South-Central Timber* concerned an Alaska statutory requirement that timber purchased from the state be partially processed within Alaska before being shipped out-of-state.[42] The Court held the statute unconstitutional, because timber is a raw natural resource[43] and because the burden on commerce created by the statute affected not only those

---

[33] There is no clear distinction drawn between the process used to extract and distribute a resource and the process that converts that resource into a produced good.

[34] The Court's perfunctory application of the distinction "precluded a serious consideration of Reeves's claim that the practical impact of the state activity on interstate commerce should be the determinative factor." Case Comment, *Constitutional Law—Commerce Clause—State Market Activity Exempt from Commerce Clause Review—State-run Cement Plant May Withhold Supply from Nonresidents to Preserve Cement for Residents*, 27 WAYNE L. REV. 1575, 1590 (1981).

[35] 460 U.S. 204 (1983).

[36] *Id.* at 208.

[37] *Id.* at 205–06.

[38] *Id.* at 217 (Blackmun, J. and White, J., concurring and dissenting).

[39] *Id.* at 206–08.

[40] *See id.* at 214–15.

[41] 104 S. Ct. 2237 (1984).

[42] *Id.* at 2239.

[43] *Id.* at 2245.

involved in the immediate transaction, but also restricted resale to third parties.[44]

Attempting to give some indication of the contours of the market participant/market regulator distinction, the Court indicated that "[t]he limit of the market-participant doctrine must be that it allows a State to impose burdens on commerce within the market in which it is a participant, but allows it to go no further."[45] This explanation only restated the issue: how broad is the market?

The Court went on to try to support its application of the market participant/market regulator distinction: "There are sound reasons for distinguishing between a State's preferring its own residents in the initial disposition of goods when it is a market participant and a State's attachment of restrictions on dispositions subsequent to the goods coming to rest in private hands."[46] The Court's reasons, however, were not persuasively articulated. First, the Court relied "simply as a matter of intuition"[47] on the notion that a state as market participant has a greater interest in the actual sales transaction than it does in the buyer's subsequent use of the goods after he acquires them.[48] The Court found it "unimportant for present purposes that the state could support its processing industry by selling only to Alaska processors, by vertical integration, or by direct subsidy."[49] Second, the Court argued that "downstream restrictions have a greater regulatory effect than do limitations on the immediate transaction."[50] As the dissent pointed out, however, this merely restates the conclusion.[51] Nonetheless, the Court concluded that Alaska was a participant only in the timber market, and when imposing conditions downstream in the processing market, it could not avail itself of the protection under the market participant doctrine.[52]

Although the Court professes to rely on precedent for its Commerce Clause distinctions, *South-Central Timber* illustrates that intuition motivates the Court's decisions. The Court's dis-

---

[44] *Id.* at 2244.
[45] *Id.* at 2245.
[46] *Id.* at 2246.
[47] *Id.*
[48] *Id.*
[49] *Id.*
[50] *Id.*
[51] *Id.* at 2248 n.* (Rehnquist, J., dissenting).
[52] *Id.* at 2246.

tinctions are not without merit, and its intuition is not without reason. Yet, the Court has failed to articulate a coherent theory which illuminates its distinctions and informs its intuition. After exploring the privileges and immunities doctrine, it will be possible to show that although the Court has not enunciated a comprehensive theory of federalism, such a theory can nevertheless be formulated.

II. The Privileges and Immunities Clause Cases

The Court has failed to apply the distinctions espoused in the dormant Commerce Clause cases to the problems of federalism raised by cases dealing with the Privileges and Immunities Clause. Instead, the Court has resorted to a fundamental rights test that forces compilation of a list of qualifying rights.

In *Toomer v. Witsell*,[53] the Court reviewed an array of South Carolina statutes that imposed significant restrictions on the ability of nonresidents to catch shrimp off its coast. The Court struck down as violative of the Privileges and Immunities Clause the imposition of a discriminatory licensing fee on boats that were owned by nonresidents and used to catch shrimp. The state argued that fish and game are the "common property of all citizens of the governmental unit and that the government, as a sort of trustee, exercises this 'ownership' for the benefit of its citizens."[54] Further, the state claimed that each government may "regulate the corpus of the trust in the way best suited to the interests of the beneficial owners, its citizens, and may discriminate as it sees fit against persons lacking any beneficial interest."[55]

The Court, in rejecting the ownership theory, stated that it is "now generally regarded as but a fiction . . . that a State ha[s] power to preserve and regulate the exploitation of an important resource."[56] Instead the Court held that the Privileges and Immunities Clause barred discrimination against nonresidents "where there is no substantial reason for the discrimination beyond the mere fact that they are citizens of other states."[57] The Court then applied its substantial reason test in a three step

[53] 334 U.S. 385 (1948).
[54] *Id.* at 399.
[55] *Id.* at 400.
[56] *Id.* at 402.
[57] *Id.* at 396.

analysis, focusing first on whether the non-residents were a "peculiar source of the evil at which the statute was aimed";[58] second, on whether the record demonstrated that in fact the legitimate interests of the statute were satisfied by the discrimination;[59] and third, on whether there were less restrictive alternatives.[60]

Prior to *Toomer*, in precedents stretching as far back as *Corfield v. Coryell*,[61] the Court interpreted the Privileges and Immunities Clause as requiring states to give equal treatment to citizens of other states when fundamental rights were at stake. This doctrine necessitated the inherently difficult process of enumerating which rights were to be deemed "in their very nature fundamental, which belong, of right to the citizens of all free governments."[62] The *Toomer* decision's analysis, focusing on underlying justifications for the discriminatory action, appeared to be a move away from the static use of a list of fundamental rights. As Professor Laurence Tribe points out, "*Toomer v. Witsell* dramatically shifted the focus of review under the Privileges and Immunities Clause from categorizing fundamental rights of state citizenship to analyzing state justifications for maintaining the challenged discriminatory burdens."[63]

In the recent case of *Baldwin v. Fish and Game Commission of Montana*,[64] however, the Court again shifted back to the fundamental rights approach, applying the three step analysis of *Toomer* only after the object of state discrimination has been found to implicate a fundamental right of the out-of-state resident.

In *Baldwin* the Court faced another discriminatory licensing scheme. In this case however, the Court sustained state legislation that charged nonresidents considerably higher premiums than it charged residents to hunt elk. Although recognizing that the ownership of a resource distinction was "no more than a 19th century legal fiction,"[65] the Court nonetheless was not con-

---

[58] *Id.* at 398.

[59] *Id.* at 403.

[60] *See id.* at 396; *see also* L. TRIBE, *supra* note 3, § 6-33 at 410.

[61] Corfield v. Coryell, 6 Fed. Cas. 546, 551–52 (No. 3230) (C.C.E.D. Pa. 1823); *see also* Slaughter House Cases, 83 U.S. (16 Wall.) 36, 75–76 (1873) (Court relied on the *Corfield* discussion of "fundamental rights" in intrepreting the 14th Amendment).

[62] *Corfield*, 6 Fed. Cas. at 552.

[63] L. TRIBE, *supra* note 3, § 6-33 at 410.

[64] 436 U.S. 371 (1978).

[65] *Id.* at 386.

vinced to "completely reject" its earlier decisions drawing on that distinction.[66] Finding that the hunting of elk was not one of those "basic and essential activities, interference with which would frustrate the purposes of the formation of the Union," the Court permitted the discriminatory treatment.[67]

Less than a month later, the Court, in *Hicklin v. Orbeck*,[68] further spelled out the role state ownership of a resource would play in future privileges and immunities doctrine and expanded the scope of analysis under that clause. Ownership would be a factor, often crucial, in deciding whether the statute violated the clause. Ownership would not, however, be sufficient to place the statute outside the scope of the clause.

In *Hicklin*, Alaska sought to require that all agreements relating to oil and gas contain a provision preferring the employment of qualified Alaska residents over nonresidents.[69] The Court in *Hicklin* applied a two step test, inquiring first whether nonresidents were a peculiar source of the evil that the law was enacted to remedy and, second, whether the discriminatory means were substantially related to the problem that the law addressed.[70] The Court found that neither step of the test was sufficiently satisfied:[71] the state had failed to show that nonresidents were a major cause of Alaskan unemployment,[72] and further had failed to tailor its solution—an absolute preference for all residents—closely enough to the problem.[73]

Alaska argued that ownership of the oil and gas was sufficient justification for the Act's discrimination against nonresidents, and that the Act was "totally without the scope of the Privileges and Immunities Clause."[74] The Court rejected this argument. Although the Court assumed that providing access to employment for state residents was an interest which the states could conduct proprietary activities to advance, the Court found Alaska's proprietary position too attenuated. Alaska "has little or no proprietary interest in much of the activity swept within the ambit of Alaska Hire; and the connection of the State's oil

[66] *Id.* at 385.
[67] *Id.* at 388.
[68] 437 U.S. 518 (1978).
[69] *Id.* at 520.
[70] *Id.* at 526.
[71] *Id.* at 527.
[72] *Id.* at 528.
[73] *Id.* at 529.
[74] *Id.* at 526–27.

and gas with much of the covered activity is sufficiently attenuated so that it cannot justifiably be the basis for requiring private employers to discriminate against nonresidents."[75]

The Court gave no indication of when ownership became too attenuated to support discrimination. As with the Court's definition of the market in *South-Central Timber*, the scope of ownership is dependent upon the Court's ill-defined intuition.

In *United Building and Construction Trades Council v. Camden*,[76] the Court again faced the situation presented in *White*. The city of Camden, New Jersey passed an ordinance requiring that forty percent of workers hired by city construction project contractors and subcontractors had to be city residents.[77] The distinction between market participant and market regulator, relied upon in *White* to dispose of the Commerce Clause challenge, was set aside as nondispositive in *Camden*,[78] because the order was being challenged under the Privileges and Immunities Clause. The Court reasoned that the Privileges and Immunities Clause bars any type of discriminatory action that impermissibly burdens fundamental rights, not merely regulatory action.[79] The Court examined the distinction in *Hicklin*, which was based on the ownership doctrine.[80] In *Camden*, the city's ownership of its money was a factor, "perhaps the crucial factor," in the Privileges and Immunities Clause analysis.[81] But this fact did "not remove the Camden ordinance completely from the purview of the Clause."[82] The Court concluded that the ordinance discriminated against a fundamental, and therefore protected, privilege.[83] The Court therefore remanded the case for an evaluation of the city's justification for the discrimination.[84] Justification is to hinge on whether there are "substantial reasons" for the discrimination and whether the discrimination is closely related to these reasons.[85]

An overview of the Privileges and Immunities Clause and the Commerce Clause cases reveals some contradictory features.

---

[75] *Id.*
[76] 104 S. Ct. 1020 (1984).
[77] *Id.* at 1023.
[78] *Id.* at 1028.
[79] *Id.* at 1028–29.
[80] *Id.* at 1029.
[81] *Id.*
[82] *Id.*
[83] *Id.*
[84] *Id.* at 1030.
[85] *Id.*

Characterization of the state as a market participant takes the state action completely outside of the scrutiny of the Commerce Clause.[86] Categorization of a right as nonfundamental removes it from the protection of the Privileges and Immunities Clause.[87] And as *Camden* illustrates, state ownership of the resource in question is one factor, possibly a crucial factor, in avoiding the Privileges and Immunities Clause analysis.[88] When taken together, the above features yield a conception of federalism that is confused; large holes in the conception illustrate that the shared concerns of the two clauses will go unsatisfied, nonresidents will often go unprotected, and interstate harmony will suffer.

The revived fundamental rights approach does little to fulfill the Privileges and Immunities Clause's purpose most clearly articulated in its predecessor in the Articles of Confederation—"to secure and perpetuate mutual friendship and intercourse among the people of the different States in this Union."[89] Discriminatory state action that does not affect fundamental rights or that is tied to a resource technically owned by the state, although that action may be without justification and deeply offensive to nonresidents, now cannot be prohibited under either of the two clauses. It is unrealistic to assume that because a state arbitrarily discriminates against nonresidents in a nonfundamental way that "serious sources of animosity and discord" would not arise.[90]

Similarly, in terms of the Commerce Clause, it cannot be seriously contended that discriminatory state legislation taking the form of participation does not have an impact on interstate commerce and does not provide a potentially significant threat to mutual friendship and intercourse. In *Reeves*, for example, South Dakota controlled all of the cement in the area through its market participant activities.[91] The Court nonetheless allowed South Dakota's preference for residents under the Commerce Clause. It is not evident why, if the state's rights as a

[86] *See supra* text accompanying note 17.
[87] *See supra* text accompanying note 67.
[88] *See supra* text accompanying notes 81–82.
[89] ARTS. OF CONFED. Art. IV.
[90] One commentator suggests that in certain circumstances nonresidents might feel greater hostility towards state discrimination in the area of nonfundamental rights, than in the area of fundamental rights. A hypothetical law excluding nonresidents from state beaches is offered as one example. *See* Varat, *State "Citizenship" and Interstate Equality*, 48 U. CHI. L. REV. 487, 511–12 (1981).
[91] *Reeves*, 447 U.S. at 432.

free trader are subject to limitation under one clause, they may not be subject to limitation under the other when both clauses share the same view of federalism. The Court's explanation, that the Privileges and Immunities Clause's concern with interstate harmony cuts across the participant/regulator distinction, does not explain why the Commerce Clause's concern with interstate harmony should not also cut across the distinction.

The Commerce Clause and the Privileges and Immunities Clause stem in part from a "shared vision of federalism."[92] It is therefore inconsistent that the two could yield opposite results in cases as similar as *White* and *Camden*. The cases discussed above illustrate that the distinction between market participant/ market regulator, resource/product, state ownership/private ownership, and fundamental rights/nonfundamental rights do not provide an adequate framework for analysis of state discriminatory legislation and tend to break down conceptually when the Court's intuition points toward a different result. The original position theory, however, provides a coherent underpinning for the Court's analysis of federalism.

III. The Original Position and Federalism

The original position is characterized by a veil of ignorance that "nullif[ies] the effects of specific contingencies which put men at odds and tempt them to exploit social and natural circumstances to their own advantage."[93] Under this veil of ignorance the parties do not know their place in society, their fortune in terms of natural assets and abilities, their conception of the good, their rational plan of life, or the particular circumstances of their own society.[94] The agreement reached in the original position sets out the principles of justice for the structure of society. Because these principles are chosen in a position of moral equality, they may be seen as fair. It is not important that the original position has never, or could never, be entered into; its primary utility is in adopting its perspective and deducing

---

[92] Arts. of Confed. Art. IV; *see* Hicklin v. Orbeck, 437 U.S. 518, 531–32 (1978). "[There is a] mutually reinforcing relationship between the Privileges and Immunities Clauses of Art. IV and the Commerce Clause—a relationship that stems [in part from] their common origin in the Fourth Article of the Confederation and their shared vision of federalism." *Id.*

[93] J. Rawls, *supra* note 11, at 136.

[94] *Id.* at 136–42.

principles to guide our behavior.[95] In a concrete case, it would be very difficult to put aside our prejudices and determine what would happen if the parties in the original position were confronted with the controversy. The original position is therefore most helpful when it is used to derive general principles that may be applied in a given case, rather than determining how parties under the veil of ignorance would react in the case itself.

For the purposes of this Comment, then, principles of federalism derived from placing the states in the original position will be applied in the context of our federal system and under our Constitution.[96]

The states in the original position would assume the positions of parties entering into a social contract, with the obvious goal of promoting their own good. Under the original position's veil of ignorance, however, the states would not know what natural resources they possessed. Ignorant of the natural resources under their soil and of the quality of their climates, the states would agree that the natural resources of any state could not be hoarded to the detriment of the union; instead, these accidents of nature would be used for the common benefit.

States in the original position would not necessarily know their citizens' conception of the good. The states would want to insure that when they emerged from the original position they would be able to serve their citizens by providing both the opportunities and the environment that their citizens would desire. Therefore, they would want maximum liberty to act affirmatively to pursue their citizens' conception of the good.

Thus, the first principle of federalism likely to be chosen by the states in the original position is the following: *each state is to have an equal right to the most extensive liberty to act affirmatively in order to provide for its citizens, so long as such action is compatible with a similar system of liberty for all states*. This principle guarantees that a state will have the greatest liberty possible to provide employment, property, political

[95] It is not necessary that the author demonstrate that Rawls's justifications for the principles individuals would choose in the original position would be equally true of states in the original position. Borrowing yet another technique from Rawls, it is enough to take on this assumption as a "provisional fixed point" and, comparing the results of principles so chosen with the Court's considered judgments of federalism, reach a "reflective equilibrium" from which we can "render coherent and justify our convictions" about proper state relations. *Id.* at 20–21.

[96] Although the theory proposed here may have broader application, this Comment attempts only to define state interaction in the absence of congressional action; it does not analyze state-federal conflicts in situations where Congress has acted.

representation, privacy, and liberty of speech and thought for its citizens.[97]

This idea is compatible with the intuitive justification for the resource/product distinction. States may not hoard natural resources because the states, by so doing, would intrude on the liberty guaranteed other states. When a state begins to develop a resource, however, the state is creating an asset for its citizens, not just benefiting fortuitiously from its natural endowment. Once the state has begun to invest its own efforts in providing for its citizens, it ought to have maximum freedom to act on their behalf,[98] so long as its liberty to do so is compatible with a similar liberty enjoyed by other states.

The first principle, standing alone, though, provides an incomplete theory of federalism. A state endeavoring to secure the best interests of its citizens might be reluctant to rely solely on

---

[97] It is significant that these liberties, corresponding to those which Rawls enumerates for individuals in the just society, J. RAWLS, *supra* note 11, at 61, roughly correspond to the attributes of state sovereignty protected by the now overruled National League of Cities v. Usery, 426 U.S. 833, 845–52 (1976). *See* Garcia v. San Antonio Metropolitan Transit Authority, 53 U.S.L.W. 4135 (U.S. Feb. 19, 1985) (overruling *National League of Cities*). Although *National League of Cities* dealt with traditional state functions protected from federal invasion, the case is significant because it attempts to protect certain traditional functions as essential to the unique identity of states.

"[A] focus on individual rights is hardly inconsistent with a concern for federalism. Since '[t]he federal and state governments are in fact but different agents and trustees of the people, constituted with different power, and designed for different purposes,' 'most of a state's rights' must ultimately be derived from the rights of its citizens." L. TRIBE, *supra* note 3, § 5-21 at 307 (1978) (quoting THE FEDERALIST NO. 46, at 330 (J. Madison) (B. Wright ed. 1951)).

Justice Brennan first acknowledged the connection between the rights of states as states and the vision of federalism guiding the Court's intuition in the state discriminatory legislation cases. Hughes v. Alexandria Scrap, 426 U.S. 794, 822 n.4 (1976) (Brennan, J., dissenting). Referring to the Court's statement that "[n]othing in the purposes animating the Commerce Clause prohibits [discrimination in the form of market participation]," 426 U.S. at 810, Brennan commented: "The absence of any articulated principle leads me to infer that the newly announced 'state sovereignty' doctrine of *National League of Cities v. Usery* . . . is also the motivating rationale behind this holding." *Id.* at 822 n.4 (Brennan, J., dissenting).

[98] The majority of the *Reeves* Court explicitly adopted a notion of "the role of each state 'as guardian and trustee for its people'." Reeves, Inc. v. Stake, 447 U.S. 429, 438 (1980) (quoting from Heim v. McCall, 239 U.S. 175, 191 (1915)) (footnote omitted). The Court refused to strike down South Dakota's residents-first cement policy, stating that "a State's ability to structure relations exclusively with its own citizens" is an "essential . . . purpose of state government—to serve the citizens of the State." *Id.* at 441–42 (footnote omitted). Although the dissenters argued in favor of limiting the market participant/state sovereignty notion to "areas of traditional governmental functions," *Id.* at 449 (Powell, J., dissenting), the majority stated that the state's interest in serving its people with funds encompassed nontraditional governmental functions as well. *Id.* at 438 n.10. Note that the original position theory of federalism is closer to the *Reeves* majority than it is to the *Reeves* dissent, because states have a liberty interest in providing for their citizens beyond the traditional areas of state sovereignty, when a like interest of other states does not circumscribe it.

a principle that allows states to discriminate in favor of their own citizens until that discrimination interferes with the liberty of other states. Such a regime would seem to permit a large degree of conduct we would find intuitively objectionable. For example, it would appear that a state law providing that nonresidents are not entitled to any employment or educational opportunities within the state's boundaries would have to be upheld under this principle; such a rule would not interfere with any other state's liberty to provide like benefits for its own citizens.

The state has a special role as provider for its citizens. While the purpose of providing extensive liberty of action under the first principle is to guarantee that the state can provide for its citizens, an equally important concern is that the citizen, as an individual, have extensive personal liberty. A state in the original position, as guardian of its citizens' desires, would realize that a citizen may not want to stay within that state's boundaries forever. The citizen would want to preserve the opportunity to enjoy the attributes of other states, either temporarily, by traveling, or on a permanent basis. The individual would also wish for a great measure of liberty with respect to other individuals, both within the state and in other states. Consistent with the goals of the state, the individual would want this liberty free of unnecessary restrictions, both in personal and in business relations.

Thus a second principle of federalism would emerge: *each citizen is to have an equal right to the most extensive total system of liberties compatible with a similar system of liberty for all individuals and compatible with the liberty of the states.*[99] Essentially, this principle would mandate that where a state's ability to provide for its citizens necessitates discriminatory action directed against nonresidents, the means used must be those that are least restrictive of individual liberties while still compatible with the state's goals. Thus, where conservation of a resource is found to justify limits on access to that resource, care must be taken that the conservationist measure adopted is not too restrictive of the individual liberties of nonresidents. If the resource depletion is not a direct result of nonresident use

---

[99] The second principle is derivative of the first. The first principle arguably contains the subprinciple that a state has a liberty interest in assuring that its citizens have the maximum liberty possible when they are in other states. Reading the first principle so broadly might lead to confusion; the second principle clearly spells out this derivation.

or if other less restrictive means of curtailing depletion are available, the state should not employ any heavy handed conservation method.

To illustrate the second principle, imagine that New Jersey puts much of its finances into an educational system. New York, on the other hand, spends its money on entertainment and neglects education. So that New Yorkers will not come into New Jersey and take advantage of its better educational system, New Jersey passes a law that only those nonresidents with a Ph.D. may enter the state. This law would be invalid under the second principle. The state can prevent this drain on its educational system by means far less restrictive of nonresidents than prohibiting the entrance of all non-Ph.D.'s. For example, New Jersey would be permitted to charge nonresidents higher tuition than residents for attending its schools because nonresidents would not have contributed tax dollars to the state's educational system. An out-of-state tuition level that is correlated to the additional expense that admitting nonresidents places on the system satisfies the state's interest in providing an educational system for its citizens, while at the same time providing maximum liberty for nonresidents.

The second principle therefore leads in some applications to a less restrictive means test. If a state can maintain its liberty to provide for its citizens by a means less restrictive of the liberties of nonresidents, it must pursue that course.[100]

---

[100] If the second principle is to be applied properly, it is necessary to look beyond the language of the discriminatory state legislation and to determine the independent objectives of the statute. There is a danger that the Court would not look beyond the wording of the statute and conclude that whatever result the statute prescribes must have been "intended" by the state. In the above hypothetical, for example, a court not giving substantive content to the second principle might conclude that the state's objective was to exclude nonresident non-Ph.D.'s. In that case, the second principle would seem to be satisfied—there is no less restrictive means of excluding non-Ph.D.'s than to exclude non-Ph.D.'s—and the statute seems perfectly tailored to its result.

The danger that the Court would look no further than the wording of the statute, and that it would adopt the circular logic whereby the law is seen as its own end, might seem insignificant. In the Court's examination of congressional legislation under the Equal Protection Clause, however, the Court has upon occasion used precisely this approach. In United States R.R. Retirement Bd. v. Fritz, 449 U.S. 166 (1980), the Court was faced with an attack on the validity of the classification of various railroad employees by the Railroad Retirement Act of 1974. The District Court in that case had found that the differentiation in the Act, "based solely on whether an employee was 'active' in the Railroad business as of 1974, was not 'rationally related' to the congressional purposes of insuring the solvency of the Railroad Retirement system and protecting vested benefits." *Id.* at 174. The Supreme Court, applying the rational relation test, reversed, upholding the classifications with the extraordinary finding that "the plain

The application of these two principles yields a coherent theory of federalism with which to analyze cases of state discrimination against nonresidents. The first principle guarantees that the state may act affirmatively to advance its citizens' interests when doing so will not inhibit other states from doing the same. In this respect, the original position theory of federalism would seem to permit states to engage in both regulation and participation. But the second principle's mandate that a state may not choose a means of action that unnecessarily restricts the freedom of nonresidents expresses a strong preference in favor of participation and not regulation. Generally it is less intrusive for a state to enter the market and favor the purchasers and sellers of its choice than it is for the state to dictate to individuals whom they must deal with and how they must do so.

It is important to recognize, however, that market participation will not always be a state's least restrictive means of achieving its goal. There are areas, such as health and welfare, where

---

language of [the section of the Act] marks the beginning and the end of our inquiry." *Id.* at 176.

Although the Court in *Fritz* was applying the rational relation test, if the language of a statute truly "marks the beginning and the end" of the inquiry, it is difficult to see why any statute would fail the more demanding, less restrictive means test derived from the second principle. Every law would be its own perfect fit. Justice Brennan's critique of this approach in his dissent in *Fritz* is equally applicable to any parallel approach to federalism analysis: "[B]y presuming purpose from result, the Court reduces analysis to tautology. It may always be said that [the state] intended to do what it in fact did. If that were the extent of our analysis, we would find every statute, no matter how arbitrary or irrational, perfectly tailored to achieve its purpose." *Id.* at 187 (Brennan, J., dissenting). The second principle, like the Equal Protection Clause, requires the Court to deduce the independent objective of the statute and to analyze whether the discrimination is necessary to the achievement of that objective.

In Edwards v. California, 314 U.S. 160 (1941), the Court faced a California statute making it a misdemeanor for a person to bring into the state any indigent who was not a resident of that state. The purpose and effect of the statute was to exclude paupers, and if the Court ended its inquiry at that point, the California statute would have to be upheld as impeccably suited to its end. Rather, the Court recognized that California could not address the problem of "health, morals, and especially finance" in a manner that placed the burden of poverty on other states. *Id.* at 173.

The Court in *Edwards* recognized another limit on the state's right to pass discriminatory legislation. That limit is the premise of the states in the original position—that they have joined together to achieve what they could not achieve alone, and have agreed to share one another's fate. As the Court in *Edwards* stated:

> It is frequently the case that a State might gain a momentary respite from the pressure of events by the simple expedient of shutting its gates to the outside world. But in the words of Mr. Justice Cardozo: 'The Constitution was framed under the dominion of a political philosophy less parochial in range. It was framed upon the theory that the peoples of the several states must sink or swim together, and that in the long run prosperity and salvation are in union and not division.' *Baldwin v. Seelig*, 294 U.S. 511, 523.

Id. at 173–74.

market participation may be impractical and regulation may be the only means of achieving the state's goal. In these circumstances, mechanistic application of the market participant/market regulator distinction does not comply with the two principles or the original position theory of federalism.

Overall, the two principles of federalism which would be chosen by states in the original position yield a unified view of federalism. The states band together to accomplish what they could not achieve alone—a greater level of prosperity and security. Agreeing to share the fate of all states, each state considers the natural resources within its boundaries as accidents of nature to be made widely available. Each state is to have an equal right to the most extensive liberty compatible with a similar system of liberty for all. Each citizen is to have an equal right to the most extensive total system of equal liberties compatible with a similar system of liberties for other individuals, and compatible with the liberty of the state. It follows that a state may discriminate against nonresidents only when discrimination is necessary for achievement of its citizens' conception of the good, and then, only by the means least restrictive of the liberty of nonresidents.

This view of federalism places the intuitive grapplings of the Court into a coherent framework. A brief review of the cases illustrates that this view explains the Court's intuitive approach and shapes its distinctions into a coherent theory of federalism.

IV. Analysis of the Cases Under the Original Position Theory of Federalism

The distinction introduced in *Hughes v. Alexandria Scrap Corp.*[101] between a state acting as a market participant and as a market regulator is elucidated by the original position theory of federalism. If a state wishes to facilitate its citizens' desires for an environment free of automobile hulks, it restricts the liberty interests of nonresidents far less if it uses a subsidy rather than a law requiring the processing of hulks in-state. In striking down the state regulation involved in *South-Central Timber Development, Inc. v. Wunnicke,*[102] the Court pointed out the less intrusive nature of a subsidy:

---

[101] *Alexandria Scrap*, 426 U.S. at 808–10. *See supra* text accompanying notes 14–19.
[102] 104 S. Ct. 2237 (1984).

> If the State directly subsidized the timber-processing industry by such an amount, the purchaser would retain the option of taking advantage of the subsidy by processing timber in the State or forgoing the benefits of the subsidy and exporting unprocessed timber. Under the Alaska requirement, however, the choice is made for him: if he buys timber from the State he is not free to take the timber out of State prior to processing.[103]

The above distinction is closely related to the distinction between a resource or commodity owned by the state and one that is privately owned. The original position theory's explanation for this distinction is also quite straightforward. The states have a liberty interest in developing state resources to provide for their citizens.[104] Allowing the state to make use of its resources before it allows them to pass out of its possession is a far less intrusive means of providing for its citizens than a requirement that state resources, once reduced to private possession, must be used in certain ways. It intrudes less upon the interests of the state's own citizens and the interests of the nonresidents who might otherwise be cut off from trade. It was significant, for example, that in *Reeves, Inc. v. State*, South Dakota did not bar resale of cement to out-of-state purchasers.[105] The Court in *South-Central* was correct in asserting that "[t]here are sound reasons for distinguishing between a State's preferring its own residents in the initial disposition of goods, when it is a market participant, and a State's attachment of restrictions on dispositions subsequent to the goods coming to rest in private hands,"[106] even if the Court could not articulate those reasons persuasively. The distinction is sound both because the state has a liberty interest in acting as a trader on behalf of its citizens and because such actions are less intrusive than regulatory means of fulfilling the state's goal.

The question of intrusiveness was close in *White v. Massachusetts Council of Construction Employees* because "[t]he power to dictate to another those with whom he may deal" is viewed with suspicion and "closely limited in the context of purely private economic transactions."[107] Perhaps the reason

---

[103] *Id.* at 2244.

[104] This liberty interest is limited by the fact that resources are viewed as a collective gain for the nation.

[105] *See* 447 U.S. at 429, 435 (1980); *see also infra* notes 17–32 and accompanying text.

[106] *South-Central Timber*, 104 S. Ct. at 2246.

[107] 460 U.S. 204, 219 (1983) (Blackmun, J., concurring in part and dissenting in part).

that the intrusion was allowed in *White* was the Court's sense that use of city funds, although intrusive, was the least intrusive means of satisfying the city's goal.

Although the question was close in *White*, the result in *Camden* should not be different. The facts of the two cases appear similar in all significant respects. In both *White* and *Camden*, the cities imposed restrictions on public construction projects which required that a substantial percentage of the work crews be made up of city residents. The Court in *White* was able to dispose of the challenge to the executive order because the case was analyzed under the Commerce Clause. In *Camden*, however, the ordinance was analyzed under the Privileges and Immunities Clause, and the Court found it necessary to remand for a justification of the discrimination. One of the central premises underlying the original position theory is that a single conceptual framework should be applied in all cases testing the constitutionality of state legislation discriminating against nonresidents. Thus, if the justification for the discrimination in *Camden* is as substantial as it was in *White*, the constitutionality of the ordinance should be upheld on remand.

The Court's confusion in its analysis under the Privileges and Immunities Clause concerning the ownership theory is the result of the tension between the liberty interest of a state in exploiting resources it technically owns and the collective nature of those resources.

The Court in *Baldwin v. Fish and Game Commission of Montana*, as it struggled to come to grips with this tension, expressed itself in perfect accord with the original position analysis: "In more recent years . . . the Court has recognized that the states' interest in regulating and controlling those things they claim to 'own,' including wildlife, is by no means absolute. States may not compel the confinement of the benefits of their resources, even their wildlife, to their own people whenever such hoarding and confinement impedes interstate commerce."[108]

Nonetheless the Court went astray in *Baldwin*; it upheld the discriminatory licensing scheme because hunting elk was not "basic to the maintenance or well-being of the Union."[109] The test for discriminatory legislation should not be whether the compromised interest is fundamental. Rather, the test should be

---

[108] 436 U.S. 371, 385–86 (1978).
[109] *Id.* at 388.

whether the state's discrimination is necessary to provide for its citizens' conception of the good, and if so, whether any other means would be less intrusive of the liberty of others.[110]

It could be argued that if the principles presented often break down into a least restrictive means test, then the state would never be allowed to discriminate, because a less restrictive means is always possible. The answer to this argument is that the Court acts in a world of practical realities when deciding whether other means would be significantly less intrusive. States desirous of avoiding constitutional obstacles will have an incentive to discover the least intrusive means to their goals. Those actions that are challenged before the Court will be evaluated in light of the perceived realities of the particular fact situation.

In *Hughes v. Oklahoma,*[111] the Court struck down an Oklahoma statute that forbade transportation of any commercially significant number of minnows for sale out of the state. The Court found it significant that the state had unnecessarily chosen a highly intrusive means of protecting the fish: "Far from choosing the least discriminatory alternative, Oklahoma has chosen to 'conserve' its minnows in the way that most overtly discriminates against interstate commerce . . . . [N]ondiscriminatory alternatives would seem likely to fulfill the State's purported legitimate local purpose more effectively."[112]

The Court in *Hughes v. Oklahoma* drew on the case law striking down state attempts to hoard natural resources.[113] Quoting at length from *West v. Kansas Natural Gas,*[114] the Court again echoed the philosophy of the original position theory of federalism:

> In such commerce, instead of the States, a new power appears and a new welfare, a welfare which transcends that of any State. But rather let us say it is constituted of the welfare of all of the States and that of each State is made greater by a division of its resources, natural and created, with every other State, and those of every other State with it. This was the purpose, as it is the result, of the interstate commerce clause of the Constitution of the United States.[115]

---

[110] Note that this is very close to the dissent's formulation, which requires a justification for discriminatory treatment. *Id.* at 402 (Brennan, J., dissenting).

[111] 441 U.S. 322 (1979).

[112] *Id.* at 337–38 (footnote omitted).

[113] *Id.* at 329–35.

[114] 221 U.S. 229 (1911).

[115] Hughes v. Oklahoma, 441 U.S. at 330 (quoting West v. Kansas Natural Gas, 221 U.S. at 256).

The final distinction made by the Court—between whether the good is a natural resource or is the product of state industry, skill, and foresight—is also rational according to the original position theory. In forming the union, the states agreed to share one another's fate and recognized that natural resources accrue to the benefit of all states. Nevertheless, as we have seen, each state has a liberty interest in developing its resources to provide for its citizens. Once a state has invested its time, energy, and finances in developing a resource, it has a liberty interest in providing for its citizens the fruits of their collective labor. As the Court pointed out in *Reeves,* allowing South Dakota's citizens to enjoy the fruits of the state's work reflected "the essential and patently unobjectionable purpose of State government—to serve the citizens of the State."[116] It was significant in *Reeves* that the state had not sought to hoard its supply of limestone, the prime element of cement.[117] According to the original position theory it would difficult to justify restriction of a resource that, by mere happenstance, is found under one state's ground.[118]

And similar reasoning was employed in *Alexandria Scrap.* The commerce at issue in that case was created by the state's subsidy. As the Court noted, "[w]e would hesitate to hold that the Commerce Clause forbids State action reducing or eliminating a flow of commerce dependent for its existence upon state subsidy instead of private market forces."[119] Similarly, in *Baldwin* the Court made frequent reference to the fact that the elk population would not exist at all were it not for the efforts of the state.[120]

V. CONCLUSION

The conception of federalism developed through the use of the original position analysis grants states a large measure of

---

[116] *Reeves,* 447 U.S. at 442 (footnote omitted).

[117] *See id.* at 444.

[118] "In justice as fairness men agree to share one another's fate. In designing institutions they undertake to avail themselves of the accidents of nature and social circumstance only when doing so is for the common benefit." J. RAWLS, *supra* note 11, at 102.

[119] *Hughes,* 426 U.S. at 809 n.18.

[120] *Baldwin,* 436 U.S. at 375–77, 388–89 (1978). "Montana's elk should be treated constitutionally as if they were the property of the state because the elk, and the opportunity to hunt elk, would not exist but for the voluntary undertaking by the citizens of Montana to preserve these animals." L. TRIBE, AMERICAN CONSTITUTIONAL LAW 38 (Supp. 1979).

freedom in their role as providers for their citizens. The means of exercising this liberty, however, are carefully limited; a state may not discriminate against nonresidents unless such discrimination is the only way to achieve the state's goal; nor may a state discriminate against nonresidents in excess of that necessary to achieve its goals. That the state's ability to discriminate to the extent necessary to achieve its citizens' conception of the good is circumscribed by its membership in the union does not detract from its role as a provider. Rather, this limitation expresses the state's conviction that by joining the union's enterprise, and by sharing its common wealth, the state may better provide for its citizens' conception of the good.

Seen in this light, the Court's distinctions between market participant and market regulator, resource and product, and state ownership and private ownership are useful indicators of whether state action is more intrusive than necessary for achievement of the state's goal. But these distinctions cannot be applied mechanistically if mutual friendship and intercourse are to be achieved. Characterization of the state as a market participant cannot end the inquiry under the Commerce Clause; instead it should inform the Court that a state's action may be one of those actions minimally intrusive of nonresidents' rights.

The fundamental rights analysis of when to apply the Privileges and Immunities Clause is not the proper approach. As the original position conception of federalism illustrates, it is not the nature of the right that prompts the scrutiny of the Privileges and Immunities Clause, but the gratuitous restriction of any right of a nonresident. It would be ironic that at a time when the Court has eschewed an analysis of the "fundamental element[s] of state sovereignity"[121] because of the "elusiveness of objective criteria"[122] in determining which state activities are fundamental, it would continue to embrace a fundamental rights approach under the Privileges and Immunities Clause.

The original position theory of federalism points out the difficulty with relying on distinctions that immunize potentially disruptive state behavior from the scrutiny of the Commerce Clause and the Privileges and Immunities Clause. As the Court stated in *Garcia*:

---

[121] Garcia v. San Antonio Metropolitan Transit Authority, 53 U.S.L.W. 4135, 4140 (U.S. Feb. 19, 1985).

[122] *Id.*

> The problem is that neither the governmental/proprietary distinction nor any other that purports to separate out important governmental functions can be faithful to the role of federalism in a democratic society. The essence of our federal system is that within the realm of authority left open to them under the Constitution, the states must be equally free to engage in any activity that their citizens choose for the common weal, no matter how unorthodox or unnecessary anyone else—including the judiciary—deems state involvement to be.[123]

The solution must be to step back and take a broad view of federalism. The conception here of states in the original position, like the conception of individuals according to Rawls, is "an intuitive notion that suggests its own elaboration, so that led on by it we are drawn to define more clearly the standpoint from which we can best interpret . . . [federal] relationships. We need a conception that enables us to envision our objective from afar . . . ."[124] The original position theory of federalism supplies that vision, and may be the key to defusing threats to interstate harmony in whatever form they arise.

---

[123] *Id.*
[124] J. RAWLS, *supra* note 11, at 21–22.

RECENT PUBLICATIONS

THE GOVERNMENT/PRESS CONNECTION: PRESS OFFICERS AND THEIR OFFICES. By *Stephen Hess*. Washington, D.C.: The Brookings Institution, 1984. Pp. xv, 160, appendices, notes, index. $28.95 cloth, $9.95 paper.

The American public has passionate feelings about its press. Members of the press are both intruders and companions in our daily lives. Whether we revere or despise the fourth estate, we are unquestionably dependent upon it for truth, protection, affirmation, inspiration, and even humor. We feel about the press much the same way we feel about government. Stephen Hess has again chosen a subject that is timely and compelling in his latest work in a series of studies involving the media.

The author's work is a compact and informative account of a year spent observing the government press operations at the White House, the State Department, the Department of Defense, the Department of Transportation, and the Food and Drug Administration. He describes the political considerations involved, the multitude of interactions, and the power dynamics of the Washington news panorama. Hess introduces the reader to the hierarchy of government officials that has evolved to guide the press corps through the bureaucracy that is government. Career press officers of major agencies and the political appointees to whom they report—press secretaries—are translators as well as information disseminators (p. 2). It is in the act of translating that the free flow of information may be disrupted and the reportorial function sabotaged. Because Hess presumes that the press is a public policy institution, he argues that it should be scrutinized no less than the executive, the legislature, or the judiciary (p. 5). Like other institutions, then, its product is influenced by faceless actors, including those who function from the government press office.

The importance of that function seems to surprise Hess himself. His conclusions confirm the 1973 study conducted by Leon V. Sigal that most Washington news comes from routine government news releases and briefings, surpassing channels such as leaks as an information source by two to one (p. 5). Moreover, that peculiar brand of newsperson, the press officer, is most often a competent and hardworking individual, who is dedicated to the needs of reporters (p. 108). The sociological

lifestyle of this type of person is among the most interesting discussions of the book, for it reveals the glamourless side of news development and the individuals who endure its long hours for scant recognition. If press officers tend not to acknowledge the political sensitivity of newsgathering, it is because press secretaries—politically appointed superiors—handle events that the administration deems to be sensitive (p. 28–29). Hess describes the press secretary as an advocate of the agency, reporting to other political appointees who must answer to the President (p. 23).

Although Hess's description of press officers is interesting, it seems that his chief area of concern focuses on those instances where politics impinges on news coverage. This is the essence of the "connection." While the title suggests that press officers have unrecognized influence over the creation and dissemination of the news, Hess concludes that their influence is limited to routine matters that do not concern media critics (p. 37). Press officers serve as an essential information link for reporters, but they are not powerful or skillful enough to manage, to manipulate, or to control the news (p. 108). Who, if anyone, is powerful or skillful enough? For that answer we look not to the relationship of reporter to press officer, but primarily to the nature of "controlled combat," which characterizes the interactions of agency advocates and media representatives (p. 23).

Hess observed different agencies and their press operations under the Reagan administration for one year beginning in September 1981. He establishes from the outset that each agency differs in the way it copes with the press's needs. Generally, however, he noted that press spokesmen are "affable good ol' boys" who are able to communicate to reporters that they enjoy the spirit of the game (pp. 19, 21). The press secretary must be likeable, but to maintain his position in the administration, he must also have the ability to control a given situation. Because reporters recognize this, an etiquette has developed, which regulates the relationship between the secretary and the reporters. Hess explains the spectrum of acceptable forms of lying and evasion, the conventions of attribution, and the reporters' struggle to "unravel half-truths" (pp. 23–26). These rules are part of a long-standing tradition and seem to give method to the madcap tournament of briefings, press conferences, and interviews.

But do the rules give method to this madness? Hess hedges. That news will be generated is doubtless; but the system renders

the quality of the news unpredictable (p. 69). This result is inevitable when the government seeks to control the media and where media attempts to maximize its access to information through formal sources, such as public service organizations and think tanks, or through informal means, such as inference, nuance, or leak.

Hess is not sufficiently critical of this catch-as-catch-can approach to obtaining the big news story. He sees the process as reflective of a pluralistic society where separation of powers produces competitiveness and occasional excesses, but where individuals basically have the same stake in society (p. 111). He concludes that no press office or political office manages the news (p. 109). They help to present information in an orderly manner and are a vital part of a free society and the relationship between state and citizen (p. 115).

More germane to the concerns of the public are the policy ramifications of news coverage. Can and does the press force the government's hand on sensitive issues—or even alter its course? To what extent does the government guide the news? These questions are provocatively raised, but not sufficiently explored. Part and parcel of examining the question whether the government manages the news is whether the media influences government policymaking in a manner distinguishable from public opinion. These inquiries are less systemic in nature and probably best undertaken with the benefit of greater hindsight. Yet Hess has at his disposal the resources necessary for an informed analysis of the most controversial events of the early Reagan administration and their coverage. Events and issues discussed in the book include the Stockman interview in *The Atlantic*, the misfortunes of Alexander Haig, United States aid to Central America, and the defense budget. But Hess's conclusions focus on agency press operations, and he notes that his observations were conducted in peacetime, when government is more receptive. But the government/press connection raises more provocative questions and demands more profound conclusions.

Perhaps Hess meant only to recount what he witnessed and to be a voice of objective reason in a chorus of inflamed government-media critics. Unfortunately, the effect is that his work has a glossy veneer—superficially attractive, but ultimately dissatisfying. The book succeeds in being informative, but it is less than enlightening. Hess attempts to find a new slant on an old

theme; he focuses on government press officers and their offices. Certainly, this discussion has its place in demonstrating the government/press connection. But Hess fails to pursue fully the essence of this link and misses the mark.

The debate concerning the media and the government continues. Regrettably, it is neither substantially altered nor enhanced by *The Government/Press Connection.*

Cecelie S. Berry

RECONSTRUCTING AMERICAN LAW. By *Bruce A. Ackerman.* Cambridge, Mass.: Harvard University Press, 1984. Pp. viii, 110, index. $16.00 cloth, $6.95 paper.

"When they speak so resolutely of 'public policy' do lawyers have the slightest idea what they are talking about?" (p. 22). In this extended essay, Bruce Ackerman attempts both to answer that question and to provide a new framework for legal discourse that might remedy such a problem.

Ackerman bases his essay on the indubitably correct assertion that the New Deal altered the terms of political discourse in America. The Supreme Court's 1937 sea change reflected the legal system's accommodation to that alteration. The legitimacy of the activist state, as it intervened to alter the results produced by the invisible hand of the market, could no longer be questioned. As Ackerman analogizes, the pre-1937 realm of legal discourse looked like a vast common law sea with a few statutory peaks rising above the waves. These peaks, however, were at risk of slipping back under the waves without notice at any moment. After 1937, that threat evaporated and the statutory peaks began multiplying so fast that it became difficult to catch glimpses of the common law sea (pp. 8–9).

Legal realism provided the new terms of discourse that enabled lawyers to adapt to swimming in this new aquatic environment (p. 17). Ackerman argues that the legal realist tenets of 1) skepticism about abstraction and 2) confidence in the intuitionistic adaptation of legal doctrine enabled the profession to adapt smoothly to the New Deal reconceptualization of American politics (p. 19). Legal realism disaggregated the law: "So long as one could talk about particular contracts without paying explicit fealty to the ideal of Free Contract, . . . the profession might survive the political crisis with its basic discursive equipment intact" (p. 17). Law could continue to function without a

thoroughgoing readjustment in doctrine; piecemeal readjustment was sufficient. As a result, the doctrines of consideration and unconscionability, negligence and strict liability, survive in uneasy coexistence in American law (p. 18).

Ackerman criticizes legal realist lawyers for their failure fully to adapt legal discourse to the new activist political discourse (p. 22). He formulates paradigms of reactive and activist lawyering that emphasize this failure. The elements of the reactive paradigm follow from the assumption (which he terms the reactive constraint) that no legal argument can question the legitimacy of the military, economic, or social structures generated by the invisible hand; the only reactive legal question is whether the challenged action deviates from institutionalized norms (pp. 24–28). The elements of activist lawyering, the lawyering appropriate to an activist political system, follow from a rejection of the reactive constraint and a concomitant willingness to question the legitimacy of social structures (pp. 28–37). Legal realism frustrates the development of activist lawyering because of its focus on individual fact situations, its rejection of overarching doctrine, and its faith in intuitionistic adjudication. He urges lawyers to begin to construct a new form of legal discourse, which he terms "Constructivism," to bring post-1937 power talk to legal discourse (p. 22).

Ackerman focuses on two areas that he considers central to the reconstruction of American law: reconstructing the statement of facts and reconstructing legal values. In an activist legal system, the statement of facts must not only include information about the particular dispute but information about the broader social background against which this legal dispute occurs (p. 66). If the legitimacy of military, economic, and social structures is open to challenge, relevant information about those structures must be provided for the fact-finder. These newly relevant facts would apparently be provided by welfare economics (pp. 69–70). The key concept of welfare economics is pareto optimality, a condition that is attained when no action can be taken that would make anyone better off without making someone else worse off. Pareto optimality is thus an economic conceptualization of efficiency. Ackerman stresses the ability of welfare economics to provide the facts activist lawyering needs by highlighting market failures and the effects of externalities. This emphasis differentiates his efforts from those of the law and economics school, particularly those from the University

of Chicago, which routinely presume that perfect markets exist (pp. 64–65). Ackerman argues that we must go beyond pure efficiency judgments and also consider normative values in order to create an activist legal system (pp. 78–79).

Ackerman then moves to the second Constructive task—reconstructing legal values. He criticizes welfare economists for not attempting to push beyond positivistic conclusions to normative issues (pp. 80–88). As the welfare economists themselves might well respond, they are hardly qualified to do so. It is equally unclear that lawyers are particularly qualified to accomplish this task, but Ackerman assumes that they are.

Ackerman's Constructive task may be summarized then as follows (pp. 97–101). Reflecting upon the appropriate structure of ongoing legal conversation is the greatest task of the Constructive lawyer, because Americans rely upon legally constrained conversation to articulate their collective rights and duties to each other. The two main features of the ongoing Constructivist legal dialogue are 1) comprehensive legal questioning of the sanctity of the current distribution of power (which Ackerman terms a generalization of procedural due process) and 2) assuring that a condition of undominated equality exists by constraining comprehensive legal questioning with two principles of Neutrality.[1] These two principles 1) forbid citizens from justifying their legal rights by asserting possession of an intrinsically superior insight into the moral universe and 2) forbid legal recognition of any right necessarily based on claims by its holders that they are intrinsically superior to their fellow citizens. Only after undominated equality is assured can the lawyer concern himself with issues of economic efficiency (p. 99).

Ackerman makes a strongly convincing argument that lawyers in the activist state cannot afford to ignore economics (pp. 65–70). The hordes of business school and public policy school graduates now being produced by American universities seem incapable of looking beyond the figures on their computer printouts to the normative issue. Someone needs to do so, or the Chicago school's utilitarian approach to law and economics will succeed, and efficiency will be valued above all else. But law-

[1] These principles are drawn from Ackerman's political philosophy treatise, B. ACKERMAN, SOCIAL JUSTICE IN THE LIBERAL STATE (1980).

yers must understand the efficiency arguments and the computer printouts in order to put them into a proper perspective. Otherwise, the opinions and ideas of lawyers risk becoming less relevant in the ongoing political discourse. Legal education must play a role in teaching lawyers these skills, and the case method is hardly an appropriate way to teach lawyers all the necessary social and scientific skills (pp. 68–69).

This message is important, yet it is not unique to the legal field. Indeed the tenets of Constructivism generally treat legal discourse as indistinct from political discourse. The activist lawyering task—to critique existing power structures, and to evaluate those institutions both in terms of economic efficiency and equity—is precisely the task of political institutions and political discourse in the activist state. Should legal institutions and legal discourse perform these same functions? Ackerman does not raise this basic question.

Illustrative of how Ackerman conflates the political and legal tasks is his discussion of the warrant of habitability (pp. 73–75). He identifies the following issues as the "decisive" legal questions:

> (w)hether, *ex ante*, the warranty will benefit slum dwellers as a class; and if so, whether egalitarian redistribution is an affirmative value in our legal system; and if so, whether it is fairer to pursue these egalitarian objectives by reshaping common law forms like warranty, or by devising new public law forms like the negative income tax, or both (p. 74).

Ackerman notes the acceptance of the legitimacy of the legislature's decisions, which is inherent in the 1937 transformation of law (p. 79), yet he fails to come to grips with the institutional restraints implicit in that recognition. The courts cannot be reaching for their own balances of economic efficiency and equity and still be faithful to the legislature's efforts to complete the same task. Ackerman's Constructivism ignores these substantial restraints upon legal discourse.

The other central problem with Ackerman's approach is the continued primacy of efficiency. He rejects the ability of efficiency to explain our legal system, and he disclaims that efficiency will always be universally endorsed in a Constructive world (pp. 92, 100). He notes the vacuity of an approach which is typified by Richard Posner's argument that slavery is to be condemned only because the dollar value of our labor as free persons is greater than the dollar value of our labor as slaves

(p. 91). Yet, as in Guido Calabresi's analysis of tort law[2] (a work cited as an extant example of Constructive legal scholarship, (pp. 57 n.11, 60 n.15)), equity forms only a background constraint upon the Constructive results achieved through the use of pareto optimal comparisons of efficiency (pp. 69–70). Ackerman criticizes the technocrats and those who lionize the great god Efficiency as simplistic, but he merely notes that, in reality, the economic answers are much more complex than they acknowledge (p. 45). Merely refining the economics used by lawyers will not produce justice.

Ackerman focuses upon refining the economics used in legal discourse but devotes little attention to defining the values of equality that should inform the law. Perhaps such a discussion is beyond the scope of Ackerman's essay, but mankind's spirituality and individual dignity are trivialized in the crunch of numbers generated by the search for efficiency. Unless Ackerman makes clear that such equality concerns predominate over the concern for efficiency and the resulting crunch of numbers, it is unclear that Constructivism would properly place economic efficiency among legal values or lead to the achievement of a more just society.

Eric T. Secoy

THE NEW AMERICAN POVERTY. By *Michael Harrington*. New York, N.Y.: Holt, Rinehart and Winston, 1984. Pp. xii, 271, index. $17.95 cloth.

In undertaking to discuss poverty in the United States, Michael Harrington admits, "It is . . . with some trepidation that I set foot in an area covered with the scattered remarks of strawmen" (p. 183).

Over twenty years ago Harrington wrote *The Other America*,[1] drawing national attention to the previously unrecognized problem of poverty in a land of plenty and inspiring the Johnson administration's spirited "War on Poverty." His new work, *The New American Poverty*, updates the plight of the poor since the inception of that war in the mid-1960's and explores reasons why, despite gains in some areas, their situation has remained disturbingly unchanged.

[2] G. CALABRESI, THE COSTS OF ACCIDENTS (1970).

[1] M. HARRINGTON, THE OTHER AMERICA (1963).

The central theme of *The New American Poverty* is that "structures of misery" impair the development of a social vision capable of effectively handling "a new poverty much more tenacious than the old" (p. 1). Before the mid-1960's, the United States could have been called the most limited welfare state in the West. With the Johnson administration, America developed into a more progressive welfare state. But poverty has persisted, despite the liberal social programs spawned by the Johnson administration and expanded by subsequent administrations. This has provoked conservatives to argue for those programs' elimination.

Harrington contends that these welfare programs were neither liberal nor comprehensive in their efforts and that they never provided for the "massive investment of billions of dollars in radical innovations that challenged the very structure of power" (p. 15). Harrington denounces them as limited and unthreatening, more a palliative to social unrest than a prescription for a social cure. He blames their failure on the government's lack of a serious and intelligent commitment. The program was oversold and underfinanced. The War on Poverty failed by design (pp. 20–21). He also blames the Vietnam War for subverting the "War on Poverty" almost before it began, distracting concern and energy from the domestic front and fragmenting society as well as its leaders.

Today, the problem of poverty proves more substantial. While the "old" poverty, the poverty of the pre-1970's, had its roots deeply embedded in American soil, the "new" poverty has grown into a global garden. Harrington asserts that "massive international and national trends have created the basis for the new structures of misery in the America of the eighties, for a poverty more difficult to defeat than the indignities of twenty years ago" (p. 11). America's participation in the world economy of the 1970's assured its participation in the imminent worldwide economic crisis. This was brought on by three factors: 1) a transformation of the international division of labor, 2) an unprecedented internalization of capital, and 3) a technological revolution (p. 239).

Harrington describes the impact of these forces on the social and economic structure of American society, delineating and defining the links that relate foreign production and international markets to America's socio-economic crisis (p. 11). At times, the connections are direct and apparent and involve such issues

as international commerce or changes in modes of production. At other times, the links are tenuous or seemingly unconnected, such as a "decline in the capacity for compassion, which is a by-product of the economic insecurities of the last decade" (p. 12). These insecurities include troubles in foreign countries, which send waves of undocumented immigrants into the American job market, and the deinstitutionalization of mental patients that releases the homeless into urban centers (pp. 100, 165).

Regardless of which arguments appear to be more credible, the effects on American society have been devastating. Members of the struggling working and middle classes, the "limbo people," frequently drop out of the middle and join the ranks of the impoverished, "which contracts the market and therefore causes more workers, who had been producing for that middle, to drop out" (p. 47). Harrington cites unemployment as the nation's most serious malady, affecting not only the jobless, but perpetuating "a vicious downward spiral for the society as a whole" (p. 47).

In addition to the "traditional" poor, those who never made it out of their enduring deprivation, *The New American Poverty* furnishes a bleak catalogue of "limbo people" precariously tottering on the brink of financial catastrophe and poverty. They include unwed mothers and their fatherless children; exploited immigrant laborers in the sweatshops of the Northeast and on the farms of the Southwest; "bag" people; young blacks whose lack of training or education hinders their entrance into the work force; and skilled workers displaced by new technology. These groups live "hand to mouth" without financial security, insurance, or hope for the immediate future (pp. 39–40).

Harrington suggests that there are between forty and fifty million Americans who live in poverty today, which approximates the number of the poor he estimated twenty years ago, although the population has increased since then. The figure, however, "represents an exceedingly modest decline in the percentage of the poor" (p. 88). The same problems of twenty years ago continue to plague today's impoverished: substandard housing, "poor on poor" crime, malnutrition, unemployment, limited opportunities, and frustration.

Harrington does not believe that the situation of the poor will change until the standard of living for the upper classes in society begins to decline (p. 62). "If this were to happen—and if there were political movements to point out the common

problems of different social classes—then it would be possible to talk about a majority movement in the United States" to confront issues of eliminating old "structures of misery" and restructuring the economic, social, and political system of America (p. 62).

Harrington recognizes a current trend in the media and in the Reagan administration to understate the number of poor, thereby erasing poverty with a calculator. The "new Gradgrinds," named for a Dickens character who saw numbers instead of people, have made "some remarkable statistical victories over social ills, which have left reality totally unchanged" (p. 68).

President Reagan's 1981 federal budget cut diminished funds to the poor by sixty percent, which has led Harrington and others to say that the President supports a "war on the poor" (p. 36). While "Reaganomics" adversely affected a number of long-standing and developing social programs, the assault on the welfare state carefully avoided challenging "programs with strong political backing, like social security" (p. 36). He suggests that advances in federal social programs have greatly aided the aged, with the largest grants going to people over age sixty-five (p. 87). Although Harrington contends that "the welfare state in the United States is primarily for people over sixty-five, most of whom are not now, and for a long time have not been poor," he does not demand cuts in social security or other entitlement programs (p. 85). Attacking the fundamental principles of Reaganomics, he insists that spending less on social programs has not increased the productivity that President Reagan suggests. He outlines convincing statistics to support his proposition (p. 88).

Harrington labels the arguments of his theoretical opponents as "simplistic" and those of his allies as "sophisticated," and challenges the conservatives, who he claims like to blame the troubles in American society on the poor and the social programs of the 1960's and 1970's, to defend their assumptions against the reality of his statistics. He relies heavily on statistics to support his premises and points, something for which he criticizes the "new Gradgrinds." As a socialist, Harrington attacks just about every aspect of the System and its "structures of misery," but he never coherently or completely defines these structures or effectively demonstrates how they form an unstable bridge connecting all of the continents.

According to Harrington, "A commitment to a new international economic order is therefore part of the solution to the poverty of Los Angeles and New York and Chicago" (p. 178). He never proposes how this is to be achieved, although he vaguely suggests total employment, intelligent planning, limitations on corporations, and the organization of the poor as means of reaching some socio-economic nirvana. More important, he does not offer a concrete or workable plan of action for his most creative idea in the book, "(a)n intellectual offensive against the conservative disinformation campaign around the War on Poverty" (p. 37).

Despite obvious gaps in some of the arguments and its often loose construction, *The New American Poverty* is a moving and thought provoking documentary on a complex and difficult subject. Unfortunately, Harrington appears to have fallen into his own trap: "The discussion of crime and poverty in the United States is, more often than not, a battle between strawmen. Each side sets up a stunning, even idiotic, simplicity that it imputes to the other side and then shreds it to pieces" (p. 182).

P. Todd Pickens

THE UNITY OF LAW AND MORALITY: A REFUTATION OF LEGAL POSITIVISM. By *M.J. Detmold.* London: Routledge & Kegan Paul, 1984. Pp. xix, 271, bibliography, index. $25.00 cloth.

Legal positivism is the name given to a group of theories about the nature of law. The common battle-cry is "a law's validity is logically separate from its moral content" (pp. 21–22). This obscure motto is an answer to those natural law theorists who state that if a rule is not consistent with certain basic moral principles, it simply is not "law."

Legal positivists try to clarify the discussion about the nature of law and legal duties. They view the assertion of a necessary connection between law and morality as a naive inability to distinguish what is from what ought to be—a mistake that clouds analysis when considering whether a given law is valid and what our legal duties are in a given situation. The positivists believe that once we determine what the law is, then we can turn to our beliefs about the moral value of that law (pp. 22–23).

Michael Detmold seeks to tear down the rigid philosophical distinctions of legal positivism in *The Unity of Law and Moral-*

ity. The book is an attack on legal positivism. While the writings of legal positivists tend to be dry and logically rigorous, Detmold's discussion is a kind of philosophical stream of consciousness, lacking the narrow scope of inquiry usual in legal philosophy. Law is not only connected with morality; Detmold insists that the problems of understanding the nature of law necessarily involve all of the basic problems of philosophy. In elaborating his view of the law, Detmold discusses Kant, Hume, and Wittgenstein, as well as his own ideas on free will, commitment, and love.

For Detmold, the paradigms of legal experience consist of an individual citizen deciding whether to comply with a particular rule and a judge adjudicating a particular case. He believes that the proper focus of legal theory is the understanding of legal rules as reasons for action (pp. 250–53). His attention in the book is thus on how we react to particular facts in particular situations. Because rules deal only with general categories, a rule-bound approach misses the basic experience of individual decision within the world (pp. 257–59): "The primacy of the single case . . . is a fundamental moral notion" (p. 219). Only from the perspective of an individual struggling to interpret and to comply with the rules, can we begin to understand the nature of law.

Detmold elaborates with the example of a judge deciding a capital punishment case (p. 22). The judge refers to the relevant legal sources and concludes that under those standards the defendant should be executed. At the same time, however, this judge believes it would be morally wrong to impose the death sentence in this case. The positivist would tell the judge that "what the law is" will not always coincide with morality, and if the judge feels that she is unable to enforce the law as it is, she should resign. Detmold rejects the positivist's view and argues that the distinction between a law's validity and its moral content necessarily breaks down (pp. 22–27).

The positivist tells the judge to "apply the law." Detmold argues that this simple command ignores difficult conceptual problems (pp. 34–39). We think of laws as saying "you should do *x*" or "if you do not do *x*, you should be punished in the following manner." These are norms—statements of what ought to happen. The problem is in determining what is meant by saying that a norm exists (for example, by saying "norm *z* is the law."). Norms do not exist in the world in a way that we can

point to them. Norms are necessarily subjective: they exist only to the extent that individuals interpret factual situations normatively.

In discussing normativity, Detmold relies heavily on the work of the most influential legal positivist, Hans Kelsen[1] (pp. 32–33, 54–60). Kelsen explained the idea of legal validity by discussing the differences in the way we view the actions of gangsters as contrasted with the way we view the actions of government officials. Their actions in the world may be objectively similar, but we view the officials' actions differently because we accept the legal rules of the society as norms that bind us. The significance of law is that most members of society view the rules normatively or as binding upon them (pp. 58–59).

Detmold contends that there is no difference between accepting a legal rule as normatively binding and accepting some extralegal rule as normatively binding (p. 98). If we accept the rule that death should not be used as a punishment, we are not going to accept death penalty statutes as binding upon us. We would see such a rule as a mere description of how some officials view part of the legal system, or as mere predictions of how legal officials might act.

Detmold's troubled judge, still presiding over the capital case, does not find the positivists' arguments helpful. The judge does not accept the death penalty rule as a binding norm. She cannot escape her judicial problems by seeking legal theories that avoid the problems of subjectivity and normativity. The objective theories describe law as merely another exercise of force or as "mere prophecies of what the judges will do in fact."[2] Yet she obviously cannot decide the case based on her prediction of what she herself will do (p. 250).

Another theorist, H.L.A. Hart, attempted to bridge the gap between the necessarily subjective normative aspect of law and the apparently stable objective nature of legal systems. Hart's "rule of recognition," to simplify his theory a great deal, states that what the law is (what it is that judges ought to apply) can be determined from the consensus of what other legal officials see as the structure and boundaries of the legal system.[3]

---

[1] H. Kelsen, Pure Theory of Law (1967).
[2] Holmes, *Path of the Law*, 10 Harv. L. Rev. 457, 461 (1897)
[3] H.L.A. Hart, The Concept of Law (1961).

As Detmold points out, Hart's theory is just another attempt to make the logically impossible connection between the descriptive and the normative (pp. 58–60). Detmold elaborates his argument by discussing a situation in which a judge presides within a legal system that has a constitution expressly granting the legislature the power to pass a statute that "all blue-eyed babies should be killed" and in which the legislature passes such a statute (p. 220). Detmold's judge admits that the written document known as the constitution includes the baby killing norm and that most other citizens and officials accept the norm as being of constitutional force. Despite this, the judge asserts that these factors are insufficient logically to force the conclusion that the baby killing norm is of constitutional significance and should control the case before her (p. 222).

The positivists' ability to state that the validity or the correct interpretation of the law can be logically separated from morality depends on the assumption that the boundaries of law and of legal systems can be, at least in theory, objectively set. This assumption faces two separate but related difficulties. First, as both Kelsen and Detmold point out, treating a rule as normatively binding is no different for legal rules than it is for extralegal rules, or for rules that do not meet specified procedural criteria. Normativity is subjective and has no necessary connections with objective factors.

Second, no matter how carefully we delineate the objective sources of law, we cannot derive from those objective elements the proper interpretation or the proper limits of the laws. Legal systems, like all sets, are subject to Bertrand Russell's proof that no set is self-contained. Professor Laurence Tribe has concluded that this truth is applicable to the problem of constitutional interpretation,[4] and it is equally applicable to legal theory in general.

Detmold provides thoughtful criticisms of the various positivist theories, but he does not seem to comprehend the reasons for the structural problems of positivism. Legal positivism is a descriptive theory. The limitations of positivism are the limitations of all theories that merely try to describe human behavior.

These limitations are shown in the way that a legal positivist's critique of natural law theory can be used against legal positiv-

---

[4] L. Tribe, Constitution as Point of View 3–4 (Apr. 1981) (unpublished manuscript on file at Prof. Tribe's Harvard Law School office).

ism itself. An early legal positivist, John Austin, set forth the following critique of natural law theories: if I am convicted of violating this rule that lacks the requisite moral content, and if I am hanged for my violation, I am no less dead because the rule did not fit the theorist's criteria for "law."[5] The response to Austin logically follows: if I am convicted of violating this rule that lacks the requisite procedural or positive criteria, and if I am hanged for my violation, I am no less dead because the rule did not fit the positivist's criteria for "law." No matter how we try to contain official action within descriptive theory, at some point an action done in the name of "law" will exceed the boundaries of what we consider valid law. At this point, it is a game only for philosophers whether to label this official action "a change in what the law is" or "an action outside the law."

Philosophical theories that seek only to offer analytical labels or logical deductions from legal behavior deal with only a small portion of the legal issues that trouble us. Legal positivism does not purport to offer us psychological or sociological truths about our behavior, nor does it attempt to establish a moral framework within which we can analyze the legitimacy of various uses of force. A theory like that of H.L.A. Hart, however, for all of its limitations and occasional lapses in logic, does accurately reflect the way most citizens view the law.

By contrast, Detmold's approach to the law fails to reflect the way most people view the law or act within it. Whatever the logical problems with their analyses, most judges and citizens do distinguish between legal reasons for action and moral reasons for action. In most countries there is a consensus in the way people describe, apply, and comply with legal rules. This consensus is not explained by the extreme intuitionism of Detmold's theory of law.

Brian S. Bix

PATHS OF NEIGHBORHOOD CHANGE: RACE AND CRIME IN URBAN AMERICA. By *Richard P. Taub, D. Garth Taylor & Jan D. Dunham*. Chicago, Ill.: University of Chicago Press, 1984. Pp. xii, 253, appendices, bibliography, notes, index. $25.00 cloth.

Inadequate and dilapidated housing, racial imbalances, and crime are problems afflicting almost every American city. The

[5] J. AUSTIN, THE PROVINCE OF JURISPRUDENCE DETERMINED 185 (Library of Ideas ed. 1954).

need to understand and to address the causes of urban decline is clear, but the task is difficult. The interaction of economic and sociological forces often leads urban theorists to confuse cause and effect, or to oversimplify in an attempt to avoid incomprehensibly complex models. In *Paths of Neighborhood Change*, Richard P. Taub, D. Garth Taylor, and Jan D. Dunham examine and try to refine current theoretical approaches to these issues by focusing on the process of change in eight Chicago neighborhoods. They then use the results of this empirical and theoretical study to support a number of policy recommendations.

The empirical portion of the study consists of a random telephone survey of approximately 400 people in each neighborhood examined. The surveys are supplemented by participant observations, local news sources, and archival data. The authors devote three chapters to the presentation of the information gathered from these studies (p. 18). In addition to statistics relating to racial concentration, property values, and crime rates, the authors present measures of individual satisfaction and neighborhood condition, stability, and cohesion. While much of this data is eventually incorporated into the theoretical discussion of chapters six through eight, a good deal of information is presented without precise explanation of its theoretical relevance or statistical significance.

According to the authors' review of the literature that has developed with regard to racial change in urban neighborhoods, most theorists believe that once substantial numbers of blacks enter a neighborhood (a process referred to as "invasion"), the proportion of blacks will continue to rise and will not level off until racial change has been complete ("succession") (p. 5). In response to this belief the authors look to their own empirical findings and report that this process of complete racial succession is in fact not inevitable. While there are neighborhoods in Chicago that are experiencing succession as described above, there are neighborhoods that are resisting invasion and succession, and others that are experiencing invasion without succession. In fact, three of the neighborhoods studied have appreciating property values and a minority population which is both sizeable and stable. Interestingly, two of these neighborhoods are enjoying this success in spite of high crime rates (pp. 24, 31, 89).

Why does it now seem to be possible for integrated neighborhoods to exist when in the 1950's and early 1960's the de-

terministic model of racial succession seemed to be correct? The authors point to a number of factors. First, there is the growth of a substantial black middle class. Second, there have been changes in the attitudes of whites toward minorities. This attitudinal change is accompanied by an increasing willingness to live in neighborhoods in which black families are present. Third, large-scale institutions such as universities, banks, and community organizations have become increasingly sophisticated at managing development and integration. For example, in 1952 the University of Chicago established the South East Chicago Commission to deal with issues of urban renewal in the Hyde Park-Kenwood community (pp. 99–102). The Commission invested capital, lobbied for particular policies, pursued code enforcement, and encouraged entrepreneurial activity. To combat crime and to alleviate community anxiety, the University created a large private security force and took action to establish a visible presence in the area.

The success of such institutional intervention leads the authors to emphasize the tremendous potential of corporate actors to initiate and to sustain racial balance, while maintaining or improving the neighborhood. In seeking to encourage and to guide such intervention the authors point out that successful community organizations generally seek to maintain a stable real estate market, to improve the quality of neighborhood schools, and to establish an anticrime package (pp. 184–86, 192–93). Full-time professional staffs, substantial resources, and "clout" are also important ingredients for success (p. 185).

In addition to an appreciation for the historic and institutional forces that establish the context in which individual decisions are made, the authors also carefully consider the importance of individual decisions as a force in determining the pattern of change in urban neighborhoods. The authors develop theories to explain how individuals decide whether to invest in their homes and whether to move in reaction to increases in the minority population (pp. 119, 142). Perhaps the most interesting aspect of this analysis is the demonstrated relationship between market conditions and the maintenance of racial balance. The authors assert that an excess supply of housing significantly increases the likelihood of complete succession (pp. 148, 158). While the argument is persuasive, the resulting recommendation to reduce the supply of housing in certain neighborhoods is less so.

The theoretical discussions are generally convincing, but the authors seem to have overemphasized the extent to which an investment in home improvement is a function of the expected financial return upon sale of the house. The homeowner's desire to enjoy the improvement herself is probably a far more important motivation. More significant, in the chapter dealing with reactions to changes in the level of racial concentration, the authors fail to discuss the ways in which people's attitudes toward other racial groups are changed by actual contact (p. 167). The authors do indicate an implicit awareness of this process, however, in their conclusion. They note that "[i]t is the self-conscious attempt to link the stable elements of both races that will help to reduce the fears associated with race and crime" (p. 193). They appear to believe that the interaction between the races can alter people's perceptions and fears and that the nature of this interaction can be affected by policymakers. The understanding evidenced by this passage, taken along with the authors' initial discussion of attitudinal changes as one explanation for the greater integration of neighborhoods, makes it surprising that more attention was not given to this issue.

Taub, Taylor, and Dunham correctly point out that housing questions and poverty questions are often confused and that "the provision of housing for the poor is a different problem from either the improvement of the plight of poor people or the promotion of solid urban neighborhoods and the improved quality of life that goes with them" (p. 192). Nevertheless, while poverty cannot be effectively fought through housing policies, the authors might have addressed the questions of whether and to what extent housing policies can be advanced through the alleviation of poverty.

Paths of Neighborhood Change offers no great theoretical contributions and provides no easily applicable or innovative prescription for public policy. But in its eclectic empirical approach and its incorporation of a number of variables often ignored by other urban theorists, the book has provided some insight into the process of urban development. It has established a method of analysis that could prove to be helpful in addressing a variety of issues in urban planning.

David J. Goldstein

Index to the Harvard Journal on Legislation

Volume 22

SUBJECT INDEX

BANKING LAW

The Proposed Uniform New Payments Code: Allocation of Losses Resulting From Forged Drawers' Signatures, *Steven B. Dow & Nan S. Ellis* 399

BANKRUPTCY

Scrambling to Define Bankruptcy Jurisdiction: The Chief Justice, the Judicial Conference, and the Legislative Process, *Vern Countryman* 1

COMMERCIAL LAW

The Proposed Uniform New Payments Code: Allocation of Losses Resulting From Forged Drawers' Signatures, *Steven B. Dow & Nan S. Ellis* 399

COMMUNICATIONS

Regulating the Interception and Disclosure of Wire, Radio, and Oral Communications: A Case Study of Federal Statutory Antiquation, *Bruce E. Fein* 47

Telephone Pornography: First Amendment Constraints on Shielding Children From Dial-A-Porn, *John C. Cleary* 503

COMMUNITY DEVELOPMENT

Small Steps on the Long Road to Self-Sufficiency for Indian Nations: The Indian Tribal Governmental Tax Status Act of 1982, *Robert A. Williams, Jr.* 335

CONSTITUTIONAL LAW

State Discriminatory Action Against Nonresidents: Using the Original Position Theory as a Framework for Analysis, *Adam B. Schiff* 583

Telephone Pornography: First Amendment Constraints on Shielding Children From Dial-A-Porn, *John C. Cleary* 503

CONSTITUTIONS

A New Generation of State Tax and Expenditure Limitations, *Justin J.T. Hughes & Garth B. Rieman* 269

CRIMINAL JUSTICE

Crimes Against the Unborn: Protecting and Respecting the Potentiality of Human Life, *Jeffrey A. Parness* 97

Understanding Criminal Investigations, *Philip B. Heymann* 315

DEVELOPMENT LAW

See COMMUNITY DEVELOPMENT

ELECTRONIC SURVEILLANCE

Understanding Criminal Investigations, *Philip B. Heymann* 315

EMPLOYMENT

See LABOR LAW

FEDERAL COMMUNICATIONS COMMISSION

See COMMUNICATIONS

FEDERALISM
State Discriminatory Action Against Nonresidents: Using the Original Position Theory as a Framework for Analysis, *Adam B. Schiff*....... 583
FIRST AMENDMENT
See LIBEL, OBSCENITY
FOREIGN RELATIONS
See INTERNATIONAL LAW AND RELATIONS
HEALTH CARE
Home Health Care for the Elderly: Programs, Problems, and Potentials, *Chai R. Feldblum* 193
HUMAN RIGHTS
Action Specific Human Rights Legislation for El Salvador, *Kenneth H. Anderson* 255
INTERNATIONAL LAW AND RELATIONS
Action Specific Human Rights Legislation for El Salvador, *Kenneth H. Anderson* 255
JURIES
Sentencing by Death Qualified Juries and the Right to Jury Nullification, *Bruce C. McCall* 289
LABOR LAW
Union Restrictions on the Right to Resign: A Proposal for a New Reasonableness Test, *Dennis C. Shea* 551
LAW REVISION COMMISSIONS
The Proposed Uniform New Payments Code: Allocation of Losses Resulting From Forged Drawers' Signatures, *Steven B. Dow & Nan S. Ellis* 399
LEGISLATIVE PROCESS
Scrambling to Define Bankruptcy Jurisdiction: The Chief Justice, the Judicial Conference, and the Legislative Process, *Vern Countryman* 1
See also STATUTORY DRAFTING AND INTERPRETATION
LIBEL
Of Things To Come—The Actual Impact of *Herbert v. Lando* and a Proposed National Correction Statute, *Dale M. Cendali*................ 441
MEDICAID AND MEDICARE
See HEALTH CARE
MILITARY
See INTERNATIONAL LAW AND RELATIONS
MUNICIPAL BONDS
Small Steps on the Long Road to Self-Sufficiency for Indian Nations: The Indian Tribal Governmental Tax Status Act of 1982, *Robert A. Williams, Jr.* 335
NEWS MEDIA
See LIBEL
OBSCENITY
Telephone Pornography: First Amendment Constraints on Shielding Children From Dial-A-Porn, *John C. Cleary*............................... 503
POLITICAL THEORY
State Discriminatory Action Against Nonresidents: Using the Original Position Theory as a Framework for Analysis, *Adam B. Schiff*....... 583

PRESS
See LIBEL
REVENUE SHARING, ALLOCATION, AND EXPENDITURE
A New Generation of State Tax and Expenditure Limitations, *Justin J.T. Hughes & Garth B. Rieman* 269
STATES
See CONSTITUTIONS, FEDERALISM
STATUTORY DRAFTING AND INTERPRETATION
Obscene Telephone Calls: An Introduction to the Reading of Statutes, *Reed Dickerson* 173
TAXATION
A New Generation of State Tax and Expenditure Limitations, *Justin J.T. Hughes & Garth B. Rieman* 269
Small Steps on the Long Road to Self-Sufficiency for Indian Nations: The Indian Tribal Governmental Tax Status Act of 1982, *Robert A. Williams, Jr.* 335
TECHNOLOGY
See COMMUNICATIONS
TORTS
Crimes Against the Unborn: Protecting and Respecting the Potentiality of Human Life, *Jeffrey A. Parness* 97
WIRETAPPING
See ELECTRONIC SURVEILLANCE

AUTHOR INDEX

Anderson, Kenneth H., Action Specific Human Rights Legislation for El Salvador 255
Cendali, Dale M., Of Things To Come—The Actual Impact of *Herbert v. Lando* and a Proposed National Correction Statute 441
Cleary, John C., Telephone Pornography: First Amendment Constraints on Shielding Children From Dial-A-Porn 503
Countryman, Vern, Scrambling to Define Bankruptcy Jurisdiction: The Chief Justice, the Judicial Conference, and the Legislative Process 1
Dickerson, Reed, Obscene Telephone Calls: An Introduction to the Reading of Statutes 173
Dow, Steven B. & Nan S. Ellis, The Proposed Uniform New Payments Code: Allocation of Losses Resulting From Forged Drawers' Signatures 399
Ellis, Nan S. & Steven B. Dow, The Proposed Uniform New Payments Code: Allocation of Losses Resulting From Forged Drawers' Signatures 399
Fein, Bruce E., Regulating the Interception and Disclosure of Wire, Radio, and Oral Communications: A Case Study of Federal Statutory Antiquation 47
Feldblum, Chai R., Home Health Care for the Elderly: Programs, Problems, and Potentials 193
Heymann, Philip B., Understanding Criminal Investigations 315
Hughes, Justin J.T. & Garth B. Rieman, A New Generation of State Tax and Expenditure Limitations 269
McCall, Bruce C., Sentencing By Death Qualified Juries and the Right to Jury Nullification 289
Parness, Jeffrey A., Crimes Against the Unborn: Protecting and Respecting the Potentiality of Human Life 97

Rieman, Garth B. & Justin J.T. Hughes, A New Generation of State Tax and Expenditure Limitations 269
Schiff, Adam B., State Discriminatory Action Against Nonresidents: Using the Original Position Theory as a Framework for Analysis 583
Shea, Dennis C., Union Restrictions on the Right to Resign: A Proposal for a New Reasonableness Test 551
Williams, Robert A., Jr., Small Steps on the Long Road to Self-Sufficiency for Indian Nations: The Indian Tribal Governmental Tax Status Act of 1982 335

TITLE INDEX

Action Specific Human Rights Legislation for El Salvador, *Kenneth H. Anderson* 255
Crimes Against the Unborn: Protecting and Respecting the Potentiality of Human Life, *Jeffrey A. Parness* 97
Home Health Care for the Elderly: Programs, Problems, and Potentials, *Chai R. Feldblum* 193
A New Generation of State Tax and Expenditure Limitations, *Justin J.T. Hughes & Garth B. Rieman* 269
Obscene Telephone Calls: An Introduction to the Reading of Statutes, *Reed Dickerson* 173
Of Things To Come—The Actual Impact of *Herbert v. Lando* and a Proposed National Correction Statute, *Dale M. Cendali* 441
The Proposed Uniform New Payments Code: Allocation of Losses Resulting From Forged Drawers' Signatures, *Steven B. Dow & Nan S. Ellis* 399
Regulating the Interception and Disclosure of Wire, Radio, and Oral Communications: A Case Study of Federal Statutory Antiquation, *Bruce E. Fein* 47
Scrambling to Define Bankruptcy Jurisdiction: The Chief Justice, the Judicial Conference, and the Legislative Process, *Vern Countryman* 1
Sentencing By Death Qualified Juries and the Right to Jury Nullification, *Bruce C. McCall* 289
Small Steps on the Long Road to Self-Sufficiency for Indian Nations: The Indian Tribal Governmental Tax Status Act of 1982, *Robert A. Williams, Jr.* .. 335
State Discriminatory Action Against Nonresidents: Using the Original Position Theory as a Framework for Analysis, *Adam B. Schiff* 583
Telephone Pornography: First Amendment Constraints on Shielding Children From Dial-A-Porn, *John C. Cleary* 503
Understanding Criminal Investigations, *Philip B. Heymann* 315
Union Restrictions on the Right to Resign: A Proposal for a New Reasonableness Test, *Dennis C. Shea* 551