NBER Working Paper Series

ADJUDICATION AS A PRIVATE GOOD

William M. Landes and Richard A. Posner
University of Chicago Law School

Working Paper No. 263

National Bureau of Economic Research, Inc.
204 Junipero Serra Boulevard, Stanford, CA 94305

July 1978

This work was supported by a grant to NBER from the National Science Foundation. The research assistance of Sharon Wallace and Donna Patterson, and John Langbein's advice and comments on an earlier draft, are gratefully acknowledged.

NBER Working Paper 263
July 1978

ADJUDICATION AS A PRIVATE GOOD

William M. Landes and Richard A. Posner

Abstract

This paper examines the question whether adjudication can be viewed as a private good, i.e., one whose optimal level will be generated in a free market. Part I focuses on private courts, noting their limitations as institutions for dispute resolution and rule creation but also stressing the important role that the private court, in its various manifestations, has played both historically and today. Part II discusses a recent literature which has argued that the rules generated in the public court system, in areas of the law where the parties to litigation are private individuals or firms and the rules of law are judge-made, are the efficient products of purely private inputs. Our analysis suggests that this literature has overstated the tendency of a common law system to produce efficient rules, although areas can be identified where such a tendency can indeed be predicted on economic grounds. Viewed as a contribution to the emergent literature on the positive economic theory of law, our finding that the public courts do not automatically generate efficient rules is disappointing, since it leaves unexplained the mechanisms by which such rules emerge as they seem to have done in a number of the areas of Anglo-American judge-made law. However, our other major finding, that the practices and law governing private adjudication appear to be strongly influenced by economic considerations and explicable in economic terms, is evidence that economic theory has a major role to play in explaining fundamental features of the legal system.

William M. Landes
University of Chicago Law School
1111 East 60th Street
Chicago, IL 60637

312/753-2434

ADJUDICATION AS A PRIVATE GOOD

William M. Landes and Richard A. Posner*

Introduction

Adjudication is normally regarded as a governmental function and judges as public officials. Even economists who assign a highly limited role to government consider the provision of judicial services an indisputably apt function of government; this was, for example, Adam Smith's view.[1] Few economists (and few lawyers) realize that the provision of judicial services precedes the formation of the state; that many formally public courts long had important characteristics of private institutions (for example, until 1825 English judges were paid out of litigants' fees as well as general tax revenues);[2] and that even today much adjudication is private (commercial arbitration being an important example). Further, most cases, both civil and criminal, in the public courts are settled out of court rather than litigated to judgment, and most of the inputs into the litigation of such cases are private.

---

* Landes is Professor of Economics, University of Chicago Law School, and Research Associate, National Bureau of Economic Research. Posner is Lee and Brena Freeman Professor of Law, University of Chicago Law School, and Research Associate, National Bureau of Economic Research. The research assistance of Sharon Wallace and Donna Patterson, and John Langbein's advice and comments on an earlier draft, are gratefully acknowledged. This paper continues our study of the creation of rules (precedents) in systems of adjudication, which is supported by the Law and Economics Program of the University of Chicago Law School, and by the National Bureau of Economic Research under a grant from the National Science Foundation for research in law and economics. An earlier draft of this paper was given at a conference sponsored by the Liberty Fund, Inc., and by the Law and Economics Center of the University of Miami Law School. We are grateful for the support of the Fund and the Center and for the advice of the Center's Director, Henry G. Manne.

1. See Adam Smith, The Wealth of Nations 231 (vol. II) (Edwin Cannan ed. 1976).

2. See Brian Abel-Smith & Robert Stevens, Lawyers and the Courts: A Sociological Study of the English Legal System 1750-1965, at 38 (1967).

This paper examines from an economic standpoint the operation of private judicial systems.[3] We do not stop with private court systems, however, since some of the concepts we develop in our examination of such systems, such as that of the competitive provision of judicial services, have interesting counterparts in public court systems and since the outcomes of public adjudication may, according to some recent theories, be privately determined.

Part I of the paper analyzes the demand for and the supply of judicial services, applies this analysis to the judicial systems of primitive societies and to commercial arbitration (both being examples of the private provision of judicial services) and examines judicial competition--both between private and public systems and among public systems (as under the diversity jurisdiction of the federal courts). Private nonjudicial substitutes for adjudication (such as liquidated-damages clauses, a substitute for judicial damage assessment) are also examined. Part II of the paper analyzes the private determinants of public judicial outcomes. Building on recent articles by Priest, Rubin, and Goodman, we examine the circumstances that may lead a public judicial system to bring about results similar to those of a private market.

I. The Market for Judicial Services

A. Economic Model

1. Introduction. A court system (public or private) produces two types of service. One is dispute resolution--determining whether a rule has been violated. The other is rule formulation--creating rules of law as a by-product of the dispute-settlement process. When a court resolves a dispute, its resolution, especially if embodied in a written opinion, provided information regarding the likely

---

3. We are not aware of any previous economic analysis of this subject, apart from a brief discussion in Gordon Tullock, The Pure Theory of Legal Procedure ch. VIII (mimeo. 1977).

outcome of similar disputes in the future. This is the system of precedent which is so important in the Anglo-American legal system although less so in other legal systems.

Both judicial services--dispute resolution and rule creation--are more accurately described as intermediate goods (inputs) than as final goods. Dispute resolution is not a good in itself but an input into compliance with socially desired standards of behavior. Rule creation is not desired in itself either but is a means of particularizing the standards of socially desired behavior in order to promote compliance with them. For the present, however, it will be more convenient to regard dispute resolution and rule formation as the final products of a judicial system rather than as an input into the real final product--which is right behavior.

The two judicial services are in principle severable and in practice often severed. Jury verdicts resolve disputes but do not create precedents. Legislatures create rules of law but do not resolve disputes. In the Anglo-American legal system rule formation is a function shared by legislatures and (especially appellate) courts; elsewhere judicial law-making tends to be less important.

2. Dispute Resolution. Imagine a purely private market in judicial services. People would offer their services as judges, and disputants would select the judge whom they found mutually most agreeable. The most popular judges would charge the highest fees and competition among judges would yield the optimum amount and quality of judicial services at minimum social cost. This competitive process would produce judges who were not only competent but also impartial--and would thus fulfill the ideals of procedural justice--because a judge who was not regarded as impartial could not get disputes submitted to him for resolution. One party would always refuse.

A voluntary system of dispute resolution does not presuppose that the dispute has arisen from a consensual relationship (landlord-tenant, employer-employee, seller-buyer, etc.) in which the method of dispute resolution is agreed on before

the dispute arose. All that is necessary is that when a dispute does arise, the parties to it choose a judge to resolve it. Though complete strangers, as in the typical accident case, the parties can still choose a judge to determine liability.

Although dispute resolution could thus be provided (for criminal as well as civil cases, be it noted) in a market that would operate free from any obvious elements of monopoly, externality, or other sources of "market failure," it may not be efficient to banish public intervention entirely. Public intervention may be required (1) to ensure compliance with the (private) judge's decision and (2) to compel submission of the dispute to adjudication in the first place. The first of these public functions is straightforward, and no more compromises the private nature of the adjudication system described above than the law of trespass compromises the private property rights system. The second function, compelling submission of the dispute to judge, is more complex. If A accuses B of breach of contract, the next step in a system of private adjudication is for the parties to select a judge. But suppose B, knowing that any impartial judge would convict him, drags his feet in agreeing to select a judge who will hear the case, rejecting name after name submitted by A for his consideration. Although a sanction for this kind of foot-dragging akin to the remedies the National Labor Relations Board provides against refusals to bargain collectively in good faith is conceivable, there may be serious difficulty in determining when the bargaining over the choice of the judge is in bad faith--it is not bad faith, for example, to reject a series of unreasonable suggestions by the other side.

Two ways of overcoming the submission problem come immediately to mind. The first is for the parties to agree on the judge (or on the method of selecting him) before the dispute arises, as is done in contracts with arbitration clauses. However, this solution is available only where the dispute arises from a preexisting voluntary relationship between the parties; the typical tort or crime does not.

The second solution is to randomize the choice of the judge; then the parties do not have to negotiate over his selection. But randomization undermines the competitive process whereby the optimal level of judicial services is encouraged. With random selection even an incompetent judge can anticipate an income as a result of the process by which judges are chosen to hear cases. And there is still the problem of compelling the party who fears the outcome of impartial adjudication to submit to the random-selection process.

Another type of private solution to the problem of enforcement and selection of private judge is available when both parties to the dispute are members of the same (private) group or association. The group can expel any member who unreasonably refuses to submit to an impartial adjudication (perhaps by a judge selected by the group) or to abide by the judge's decision. To the extent that membership in the group confers a value over and above alternative opportunities, members will have incentives to bargain in good faith over selection of the judge and to abide by his decision. In these circumstances dispute resolution can operate effectively without public intervention.[4]

This solution is an instance of the Becker-Stigler model of employer control over employee malfeasance.[5] The sanction for malfeasance is assumed to be dismissal (expulsion) and the employer pays a salary above the competitive level for the position in order to make dismissal a costly sanction to the employee.[6]

---

4. A qualification is necessary, however: there is a danger of a faction's seizing control of the association and using its control to expropriate the rents of the other members.

5. See Gary S. Becker & George J. Stigler, Law Enforcement, Malfeasance, and Compensation of Enforcers, 3 J. Legal Studies 1, 6-13 (1974).

6. Monopoly profits to employees can be prevented by, for example, auctioning off positions.

Obedience to the employee's contractual obligations is thus secured without public intervention.[7]

3. Rule Production. Private production of rules or precedents involves two problems. The first is that because of the difficulty of establishing property rights in a precedent, private judges may have little incentive to produce precedents. They will strive for a fair result between the parties in order to preserve a reputation for impartiality, but why should they make any effort to explain the result in a way that would provide guidance for future parties? To do so would be to confer an external, an uncompensated, benefit, not only on future parties but also on competing judges. If anything, judges might deliberately avoid explaining their results because the demand for their services would be reduced by rules that, by clarifying the meaning of the law, reduced the incidence of disputes. Yet despite all this, private judges just might produce precedents. We said earlier that competitive private judges would strive for a reputation for competence and impartiality. One method of obtaining such a reputation is to give reasons for a decision that convince the disputants and the public that the judge is competent and impartial. Competition could lead private judges to issue formal or informal "opinions" declaring their interpretation of the law and these opinions, though intended simply as advertising, would function as precedents, as under a public judicial system.[8] But this scenario is no more than plausible. If there were cheaper methods of advertising one's impartiality as an adjudicator than by writing opinions, those

---

7. Though a residual public machinery is implicit to protect the employee if the employer dismisses him wrongfully.

8. An alternative possibility would be to give judges property rights in precedents--e.g., a royalty every time one of their decisions was cited. See William M. Landes & Richard A. Posner, Legal Precedent: A Theoretical and Empirical Analysis, 19 J. Law & Econ. 249, 272 (1976). But it would be difficult to prevent lawyers and judges from using, without actually citing, prior decisions, especially in out-of-court settlement negotiations by which most legal disputes are terminated.

methods would be chosen and precedents not produced. Some evidence that this may be the case is presented later.

The second problem with a free market in precedent production is that of inconsistent precedents which could destroy the value of a precedent system in guiding behavior. If there are many judges, there are likely to be a bewildering profusion of precedents and there is no obvious method of harmonizing them. An individual contemplating some activity will have difficulty discovering its legal consequences because they will depend on who decides any dispute arising out of the activity. Stated otherwise, there would appear to be substantial, indeed overwhelming, economies of standardization in the precedent market, akin to those that have given us standard dimensions for electrical sockets and railroad gauges. While many industries have achieved standardization without monopoly, it is unclear how to achieve standardization or commonality in the precedent market without a single source for precedent production,without, that is to say, a monopoly. We find it hard to visualize a competitive process by which the precedents of competing judges would converge to a social optimum. Suppose the socially optimal rule for internalizing the costs of sonic boom is to make the airplane owner strictly liable for sonic-boom damage. Any judge who adopted such a rule could be certain to have no sonic-boom business, for the airplane owner would never submit to the jurisdiction of such a judge. The problem is that the private benefit of the rule to a "customer" whose agreement to purchase judicial services from the judge is essential to the judge's ability to sell his services diverges radically from the social benefit of the rule.

A related problem is that a system of voluntary adjudication is strongly biased against the creation of precise rules of any sort. Any rule that clearly indicates how a judge is likely to decide a case will assure that no disputes subject to the rule are submitted to that judge since one party will know that it

will lose. Judges will tend to promulgate vague standards which give each party to a dispute a fighting chance.

This problem disappears if the parties agree in advance to the submission of any disputes arising to a particular judge applying a known set of rules. It can also be overcome, in the association setting, simply by the association's monopolizing the production of the relevant rules or precedents. It is overcome in the traditional family by the monopoly of authority enjoyed by the head of the family.

Of course, without a rigorous empirical study of the costs of public dispute resolution and precedent production relative to those of private provision of these services, one cannot conclude that private provision, with all its problems, is less efficient than public. However, one can conclude that, outside of the association setting (an important qualification), a private market is more likely to emerge in dispute resolution than in rule creation; we shall see this hypothesis corroborated later on when we examine private adjudication in the real world.

The distinction just noted suggests the possibility of severing the two judicial functions and committing the rulemaking function to a public body, a legislature, that does not engage in dispute resolution, and the dispute-resolution function to private judges. But if there are economies of scope that enable the functions to be provided more cheaply by one than by separate "firms," this would be an inefficient solution. The literature on the common law system suggests that there may be such economies, which may explain why, in the Anglo-American system, the same public officials--the judges--are vested with both dispute-resolution and precedent-production functions.

4. The Financing of Judicial Services. That unlimited competition in the provision of judicial services would be inefficient does not necessarily imply,

of course, that a public monopoly is the most efficient method of providing these services. Various intermediary possibilities come to mind, including competitive bidding for a judicial monopoly along lines once suggested in another context by Harold Demsetz.[9] We shall not consider these possibilities here, but will examine the related question of how to finance publicly provided judicial services.

Ronald Coase has shown that even if (because of free-rider problems), compulsion is required for the optimal provision of lighthouse services, it does not follow that such services must or should be financed out of general tax revenues.[10] An alternative is to finance them out of user fees. This was indeed done in the early history of the English lighthouses. The analogous choice in the judicial arena is between paying the judges out of general tax revenues and paying them out of litigants' fees. The choice of the latter method is, however, less attractive in the judicial than in the lighthouse context. The social benefits of lighthouses are limited to the shipping industry, which paid the lighthouse fees. But much of the social benefit of litigation, viewed as a rule-creating activity, is received by people who may never be involved in any litigation. The existence of this external benefit may justify externalizing some of the costs of litigation by financing judges' salaries out of general tax revenues and keeping litigant fees low.

In England, until comparatively recently, judges received not only salaries paid out of general tax revenues but also a portion of the fees charged the litigants in the cases they heard.[11] This is an inefficient method of compensating judges who are engaged in precedent production, so one is not surprised

---

9. Harold Demsetz, Why Regulate Utilities?, 11 J. Law & Econ. 55 (1968).

10. R. H. Coase, The Lighthouse in Economics, 17 J. Law & Econ. 357 (1974).

11. See note 2 supra.

that it has been abandoned (in England and, so far as we know, everywhere else). We noted earlier that dispute resolution is really an intermediate rather than final good, the final good being compliance with proper standards of behavior. The other input into this final good is precedent. The judge who is not paid proportionately to either his final output or his precedent production, but solely according to the number of cases he decides, will have an incentive to overproduce that input. He may write confusing opinions that generate unnecessary disputes; he may create unmeritorious rights; he may even (as we shall see) promulgate rules that discourage the growth of nonjudicial substitutes for judicial dispute resolution.

To be sure, if there were many competing judges the gains to each from deciding cases in a way that increased the demand for judicial services would be small. This is particularly true if any added demand were met simply by appointing new judges. However, as noted earlier, the production of precedents tends to be monopolized (for example, by having a supreme court) in order to avoid the confusion that competitive production of precedents would engender; and a monopolist of precedents who is paid by the number of cases decided will have the undesirable incentives mentioned above.

The foregoing discussion suggests an economic reason why the movement to abolish the payment of litigant fees to judges, coincided, in England at least, with the movement[12] to place decision by precedent on a more systematic basis by regularizing the reporting of judicial opinions. The fee system is efficient only if the judges are engaged in dispute settlement alone. Commercial arbitrators--who, as we shall see, are not engaged in the production of precedents--are still paid on a fee basis and this makes economic sense.

---

12. Described in John P. Dawson, The Oracles of the Law 80-83 (1968).

An alternative, in principle at least, to paying judges flat salaries would be to pay them for their final output. Such a compensation system would require both (1) a dollar-weighted measure of compliance with law and (2) a determination of the individual judge's contribution to compliance through his decisions resolving disputes accurately and creating helpful precedents. The measurement problems created by this system would be intractable, however. They would be severe even if judges were instead paid separately according to both disputes resolved and precedents created. Even though, as we have argued elsewhere, the number of citations in subsequent judicial opinions may be a tolerably accurate proxy for the precedential value of a decision,[13] it would be difficult to attach dollar values to citations so as to weight them against number of disputes resolved.

A further consideration is that monetary compensation may not be necessary in order to induce judges to produce precedents. The production of precedents may yield substantial nonpecuniary rewards to judges[14]--especially in a system where they are paid salaries unrelated to the number of disputes resolved.

The question how to finance adjudication feeds back into the original question whether judicial services can be efficiently provided by the market. Insofar as it is difficult or impossible to finance these services out of user fees, as in the case of precedent creation, the attractions of private provision are reduced, because it is difficult to see how these services would be financed in a free market except on a user-fee basis. This point reinforces our earlier suggestion that the precedent-creating function of adjudication, more than the dispute-resolving function, may invite public intervention in the judicial-services market.

---

13. See William M. Landes & Richard A. Posner, Legal Precedent: A Theoretical and Empirical Analysis, supra note 8.

14. See id. at 272-73.

B. Examples of Private Adjudication: Primitive Societies and Commercial Arbitration

1. Adjudication in Primitive Societies. The anthropological literature on the judicial institutions of primitive societies[15] provides one source of data for testing the economic analysis in the previous subpart.

The literature shows that adjudication is not dependent on the existence of a state as we would understand the term. The governmental institutions of primitive societies are often rudimentary to the point of nonexistence. There may be no legislature, no permanent executive (as distinct from a chief who leads in wartime), no government bureaucracy, no public judges, no public prosecutors or police--indeed, no concept of public law. Yet even in such societies, there will often be adjudication.[16] For example, the Yurok Indians of California had no government at all but a well developed system of private judging.[17] A Yurok who wanted to prosecute a legal claim would hire two, three, or four nonrelatives from a community other than his own and the "defendant" would do likewise. These men (called "crossers") would go between the litigants (who generally did not meet in person during this period) to ascertain

---

15. See, e.g., Henry Maine, Ancient Law (Everyman's Lib. ed. 1917); E. Adamson Hoebel, The Law of Primitive Man: A Study in Comparative Legal Dynamics (Atheneum ed. 1968); Paul Bohannan, Law and Warfare: Studies in the Anthropology of Conflict pts. I-II (1967); Leopold Pospisil, Anthropology of Law, A Comparative Theory (1971).

16. The anthropological literature distinguishes, quite properly, between mediation and adjudication, the former referring to negotiation through intermediaries who have no decision-making powers. Our analysis does not include mediation.

17. See A. L. Kroeber, Law of the Yurok Indians, 2 Atti del XXII Congresso Internazionale degli Americanisti 551 (1928); E. Adamson Hoebel, supra note 15, at 24-25; Robert Redfield, Primitive Law, in Paul Bohannan, supra note 15, at 3, 9-10, 17; Walter Goldschmidt, Ethics and the Structure of Society, 53 Am.Anthropologist 506 (1951); Harold E. Driver, Indians of North America 334, 361 (1961).

claims and defenses and collect evidence. After hearing all of the evidence the crossers would render a judgment for damages. Each crosser received some shell currency from the litigants for his work.

Three questions immediately arise in analyzing the Yurok system of adjudication from an economic standpoint. First, what was to prevent each party from hiring as his "crossers" people he would be sure would act as his agents? Then a guilty defendant would never be convicted.[18] Second, how was the crossers' judgment enforced in the absence of public coercive authority? Third, how were the rules of law applied by the crossers created in the first place given the absence of a legislature or a permanent judiciary?

The first and third of these questions are not directly answered by the anthropological literature on the Yuroks but can be inferred from what is known of other primitive societies. The answer to the second is straightforward: someone who failed to pay a judgment rendered by the crossers became the plaintiff's wage slave and, if he refused to submit to this punishment for "contempt of court," he became an outlaw--that is, he could be killed by anyone without any liability attaching to the killer. This is an example of a sanction that is effective though not backed by the coercive power of any state. It is an extreme form of the ostracism or boycott sanction discussed earlier.

The first question is tougher, but the essential clue to its answer is probably to be found in the requirements (1) that each party had to pick at least two crossers and (2) that each crosser had to be a nonrelative of the party choosing him and had to reside in a different community from the party. These requirements minimized the probability that either litigant would be able to

---

18. This is not a strategy the plaintiff could usefully employ. Assuming he hired stooges and defendant likewise, the plaintiff would have only half the votes. Presumably a tie vote among the crossers resulted in judgment for the defendant, although this question is not addressed in the literature.

find a friend to serve as crosser. To be sure, they did not eliminate the possibility of a corrupt crosser--i.e., one paid by the litigant to render a judgment in his favor irrespective of the merits of his position. Bribery, however, is a general threat to judicial integrity rather than a special problem of primitive society. Whether Yurok society had a rule against bribery of crossers and, if so, how such a rule was enforced--since crossers were the only judges--are questions that the literature does not answer.

Despite the absence of any formal rulemaking machinery, the Yurok Indians had a well-developed body of rules--not of criminal law, because that presupposes a state, but of tort and property law. Where did these rules come from? The answer suggested by the anthropological literature, and one consistent with the economics of the problem, is custom. Societies that adhere to certain rules will enjoy an advantage in the competition with other societies. For example, a society that had no strictures against unlimited killing within it would be at a disadvantage in competition with other societies and would tend to disappear. The societies that survive into recorded history can be expected to have rules that promote social coherence and effectiveness. The rules emerge by a competitive or evolutionary process without need for a formal organ to promulgate them in a deliberate or self-conscious fashion.[19] (As we shall see in Part II, an evolutionary perspective may also have some value in explaining rule creation in our own, more self-consciously rule-creating, society.)

Presumably, as the conditions in a society change, the customary law will change in an adaptive fashion.[20] A problem arises, however, because there is

---

19. There is no evidence, for example, that the crossers issued any form of oral or written "opinion" that might have contributed to the evolution of the customary law of Yurok society.

20. If social change is rapid, however, a system of customary law will probably break down, as its evolutionary character would seem to preclude rapid change.

no mechanism for declaring a custom repealed. This is the source of the "legal fiction,"--the pretended adherence to one rule when a contrary rule is in fact being followed.[21] Legal fictions are found primarily in legal systems where the mechanics of formal rule promulgation, and hence of formal rule change, are deficient. The custom changes but the statement of the former custom persists. Maine associates legal fiction as the method of law change with the earliest stage of legal system's development.[22] It might be more accurate to regard legal fiction not as the agent of change but as a symptom that change in the formal statement of law is lagging behind actual change in the law. The lag is apt to be longest in a system of customary law.

Another striking feature of primitive law, and one closely related, we conjecture, to the character of private adjudication, is its extreme, and to modern eyes excessive, exactness. Yurok law, for example, consisted of a series of very precise rules unsoftened by any general principles which might allow for flexible application. These are familiar features of primitive legal systems.[23] One possible explanation is that the "primitive mind" is incapable of conceptualization, but another, more congenial to economic analysis, is that the Yuroks' adjudicative machinery, consisting as it did entirely of the ad hoc crossers, lacked incentives for the promulgation of general principles that would modify the customary rules of the society. The benefits of such principles would adhere not to the parties who had hired the crossers but to future parties. This problem is avoided in modern commercial arbitration by the fact

---

21. See generally Lon L. Fuller, Legal Fictions 56-80 (1967). A famous example of fiction in English law is the use of the action in ejectment to try title to real estate by the fictitious allegation that the land is occupied by a tenant who has refused to leave upon the expiration of his lease.

22. See Henry Maine, supra note 15, at 15

23. See, e.g., id. at 184, 218.

that, as we shall see, the law applied by the arbitrators is, normally, either the relevant formal law (which includes the precedents created by public judges) of the society or the formal rules of some group or association. The problem is also avoided in those primitive societies where adjudication is the responsibility of a chief or other official who has continuity as an adjudicator and obtains material or other rewards not only from resolving disputes as they arise but also from maintaining the basic harmony and effectiveness of the society. We conjecture that such societies have more flexible rules than those found in pre-political societies such as that of the Yuroks.

2. Commercial Arbitration. We could provide additional examples of pre-political judicial systems similar to those of the Yurok Indians but rather than do so we shall move on to a modern counterpart of primitive courts--the system of commercial arbitration.[24] The main parallel is that arbitrators are private individuals compensated by the parties rather than by the state. (A modern, though only approximate, parallel to the Yurok Indians' method of decision by "crossers" is the type of arbitration in which each disputant appoints one arbitrator and the two then appoint a third.) As in primitive law, arbitrators are appointed anew for every controversy, and do not write opinions[25] and hence are not a source of rules or precedents.

When we examine the way in which modern arbitration seeks to overcome the problems identified in our discussion of primitive private judging, some illuminating differences--as well as further parallels--emerge. One problem is the selection of the arbitrator. As discussed earlier, if one party to a dispute expects that an impartial arbitrator would rule against him, he has an incentive

---

24. We exclude--for no good reason other than limitations of space--the field of labor arbitration.

25. With an important exception noted later.

to drag his feet in agreeing to the appointment of an arbitrator. Consistently with this point, writers on arbitration agree that the problem of selection makes arbitration a virtually unusable method of dispute resolution where there is no preexisting contractual or other relationship between the disputants.[26] This suggests a clue to the superior ability of primitive compared to advanced societies to function without public institutions of adjudication. Primitive communities tend to be quite small[27] and their members bound together by a variety of mutually advantageous relationships and interactions. Expulsion, outlawry, ostracism, and other forms of boycott or collective refusal to deal are highly effective sanctions in these circumstances. Another way of putting this point is that reputation, a factor recognized in the literature as deterring people from breaking contracts even in the absence of effective legal sanctions[28] is a much more effective deterrent in a small community, where news travels rapidly throughout the entire circle of an individual's business and social acquaintances, than in large, modern, impersonal societies.

Yet even in modern society, certain trade, religious, and other associations correspond, to a degree, to the close-knit, primitive community. For example, securities or commodities exchange whose members derive substantial benefits from membership can use the threat of expulsion as an effective sanction to induce members to submit to arbitration.[29] So can a religious associa-

---

26. See, e.g., Frances Kellor, American Arbitration--Its History, Functions and Achievements 64 (1948); Britt-Mari Blegvad et al., Arbitration as a Means of Solving Conflicts 94 (1973).

27. See, e.g., the statistics on the population of Yurok villages in Harold E. Driver, supra note 17, at 334 (50-165 inhabitants).

28. See, e.g., Arthur Allen Leff, Injury, Ignorance and Spite--The Dynamics of Coercive Collection, 80 Yale L. J. 1, 26-33 (1970).

29. This process, as it operated in the New York Stock Exchange before recent developments in antitrust interpretation impaired it (see text at notes 56-57 infra), is described in Howard C. Westwood & Edward G. Howard, Self-Government in the Securities Business, 17 Law & Contemp. Prob. 518-25 (1952).

tion in which excommunication is regarded by members as a substantial cost;[30] so can a university. Exchanges, religious associations, and (private) universities are in fact important examples of modern "communities" in which private adjudication (whether called arbitration or something else) is extensively utilized in preference to public adjudication. But one cannot generalize from such small close-knit communities to the state or national "community" to which individuals and organizations also belong. For example, if there were no public judicial remedies for breach of contract, one who breached a contract in curcumstances where he expected an impartial arbitrator to rule against him would have an incentive simply to refuse to agree to the appointment of any arbitrator suggested by the other party to the contract. To be sure, if other potential contracting parties learned of his behavior, they would be reluctant to make contracts with him; but if the circle of potential contract partners was a very large one, this reputational cost might not be sufficiently great to induce him to submit to arbitration. Accordingly, one is not surprised to find that, in our society, private arbitration[31] (or its equivalent) is largely limited to two types of case: (1) those where a preexisting contract between the parties requires submission to arbitration according to specified rules for selecting the arbitrator, and (2) those where the disputants belong to an association which provides both arbitration machinery for its members and a set of effective private sanctions for refusal to submit to arbitration in good faith or to abide by its results.

---

30. The threat of excommunication was, for example, the ultimate sanction for refusal to submit to, or obey the decision of, the medieval English ecclesiastical courts, which had an immense jurisdiction covering matrimonial disputes, perjury, and a variety of other matters as well as strictly religious disputes. See Jane E. Sayers, Papal Judges Delegate: the Province of Canterbury 1198-1254, at 157 (1971); Brian L. Woodcock, Medieval Ecclesiastical Courts in the Diocese of Canterbury ch. X (1952).

31. We exclude cases where a statute compels arbitration, as in small claims in Pennsylvania. See Martin Domke, The Law and Practice of Commercial Arbitration 7-8 (1968). These are not examples of genuinely private adjudication.

Even (1) is effective, in major part anyway, only because the public courts enforce such contracts; if they did not, there would often be no effective sanction against the party who simply breached the contract to arbitrate. In sum, unless ostracism or some similar form of reputation-related private remedy is available because the dispute is between members of a close-knit community, a public remedy will often be necessary to induce a disputant who is pessimistic about the outcome of arbitration to submit to it.

The existence of sanctions for refusal to submit to arbitration is not, however, sufficient to assure that the selection process will operate effectively. As mentioned earlier, it may be difficult to determine whether a party who rejects an arbitrator suggested by his opponent is acting in bad faith. The opponent might be proposing a totally unsuitable candidate precisely in order to cast on the other party the onus of refusing to arbitrate. Given these difficulties, ascertaining bad faith--the essential predicate for applying a sanction, whether public or private--may be quite costly.

This explains, we believe, why students of arbitration tend to disfavor the familiar selection process whereby each party appoints an arbitrator and the two party-selected arbitrators then agree on a third.[32] This process is said to lead to foot-dragging. The why and how are easy to see. The party who fears the outcome of an impartial arbitration need only instruct "his" arbitrator to refuse to agree to the selection of an impartial third arbitrator.

This problem is overcome and the costs of arbitration reduced by the procedure used by the American Arbitration Association, a private association that

---

32. See Robert Coulson, How to Stay Out of Court 153 (1968); Ernest J. Cohn, The Unification of the Law of Commercial Arbitration, 24 Trans. Grotius Soc'y 1, 15 (1939); Lionel S. Popkin, Practical Problems Confronting the Practicing Lawyer, 19 Law & Contemp. Prob. 653, 654 (1952).

offers an arbitration service.[33] The AAA invites parties to contracts who wish to include an arbitration clause to specify that the arbitration will be conducted in accordance with the rules of the AAA. These rules provide (among other things) that should a dispute triggering the arbitration clause arise AAA will send each party a list of arbitrators selected by the AAA's staff on the basis of the nature of the dispute. Each party is free to cross off any names from the list it deems unsuitable and to rank its preferences for the others. The staff then selects (as sole arbitrator) the individual most preferred by the parties. A party who crosses everyone off the list hurts only himself, by guaranteeing that the arbitrator will be selected from among the names not deleted by his opponent.

The problem of enforcing the arbitrator's award once made is parallel to that of compelling the recalcitrant party to submit to arbitration in the first place. But, if anything, the problem of enforcement is less serious (or at least less complicated) than that of submission, because refusal to submit to an award is a clearer signal of bad faith than foot-dragging in the selection of the arbitrator.

The third problem addressed in our discussion of primitive adjudication was the creation of the rules that the private adjudicator applies. Arbitrators generally do not write opinions; nor is an arbitrator's award appealable to an "appellate court" of arbitrators. As a result, arbitration awards are not a source of rules or precedents. This is understandable in the case of general commercial arbitration because of the public-good character of precedent. A system in which arbitrators wrote opinions would be at a competitive disadvantage vis-à-vis one in which they did not write opinions; the former would cost

---

33. The AAA is described in Frances Kellor, supra note 26. See also Soia Mentschikoff, Commercial Arbitration, 61 Colum. L. Rev. 846, 862-65 (1961).

more but would yield no greater private benefits and, being private, could not coerce the necessary financial support by invoking the state's taxing powers. This point also explains why, to the extent that the value of appeal procedures lies primarily in the creation of precedent, arbitration does not have an appellate stage.

Here is important evidence, incidentally, that a free market in judicial services would not lead to the production of precedents as a by-product of the efforts of the competing judges to demonstrate their competence and impartiality through the issuance of judicial opinions. The American Arbitration Association does not employ this mode of advertising; nor do we know of any other group that does.

The situation with regard to the incentives to produce precedents in a regime of private arbitration is different where arbitration is prescribed by a tightly knit religious or commercial association which can presumably "tax" the membership to support rule creation by the association's judges. An example is the elaborate "common law" of Jewish religious duties evolved by the Rabbis in their interpretation of the Old Testament in disputes brought before them--a body of law codified in the Talmud. Of course, in many associations (as in many foreign legal systems) the rule-creation function is separated from the adjudicative, and then we have an independent reason for not observing written arbitration opinions. Nonetheless, it has been said that arbitration decisions are accorded precedential value in trade association but not in general commercial arbitration and that trade associations often provide appellate tribunals as part of their arbitration machinery and general commercial arbitration does not.[34] These observations support our analysis.

Where, then, do the rules applied in general commercial arbitration come

---

34. On these points see Soia Mentschikoff, supra note 33, at 857; and G. Ellenbogen, English Arbitration Practice, 17 Law & Contemp. Prob. 656, 673 (1952).

from? They usually come from the courts and other sources of public law; arbitrators typically apply the same rules as courts deciding similar questions, often because the arbitration contract will specify that the arbitrator is to apply the contract law of a particular jurisdiction. Indeed, arbitration is generally limited to disputes where the rules are perfectly clear and the only issue is their application to the facts.[35] In either case, the parties may be said to be taking a "free ride" on the public legislative-judicial system; but this is not a very fruitful invocation of the free-rider concept. Since public adjudication is financed out of general tax revenues rather than litigant fees, the parties to arbitration receive no greater net benefit from the public court system than the litigants in that system do; they actually receive less benefit because they have to pay for the arbitrator whereas the state pays for his counterpart in the public court system.

Because arbitration is a voluntary service provided in a competitive market, it may appear that the procedures widely used in arbitration must be efficient procedures for deciding the type of dispute submitted to arbitration. If so, arbitration procedures could be used as a criterion for evaluating the efficiency of the public judicial system in areas such as contract and commercial law where most arbitrable disputes arise. Before invoking such a presumption, however, one would want to know how arbitration was actually faring in competition with the public judicial system. If only a minute fraction of commercial disputes were submitted to arbitration--and especially if the fraction were declining over time--one might conclude that, whatever a priori reasons there might be for believing a private service to be more efficient than its public counterpart, these reasons had been refuted by the verdict of the market in

---

35. See American Management Ass'n, Resolving Business Disputes--The Potential of Commercial Arbitration 46, 115-16 (1965); Soia Mentschikoff, supra note 33 at 866.

favor of the public service. To complicate the picture still further, it is necessary to factor out the artificial competitive advantage that the public competitor enjoys by virtue of being supported out of general tax revenues and providing its services at no charge[36] --although, for reasons suggested earlier, this may not be an unfair (inefficient) advantage if arbitration is taking a "free ride" on the precedent-creating activities of the public courts.

Still a further complication is that, even if the precedents or procedures used in the public court system were as or more efficient than those of arbitration, the existence of a long court queue--which is equivalent to a price for judicial services--might reduce the net value of the judicial services provided by the public system to the point where arbitration, though socially less efficient, was an attractive substitute. Finally, uniformity of procedure between arbitration and the courts may yield significant economies since the same lawyers participate in both sorts of procedure[37] and may tend to homogenize them. This may be why arbitration procedures are said to resemble the judicial procedures found in the same jurisdiction.[38]

The foregoing considerations make it difficult to use arbitration as the benchmark for judging the court system--and it is no help that the statistics of the number of cases submitted to arbitration are inadequate. Although the American Arbitration Association maintains statistics of the number of commercial arbitrations it conducts, there are no statistics on the number of trade

---

36. However, this would primarily affect the proportion of disputes submitted to arbitration rather than the rate of change of this proportion.

37. Parties to commercial arbitration are generally represented by lawyers. See American Management Ass'n, supra note 33, at 26 n. 5; Hal M. Smith, Commercial Arbitration at the American Arbitration Association, 11 Arbitration J. (n.s.) 3, 12 (1956).

38. See Britt-Mari Blegvad, supra note 26, at 112; cf. Soia Mentschikoff, The Significance of Arbitration--A Preliminary Inquiry, 17 Law & Contemp. Prob. 698, 707 (1952).

association, exchange, etc. arbitrations conducted outside of the AAA.[39] A further difficulty is the absence of statistics of the number of court cases which might have been arbitrated--i.e., the number of contract and commercial-law cases.

For what it is worth--which is not a great deal in light of these difficulties--Table 1 presents a comparison of the number of commercial arbitrations (excluding insurance) conducted annually by the AAA with the number of contract (excluding insurance) cases filed in the federal district courts under the diversity jurisdiction. Table 1 indicates that since 1959 commer-

[Insert Table 1 about here]

cial arbitrations have been increasing at about the same rate as court cases.[40] But more work must be done to determine the survival characteristics of arbitration versus litigation, since our figures may be dominated by shifts in the AAA's market share of arbitrations or in the federal district courts' market share of adjudications.

Notwithstanding all the above reservations, the use of arbitration as a benchmark for evaluation of the judicial system may help resolve a recent controversy between Gordon Tullock and others regarding the relative efficiency of the Anglo-American adversary and Continental inquisatorial procedural systems.[41] It appears that most arbitrations are conducted according to English or American

---

39. A study conducted in the mid-1950's on the basis of a mail questionnaire sent to trade associations found that the AAA accounted for 27 percent of all commercial arbitrations conducted either by it or by the trade associations sampled. Soia Mentschikoff, supra note 33, at 857.

40. We use 1959 rather than 1957 as our base year because of the increase in the jurisdictional minimum amount for bringing diversity cases in the federal courts from $3,000 to $10,000 in 1958.

41. See Gordon Tullock, On the Efficient Organization of Trials, 28 Kyklos 754 (1975); Fred S. McChesney, On the Procedural Superiority of a Civil Law System, 30 id. at 507 (1977); J. A. Ordover & Phillip Weitzman, On the Efficient Organization of Trials: A Comment, 30 id. at 511 (1977).

TABLE 1

Comparison of Judicial and Arbitration Caseloads

1957 - 1976

	Contract (excluding insurance) cases filed in federal district courts under the diversity jurisdiction	Commercial (excluding insurance) cases filed with the American Arbitration Association
1957	3918	677
1958	4329	711
1959	2687	707
1960	2624	820
1961	3090	N.A.
1962	3180	N.A.
1963	3470	N.A.
1964	3599	N.A.
1965	3538	N.A.
1966	3931	N.A.
1967	4252	N.A.
1968	4532	1634
1969	4642	1964
1970	5619	2658
1971	6882	2672
1972	7310	3092
1973	8344	3228
1974	9619	3809
1975	11883	4128
1976	12229	4093

Sources: American Arbitration Association; Annual Reports of the Director of the Administrative Office of the United States Courts

arbitration procedure[42]--and, as mentioned, a nation's arbitration procedures tend to follow its judicial procedures. Here, then, is some, albeit limited, market evidence of the superiority of the adversary system.

Were it possible to use arbitration procedures as a standard of judicial efficiency, several points would be notable.[43] First, arbitration never involves trial by jury, and rarely is there more than one arbitrator; yet commercial arbitrators normally are not lawyers either, but are rather businessmen expert in the particular industry in which the dispute to be arbitrated arises. So evidently it is not the lay character of the jury that is inefficient (using arbitration procedures as the benchmark for judging the efficiency of a procedural system), but the jury's lack of expertise with regard to the subject matter of the litigation. The evidence from arbitration is that a single qualified lay judge is superior to six or 12 randomly selected laymen--on reflection, a not implausible suggestion.

The second point to be noted is that there is no appeal in general commercial arbitration. This suggests that the principal value of appellate proceedings is not to correct errors at the trial level but to formulate rules of law.[44] Third, although there is pretrial discovery in arbitration, it is much less

---

42. See Michael Marks Cohen, A Venue Problem With the Arbitration Clauses Found in Printed Form Charters, 7 J. Maritime L. & Comm. 541 (1976); Lynden Macassey, International Commercial Arbitration: Its Origin, Development and Importance, 24 Trans. Grotius Soc'y 179, 199 (1939); Donald E. Zubrod, Arbitration from the Arbitrator's Point of View, 49 Tul. L. Rev. 1054, 1055 (1975).

43. A description of these procedures may be found in the Domke treatise, supra note 31, and in Soia Mentschikoff, supra note 33, at 862-65.

44. It is consistent with this analysis that appellate tribunals are sometimes provided in trade association arbitration. See note 35 supra. It is also possible, of course, that there is less emphasis on appeal in arbitration because the arbitrator is a more expert factfinder than a jury. But the presence of appeals in trade association arbitration argues against this view, as does the traditionally restricted character of appellate judicial review of purely factual questions.

extensive than under the Federal Rules of Civil Procedure--which are much criticized for excessive liberality in this regard. Fourth, the rules of evidence of ordinary litigation are observed--in spirit albeit not in letter--in arbitration, even though the arbitrator is not a lawyer. This suggests that trial by jury is not a sufficient explanation for the Anglo-American rules of evidence. A fifth point is that, judging from arbitration practice, the judicial rules denying specific performance of particular types of contract, e.g., personal-service contracts, are inefficient since arbitrators are not bound by these rules and do not in fact observe them.[45] Sixth, while the arbitrator's fee is typically added to the arbitration award (i.e., is paid by the defendant if the plaintiff wins, and split between the parties if the defendent wins), unless the arbitration contract specifically provides that the winning party's attorney's fees are to be indemnified by the loser they are not.[46] This is some evidence that the English and Continental rule of indemnity may not be more efficient than the American rule, which does not provide for indemnity. It is important evidence since the theoretical economic analysis is indeterminate on the question,[47] but inconclusive since it is possible that most arbitration contracts do provide for indemnity. We know of no evidence on that question.

C. Competition in the Judicial-Services Market

1. Nonadjudicative Substitutes for Public Adjudication. We have thus far assumed that the only substitute for public adjudication is some sort of private adjudication. But this is obviously incorrect, though, since our major interest

---

45. See Staklinski v. Pyramid Electric Co., 6 N.Y. 2d 159, 188 N.Y.S. 2d 541, 160 N.E. 2d 78 (1959); Grayson-Robinson Stores, Inc. v. Iris Construction Corp., 8 N.Y. 2d 133, 202 N. Y. S. 2d 377, 168 N.E. 2nd 377 (1960).

46. See Martin Domke, supra note 43, at §§42.01 - .04.

47. See Richard A. Posner, Economic Analysis of Law 450-53 (2d ed. 1977).

in the paper is in adjudication, we will not elaborate this point fully. If there were no private courts yet the substantive rules or the procedures of the public courts were inefficient, substitution away from dispute resolution would take place. An important example of the substitution possibilities is the liquidated-damages clause, whereby the parties substitute a damage-assessment formula of their own choosing for whatever rules of contract damages or methods of damage assessment the courts employ. The more costly or less accurate the judicial methods of damage assessment, the more we would expect parties to resort to liquidated-damages clauses. Thus, in principle at least, ebbs and flows in the popularity of such clauses could be used to measure the efficiency of judicial damage assessment methods.

Many other examples of substitutes for judicial dispute resolution come to mind, but we will mention only two. (1) Inefficient judicial debt-collection remedies can be expected to induce a substitution of cash for credit transactions--or simply an increase in lenders' bad-debt reserves and higher interest charges. (2) If accident victims could not obtain reasonably prompt and complete compensation in the courts, they could be expected to reduce their involvement in activities giving rise to accidents or to increase their purchase of accident insurance. The relationship between nonadjudicative substitutes and public courts is analyzed in greater detail in Part II of this paper.

2. Competition Between Court Systems. Not only does the public court system face potential competition both from private methods of dispute resolution and from substitution away from activities that lead to judicially cognizable disputes; there is also the possibility of competition between public court systems. In the American judicial system, for example, many plaintiffs have a choice between a state and a federal court for a variety of disputes

because the state and federal courts have concurrent jurisdiction of, for example, disputes between citizens of different states and accidents subject to the Federal Employees Liability Act); and some plaintiffs have a choice between several different state courts.[48] Contracting parties can stipulate the jurisdiction whose law they want applied to the contract--which might even be a foreign nation--and these stipulations are generally honored.[49] In the early history of English law, the three royal courts were in competition with each other, with the court of Chancery, with the ecclesiastical courts, and with a variety of manorial and other local courts, for litigants.[50] We have now to consider the character of judicial competition.

Imagine a system in which there are several courts, public or private, with overlapping jurisdictions, and the judges are paid out of litigant fees and therefore have a direct pecuniary interest in attracting business away from competing courts. If we put to one side the problem discussed earlier of precedent production under conditions of judicial competition and fee-per-case judicial financing, it might seem that competition would lead to an optimal set of substantive rules and procedural safeguards. But this is incorrect. The competition would be for plaintiffs, since it is the plaintiff who determines the choice among courts having concurrent jurisdiction of his claim. The competing courts would offer not a set of rules designed to optimize dispute resolution but a set designed to favor plaintiffs regardless of efficiency.[51]

---

48. Where cases instituted in state court are removable to a federal court, it is the defendant who has the choice.

49. See American Law Institute, Restatement of the Laws Second, Conflict of Laws 183-200 (1968).

50. See Theodore F.T. Plucknett, A Concise History of the Common Law 81, 98 (1956); C.R. Cheney, From Becket to Langton: English Church Government 1170-1213, at 108-17 (1956).

51. Although this tendency would be held in check to some extent by the ability of potential defendants to substitute away from activities generating litigable disputes.

This problem could be overcome by allowing the defendant to opt out of the forum chosen by the plaintiff. This would mean that the parties would have to agree on the court to which their dispute was submitted. However, this "solution" would simply reintroduce, in a different form, the problem discussed at length earlier in the paper of the party who, fearing the outcome of an impartial adjudication of his dispute, refuses to agree to the selection of an impartial adjudicator. This problem can be overcome only if the parties to a contract agree in advance to the submission of any dispute arising from the contract to a particular tribunal.

The foregoing analysis predicts the pattern that we in fact observe in the history of English and American law. The public judicial system tends toward monopoly save in cases where the parties agree in advance to submit any dispute to another tribunal (such as some system of private arbitration). The rare exception to the public judicial monopoly is where plaintiff choice (with the danger of favoritism it imports) is thought to be required to offset some bias toward defendants. This is the theory of the federal diversity jurisdiction. Even so, competition between state and federal courts for diversity plaintiffs might be intolerable if judges were compensated out of litigant fees (as they are not), for then the competing court systems might outdo each other in offering plaintiffs procedural advantages beyond any required to offset prejudice against them as citizens of other states.

Left unexplained by this analysis is the actual pattern of competition in the English courts during the centuries when the judges were paid out of litigant fees and plaintiffs frequently had a choice among competing courts. There is evidence of competition among the courts for plaintiffs through

substantive and procedural innovation,[52] but none (of which we are aware) of the kind of blatant plaintiff favoritism that our economic analysis predicts would emerge in such a competitive setting. Why it did not emerge (assuming it has not simply been overlooked by legal historians) presents an interesting question for further research.

Analysis of competitive forces suggests another respect in which paying judges out of litigant fees might be inefficient. If one assumes that an effective system of arbitration requires judicial enforcement both of agreements to submit to arbitration and of arbitration awards, then the judges stand in two relations to arbitration: as suppliers of an essential input into arbitration services--namely public enforcement of the agreement to arbitrate and of the award[53]--and as competitors for cases and fees. The judges can reduce competition from arbitration by refusing to enforce the agreement to arbitrate or the award. Although they lose the fees they would have charged the parties to the arbitration enforcement suit, they gain added fees from the diversion of disputes from arbitration to the courts.

Such an analysis provides an economic basis for allegations that the refusal of the common law to order specific performance of agreements to

---

52. Though, for a skeptical view regarding a famous example of this alleged competition (Slade's Case), see A.W.B. Simpson, A History of the Common Law of Contract: The Rise of the Action of Assumpsit 294-95 (1975).

53. There is another, less important respect in which judicial intervention may be necessary in arbitration: to compel attendance of disinterested witnesses (i.e., witnesses indifferent to the grant or denial of the arbitration award). See e.g., United States Arbitration Act, 9 U.S.C. §7. Of course, a party could be required to negotiate with a witness over the terms of his attendance, but that would create potentially serious problems of bilateral monopoly and of incentives to lie.

arbitrate,[54] a refusal that dates from the time when the English judges received litigant fees, was in fact motivated by hostility to the competition of arbitration.[55] The analysis, however, is incomplete. Since the judges supply an essential input into arbitration, namely enforcement of agreements to arbitrate and of arbitration awards, they could in principle fix a schedule of fees for these services that would extract the entire profits of arbitration. If arbitration were a more efficient method of dispute resolution than formal adjudication, those profits would exceed what the judges could obtain by snuffing out arbitration.

The refusal (mysterious to an economist) of the courts to enforce penalty clauses in contracts is sometimes attributed to a competitive hostility to the liquidated-damages clause, a substitute, as we have seen, for judicial dispute resolution. But the same puzzle recurs. Judges could extract the profits of substituting liquidated-damages clauses for judicial damage assessment in the fees they charged for enforcing such clauses.

The mystery is dispelled by reflection on the actual nature of judicial compensation. Even when English judges received a substantial part of their incomes in the form of litigant fees, they also received public salaries;

---

54. Breach of a contract to arbitrate was actionable, but damages could never be proved, as the only injury to the victim of the breach was to be remitted to his legal remedies for whatever wrong he wanted the arbitrator to remedy. See discussion in Kulukundis Shipping Co. v. Amtorg Trading Corp., 126 F. 2d 978 (2d Cir. 1942). For a summary of the common law of arbitration see 6A Corbin on Contracts 383 (1962), and for detailed histories, Julius Henry Cohen, Commercial Arbitration and the Law (1918); and Morton J. Horwitz, The Transformation of American Law, 1780-1860, Ch V (1977).

55. See, e.g., Scott v. Avery, 25 L.J. Ex. 308, 313 (1855). Judicial hostility to arbitration, although well documented, was by no means uniform. See Morton J. Horwitz, supra note 54; John H. Langbein, Book Review, 18 Am. J. Legal Hist. 88, 91-92 (1974).

nor, so far as appears, had the judges much, if any, control over the setting of the fees. If the judges did not control the fee levels, the alternative (to discouraging arbitration) of setting fees for enforcing arbitration awards or compelling submission to arbitration would be unavailable to them, and the second-best solution of "foreclosing" competition might be chosen.

More mysterious is judicial hostility to competition in an era when judges' compensation is no longer proportioned to their output of disputes resolved. An example is the judicial antipathy to boycotts. One can understand why a boycott designed to enforce a cartel would be struck down under the antitrust laws, but not why a boycott designed to create an effective system of arbitration would be, as was done in Paramount Famous Lasky Corp. v. United States.[56] A group of motion-picture producers agreed to include in their contracts with exhibitors an arbitration clause and, in effect, to boycott exhibitors who refused to arbitrate or to abide by the terms of an arbitration award. There was no evidence (at least mentioned by the Court) that the purpose or likely effect of the agreement was other than to make arbitration effective by imposing an effective sanction--termination of contractual relations with other members of the producer group--for refusal to arbitrate.

A superficially more liberal attitude toward the use of the boycott to effectuate a system of private adjudication appears in the later case of Silver v. New York Stock Exchange,[57] which held that enforcement of stock-exchange rules by boycott would not violate the antitrust laws if adequate procedural safeguards were afforded in the exchange's disciplinary proceeding.

---

56. 282 U.S. 30 (1930).

57. 373 U.S. 341 (1963).

Yet before Silver, it was generally assumed that the antitrust laws had no application to the private self-government schemes of the regulated exchanges.[58]

The result in Paramount Lasky is sometimes explained on the basis of hostility to private government. But this formulation simply restates the result of the decision. It is quite true that expulsion, ostracism, and similar forms of boycotting are, as we have seen, methods by which private systems of adjudication compensate for their lack of public coercive powers. But the fact that the boycott is an important, perhaps indispensable method of enabling such systems to operate simply indicates that there is a legitimate justification for the use of the device and makes it improper to treat boycotts as per se violations of antitrust law when their purpose is not to restrain trade.

A somewhat more plausible objection to the boycott device is that it could be abused by the dominant faction in an association.[59] That is a serious problem where the association does not operate within the framework of a larger polity having coercive authority, but that is not the case with modern associations. Any contractual rights that members of an association may have against majority oppression (analgous to those of minority shareholders in corporations) would be enforceable in the public courts.

Three features of judicial competition remain to be discussed briefly.

(1) Where parties can feasibly stipulate the forum, public or private, for adjudicating disputes arising between them, competition is feasible and we would expect efficient rules of substantive law to emerge. A famous

---

58. See Howard C. Westwood & Edward G. Howard, supra note 29.

59. See note 4 supra.

example of this is the absorption of the law merchant by the English courts. Throughout the middle ages European merchants had their own private courts for the adjudication of commercial disputes--the system known as the lex mercatoria.[60] Gradually, the doctrines developed by these courts to deal with contract and commercial matters were absorbed into the common law and the official courts began winning business from the merchant courts.[61] Conceivably the financial self-interest of the English judges, who, as previously noted, were paid in part out of litigant fees during this period, was a factor in the absorption of the law merchant into the common law. In similar vein English procedural reform in the nineteenth century has been attributed in part to the competition from private arbitration.[62]

Notice that the English courts could not have won merchant business simply by favoring plaintiffs, for merchants would terminate relations with fellow merchants who utilized a biased court system. Notice further that the process by which competition between a public and a private judicial system yields efficient rules of substantive law will work even if, because of the method of judicial compensation or for other reasons, the judges are indifferent to the siphoning off of cases to competing public or private adjudicative systems. After siphoning has occurred, the cases observed will be those decided in systems which offer efficient rules. Those cases will shape the observer's impression of the content of the law. The general conclusion is that we can expect more efficient rules of contract and commercial law (including corporation law, which is also based on consensual arrangements) than

---

60. See Wyndham Anstis Bewes, The Romance of the Law Merchant pt. I (1923); Julius Henry Cohen, supra note 54, at 73-81; William Mitchell, Essay on the Early History of the Law Merchant (1904).

61. Theodore F.T. Plucknett, supra note 50, at 657-70.

62. See Brian Abel-Smith & Robert Stevens, supra note 2, at 39.

of tort or criminal law, because parties to contracts face a competitive supply of court systems.[63]

(2) As mentioned earlier, state and federal courts are competitors with regard to dispute resolution in the areas of their overlapping jurisdiction (for example, the diversity jurisdiction), but there is a monopoly so far as the production of precedent is concerned. A federal court deciding a tort case under the diversity jurisdiction uses its own factfinding procedures but applies the law of the state in which the tort took place; it may not create its own precedents in competition with the courts in that state. Litigants can opt out of the public dispute settlement machinery entirely through arbitration but the body of legal principles known as the conflict of laws (or choice of law) prevents disputants from shopping among jurisdictions to find a more favorable body of substantive law. In general, though not invariably, "forum shopping" is limited to a search for an alternative dispute settlement procedure or tribunal. Further, although there are many judges within a court system, a single tribunal--a supreme court--will monopolize the precedent production of the system through its power to reverse any discordant lower-court decision.

(3) We may be able to explain the movement toward centralization of the judicial function--e.g., the growth of the royal courts in England and of the federal courts in the United States. In a society with little mobility, a system of local courts will not generate intolerable competition among systems of substantive law. But with a mobile population, the system of regional court monopolies breaks down and must be superseded if competition in substantive law with resulting information overload is to be avoided.

---

63. This point is analyzed more fully in Part II of this part, infra.

II. Are Public Judicial Outcomes Also Privately Determined?

A. Introduction

Where a judicial system is established and financed by private parties without public regulation or intervention, the procedures and outcomes of that system may fairly be regarded as privately rather than publicly determined, and a set of substantive rules evolved by such a system would have the highest claim to be regarded as efficient. One would not expect public court systems to have an equally private character. Yet recent writers on public adjudication--Rubin, Priest, and Goodman--have argued that the outcomes of public judicial processes are indeed privately determined.[64] The basic reason for this result is that in their analysis the only inputs into public, as into private, adjudication that count are private. The decision of parties to litigate their dispute in a public court or settle it out of court is assumed both to be determined by the efficiency of the rules they are contending for and to determine the outcome of litigation and ultimately the legal rule themselves. Rubin and Priest argue that under certain conditions this process leads to the development over time of more efficient legal rules.[65] Inefficient rules--ones that impose greater costs on parties in, for example, accident avoidance or contract formation--are less likely to survive not because judges favor or even understand principles of

---

64. Paul H. Rubin, Why Is the Common Law Efficient?, 6 J. Legal Studies 51 (1977); George L. Priest, The Common Law Process and the Selection of Efficient Rules, 6 id. at 65; John C. Goodman, An Economic Theory of the Evolution of Common Law (unpublished).

65. For this conclusion to hold, Rubin requires that both parties have significant (and approximately symmetrical) on-going interests in future cases similar to the one currently being litigated. Priest's analysis does not incorporate this requirement, but does require a higher litigation rate when legal rules are inefficient compared to when they are efficient.

efficiency but as a consequence of systematic differences in the trial-settlement choices of parties subject to efficient and inefficient rules. Thus the rules that survive are determined by private choices.

The approach assimilates public to private courts. Moreover, if correct, it has important implications for the economic theory of the common law, which predicts that common law rules are efficient but has been unable to discover the mechanism that generates these results. In this part of our paper we modify and extend the Rubin-Priest approach. Because our analysis, like theirs, is formal, we include here a brief, nonmathematical summary.

The economic model of litigation[66] is concerned with identifying the circumstances under which a legal dispute will be litigated rather than settled out of court. Since settlement is assumed to be cheaper than litigation, why would disputants ever go to trial? The answer requires identifying the expected gain to the plaintiff and the expected loss to the defendant from litigating. If the expected gain is less than the expected loss--if defendant expects to lose more than the plaintiff expects to gain--both parties will be eager for settlement. But if the plaintiff's expected gain is greater than the defendant's expected loss, the defendant will be unwilling to make an attractive settlement offer and litigation will ensue. The determinants of the expected gains and losses are the subjective probabilities of victory by each of the parties and the stakes to each of them.

The parties' decision calculus may be affected by the precedential character of the decision in their case, if the case is litigated to judgment. A decision in plaintiff's favor will increase the probability of the plaintiff's

---

66. See, e.g., John P. Gould, The Economics of Legal Conflicts, 2 J. Legal Studies 279 (1973); William M. Landes, An Economic Analysis of the Courts, 14 J. Law & Econ. 61 (1971); Richard A. Posner, An Economic Approach to Legal Procedure and Judicial Administration, 2 J. Legal Studies 399 (1973). A very simple version of that model appears in Richard A. Posner, supra note 47, at 434-36.

winning similar cases in the future, and a decision for defendant will increase the probability of his winning similar cases in the future. Either way, the decision will alter (though, in an incremental, common law system, we assume slightly) the ratio of favorable to unfavorable precedents applicable to the parties' future activities.

Many plaintiffs and defendants will, of course, be indifferent to the precedential effect of a judgment in their case because they do not expect to be involved in a similar dispute in the future; for them, precedent does not enter significantly in measuring the stakes of victory and defeat. But for others, precedential effect is an important dimension in deciding whether to litigate the present case because they do anticipate future similar disputes. These are the parties who in our terminology have "future stakes" in the current dispute.

Rubin concludes that where future stakes of comparable importance are present for both parties the tendency of the common law process is to generate efficient rules. We argue that this conclusion is determined by the particular fashion in which, in his model, the decision in the present litigation operates as a precedent for future disputes. In Rubin's model, if a decision is against efficiency it has no impact on the state of the law, while if it is for efficiency it dramatically alters the balance of precedents in favor of the efficient rule. When an inefficient rule is in force, the party hurt by the rule has a strong incentive to challenge it because if he wins the rule will be replaced by an efficient rule which (by definition) will confer greater benefits upon him than losses on his opponent. But if an efficient rule is in effect, the party who loses from that rule does not have an equivalent incentive to litigate because a victory will confer smaller future benefits upon him than losses on his opponent.

The net effect is that disputes governed by inefficient rules are litigated until the rule is reversed in favor of efficiency.

We, in contrast, assume that current decisions have small and symmetrical effects on the state of the law or balance of precedents. We are therefore led to expect more litigation in areas where the rules are already efficient. For, in such areas, the likely outcome of the litigation will be a decision placing liability on the party who is the cheaper cost avoider (the party who will be induced by liability to take steps that minimize the relevant costs), thereby reducing the future costs of the activity. If, however, the dispute arises in an area where the likeliest outcome is a precedent that will strengthen an existing inefficient rule, litigation will be avoided because its expected yield is negative. Therefore, we expect litigation to arise mainly in areas where there is already a tendency toward efficiency, and this tendency will be further strengthened by litigation that creates additional precedents. Areas dominated by inefficient rules will tend to become dormant in terms of litigation activity.

The analysis as we have outlined it thus far rests on rather austere assumptions: that both parties have future stakes in the litigation; that these stakes are approximately equal; that the parties' current stakes and perceived probabilities of victory are the same; and that their expenditures on litigation are fixed. When we relax these assumptions, we find, first, that if the parties have asymmetrical stakes, the conclusion that there will be little litigation in areas dominated by inefficient rules is weakened. This follows because if the people benefitted by these rules have much larger future stakes than those

hurt by them, they will have incentives to litigate in order to strengthen the rule by adding precedents supporting it. Second, if neither party has any future stake, the analysis collapses, for then neither party has an interest in precedent, so the costs and benefits generated by precedent will not enter into their decision to litigate. This conclusion parallels the analysis in Part I of this paper. Precedent has "public good" aspects that may result in underproduction in a private market. However, to the extent that the costs and benefits of precedent will be borne (in the future) entirely by the parties to the suit in which the precedent is created, precedent is a private rather than public good.

When the parties can, if they wish, contract around a rule of law, either by expressly stipulating to the contrary or opting out of the public court system entirely, as through an agreement to arbitrate, the effect is to place a floor under any tendency of the legal system to produce (for example, in circumstances where people who benefit from an inefficient legal rule have much greater future stakes than those hurt by it) inefficient rules. Such a rule will be followed only where the inefficiency is less than the cost of contracting around the rule. But the prediction of a tendency toward efficiency in areas where the legal rules are already efficient will not be affected. Hence, as one would expect, in areas where contracting around is feasible (or, in the terminology of Part I of the paper, where there are private alternatives to the public court system), the tendency of the common law toward efficiency is accelerated even if the judges are indifferent to the loss of business that contracting around entails.

If the assumption that litigation expenditures are exogenous (i.e., independent of any of the variables in our model such as the efficiency

of the legal rules or the current stakes of the parties) is relaxed, our conclusions are actually reinforced. In general, treating litigation expenditures as endogenous implies that an increase in the expected gains of suit will cause a party to spend more on litigating his case, which in turn will increase his chances of winning. This magnifies, for example, the effects of assuming asymmetrical stakes. The party with the larger future stakes will spend much more on litigation than the party with smaller future stakes and this in turn will increase still further the first party's probability of winning.

Priest's model, unlike Rubin's, assumes that the parties to the current dispute have no interest in precedent. Future stakes are not an element of his analysis -- he focuses on present stakes entirely. As noted earlier, other things being equal, an increase in the stakes in a case will increase the tendency to litigate. Because an inefficient rule is by definition more costly than an efficient one, Priest concludes that the stakes will tend to be greater in cases where the governing rule is inefficient. Therefore there will be a greater tendency to litigate such rules and, in the process of litigation, judges will occasionally reexamine the rule, overrule it, and replace it with an efficient rule. Efficient rules, in contrast, because they involve smaller stakes on average, will tend to be relitigated less often and hence overruled (and replaced by inefficient rules) less often.

We believe, to the contrary, that once precedent is introduced into the litigation model, Priest's results are reversed. Even though the parties themselves have no interest in the precedential significance of the decision in their case, that decision will both be influenced by and influence the balance of precedents and hence the efficiency of the

rule. If, as Priest suggests, litigation will be more frequent in areas of inefficient rules because the costs of such rules are greater, then the tendency of these rules toward inefficiency will be strengthened by each successive decision. To say that the governing rule is inefficient is to say that cases within the domain of the rule are more likely to be decided inefficiently than efficiently, and every time that happens the rule is strengthened by the greater accretion of precedents. There will be less litigation of efficient rules and hence a smaller accumulation of precedents confirming and thereby (in a system of decision according to precedent) strengthening those rules.

The rest of this paper develops the analysis sketched above. After setting out some general considerations and assumptions, we next consider (subpart C) the relative survival properties of efficient and inefficient rules when the disputants have no prior contractual relationship with each other and litigation expenses are exogenous. We then generalize the analysis both to include parties that have a prior contractual relationship with each other and thus are able to contract around inefficient rules (subpart D), and to consider the situation where litigation expenditures are endogenous, being determined, in part, by the relative stakes of the parties (subpart E). Finally, we attempt to develop more rigorously Priest's analysis of the long-run equilibrium of the legal system when litigation rates are related to the relative efficiency of legal rules(subpart F).

B. Some General Considerations

We use the following notation in our analysis:

A	plaintiff
B	defendant
x	amount of damages that A is seeking from B
j	activity (e.g., driving) that gave rise to A's claim for damages
p	probability that in the event of a trial B will be liable for damages of x (thus 1-p is the probability that B will not be liable)
λ	ratio of precedents (or legal capital) favoring A's claim to those favoring B's claim
r^a, r^b	litigation expenses of A and B, respectively
S	sum of present values of all future damages and expenditures to reduce both the probability and amount of damages in activity j (including losses from reduced participation in j).

We assume that the assignment of liability in the event of a trial is a function of both λ and the parties' litigation expenses as in

$$p = p(\lambda, r^a, r^b) \qquad (1)$$

where $\partial p/\partial \lambda > 0$, $\partial p/\partial r^a > 0$ and $\partial p/\partial r^b < 0$. To simplify further, let A and B's litigation expenses be equal and fixed.[67] If $\lambda = 1$, the precedents are evenly divided in support of the two parties and hence $p = .5$ given that $r^a = r^b$. Similarly, if $\lambda > 1$, the precedents on balance support A and $p >$

---

67. This assumption is consistent with the Rubin and Priest papers since neither author systematically treats litigation expenses as an endogenous variable. In contrast, Goodman analyzes litigation expenditures as the outcome of a noncooperative game, resulting in some modification of Rubin's conclusion regarding the efficiency of the common law.

.5, while if $\lambda < 1$, $p < .5$.[68] We assume that both parties have equal access to the relevant precedents and hence identical and unbiased estimates of p.

Precedents provide information not only on the expected outcome of the current dispute between A and B but also on the likely outcome of similar disputes in the future. This information will in turn affect the allocation of resources across activities. For example, the likely assignment of liability for accident losses in activity j (assuming that transaction costs prohibit negotiation between the parties prior to the accident) will affect an individual's decision whether or not to participate in the activity, the level of his participation, the amount of resources he will allocate to reducing the likelihood and size of damages given participation, and, finally, the number of accidents. S, the present value of the sum of future damages and avoidance costs, will therefore depend in part on expectations of the likely assignment of liability in j. These expectations, in turn, depend on the stock of precedents.

To analyze the survival properties of efficient and inefficient rules, it is necessary first to specify which liability assignment leads to a more efficient resource allocation. Using Calabresi's terminology, suppose that B is the "cheaper cost avoider" in activity j so that S is minimized when liability is assigned to B (and similarly situated defendants in general). Since p is the probability that B is liable, S will tend to decline as p

---

68. This formulation implicitly allows for any biases judges may have in favor of plaintiff or defendant. For example, if judges favor plaintiff, this would reduce the number of precedents favorable to A that are required to make $\lambda = 1$ (holding constant the number favorable to B).

It should be emphasized that λ is not simply a ratio of the number of precedents favoring A and B respectively. Because of differences in the recency and authority of precedents, a smaller number of precedents in A's favor might outweigh a great number favoring B, and this, in our analysis, would result in a λ greater than 1. On the factors which influence the weight of a precedent see William M. Landes & Richard A. Posner, Legal Precedent: A Theoretical and Empirical Analysis, 19 J. Law & Econ. 246 (1976).

increases, as in

$$S = S(p) \qquad (2)$$

where $\partial S/\partial p < 0$. When p increases, the marginal return of, and hence B's expenditures, on damage avoidance rises, since that marginal return is simply the reduction in damages brought about by those expenditures discounted by the probability that B will be held liable. By similar reasoning, when p increases, the marginal return of A's accident-avoidance expenditures falls.[69] This shift in avoidance expenditure between A and B lowers S because of our assumption that B is the cheaper cost avoider.

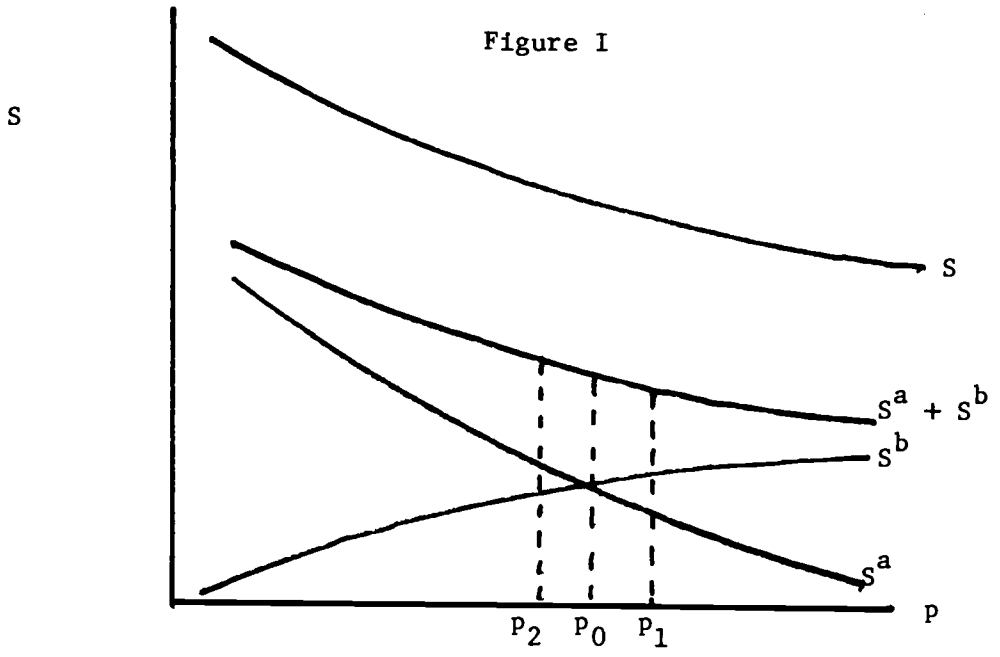
Figure I illustrates both the relationship between S and p and the effect of assuming that the parties have significant and symmetrical on-going interests in activity j. B's future costs (S^b) of accident avoidance increase as p increases because B is more likely to be liable for damages in the future and so will spend more on accident avoidance, but since these costs are more than offset by A's lower future costs (S^a), $S^a + S^b$ decreases as p increases.[70]

---

69. Notice that we are implicitly analyzing the choice between a strict-liability (i.e., B always liable) and a no-liability (B never liable) rule. We do this to simplify the analysis; our conclusions would be unaffected by extending the analysis to other liability rules (e.g., negligence).

70. Rubin, in contrast, assumes that if $p < .5$, A and B act as if A is always liable and therefore S takes the value S(0) for all p's between 0 and $< .5$. Alternatively, if $p > .5$, A and B act as if B is always liable and S equals S(1) for all p's $> .5$. Thus, the S curve would be horizontal at S(0) until $p = .5$; at $p = .5$ it would be discontinuous; and at $p > .5$ it would be horizontal at a lower value of S(1). We assume, on the contrary, that cost-minimizing behavior on the part of plaintiffs and defendants will generate marginal changes in S in response to changes in the likely assignment of liability (p), giving rise to a declining and approximately continuous S function as in Figure I.

Figure I



Some examples will illustrate the relationship between future costs (S) and expected trial outcome:

(1) Let the alternatives in eminent-domain proceedings be that either the government (B) pays or it does not pay the plaintiff (A) for taking his property. In the former instance, landowners will allocate land to its most efficient use, thereby maximizing its present value. To be sure, there will be social costs from raising the tax revenue necessary to compensate landowners in eminent-domain proceedings (e.g., distorting labor-leisure choices and tax-collection costs) but, when balanced against the gain in value from more efficient land use, assignment of liability to the government will probably minimize S. As the likelihood of government liability declines (i.e., as p falls), landowners will begin to discount the expected return from investments in land and at the margin will reduce or alter the timing of these investments (e.g., biasing them toward the present), producing deadweight losses. These deadweight losses, and hence S, will increase as a negative function of p, and will reach their maximum when the

legal rule allows the government to take property without compensation.[71]

(2) When a ship is in distress at sea, exigencies of time may preclude voluntary negotiations between the ship's master and potential rescuers (salvors). Under one possible assignment of liability, the rescuer is never entitled to compensation for his costs. Since this rule will significantly reduce the incentives of rescuers, ships would undertake excessive investments in safety or make other adjustments to avoid hazards that might give rise to losses at sea.[72] If instead rescuers are entitled to salvage awards to cover their expenses, the allocation of resources to safety and rescue will approximate the results that a competitive market would achieve if transactions between rescuees and rescuers were feasible. In terms of Figure I, if p were defined as the probability that a salvor would be compensated (assuming he was successful in his rescue efforts), then, as p increased, the social costs (S) associated with shipping activity would decrease.

(3) B, a professional photographer, contracts with A, to develop B's film. A loses B's film and B sues for damages that include the costs of B's trip to Africa where the photographs were taken. Assuming B is the cheaper cost avoider (perhaps because B could have taken a second roll of film to Africa or explicitly informed A of the value of the film and obtained extra care in exchange for a premium above the usual costs of development), the assignment of liability to A for B's full damages would increase the costs

---

71. The resource-allocation effects of inefficient rules are similar to those of factual errors in the application of efficient rules, analyzed in Richard A. Posner, supra note 66, at 402-05, 452-55.

72. Rescues may not decline to zero, however, since altruism would still be a factor motivating some rescues. For a detailed analysis of the economics of rescue and its application to salvage awards see William M. Landes & Richard A. Posner, Salvors, Finders, Good Samaritans, and Other Rescuers: An Economic Study of Law and Altruism, 7 J. Legal Studies 83 (1978).

of developing compared to assigning liability to B. Thus, the damages and avoidance costs would decline as p increased.

These examples illustrate the proposition that alternative liability rules affect resource allocation. Examples (1) and (2) differ from (3), however, because in (3) the parties had an existing contractual arrangement when the mishap that created damage arose and therefore could have contracted around an inefficient liability rule in advance of the mishap. Thus A, if liable in (3), could insert in its developing contracts an explicit clause limiting liability to the costs of the film. This, in turn, would induce B either to inform A of the value of the film and negotiate a waiver of the limited liability clause or to take several rolls of film with him to Africa. The possibility of recontracting limits the costs of an inefficient legal rule to the costs of contracting around the rule. In the analysis that follows we first assume that no recontracting is possible and then allow for the possibility of recontracting around an inefficient legal rule.

C. No Contractual Relationship Between the Parties

1. Trial-Settlement Decision. To facilitate exposition, assume A is a pedestrian and B a driver, and A is suing B for damages of x arising out of an accident. B is by definition the cheaper cost avoider and therefore the efficient legal rule is for B to be liable. We assume that there are three possible outcomes of A's legal action against B:

state 0	out-of-court settlement--A voluntarily drops suit or settles with B for a sum $\leq$ x
state 1	trial--A wins and B is liable for x; p is the probability of state 1 given a trial
state 2	trial--B wins and is not liable for x; (1 - p) is the probability of state 2 given a trial

We use the following additional notation:

w^a, w^b	A and B's wealth, excluding their share of S from participation in activity j
S_o^a, S_o^b	A and B's future costs (present value of damages plus expenditures to reduce probability and amount of damages) if an out-of-court settlement is reached
S_1^a, S_2^b	A and B's future costs if state 1 occurs
S_2^a, S_2^b	A and B's future costs if state 2 occurs

A's expected wealth if he pursues his claim against B and goes to trial is

$$\bar{w}^a = w^a - r^a + px - (pS_1^a + (1 - p)S_2^a) \tag{3}$$

and if he does not pursue his claim is

$$w_o^a = w^a - S_o^a. \tag{4}$$

Thus, A's net expected gain from going to trial ($\bar{w}^a - w_o^a$) is

$$G^a = px - r^a + [S_o^a - (pS_1^a + (1 - p)S_2^a)] \tag{5}$$

where the expression in brackets denotes the expected change in A's future costs if the current dispute between A and B goes to trial.[73] Observe that

$$S_1^a < S_o^a < S_2^a \tag{6}$$

because if a trial occurs and B is held liable (state 1), the ratio of precedents (λ) will shift marginally in favor of plaintiffs, which in turn will increase the probability that defendants will be held liable for damages in

---

73. We assume that all parties are risk neutral. Risk aversion and risk preference have been incorporated into previous models of the litigation process but to do so here would unnecessarily complicate the analysis.

the next period.[74] Since A is presumably a member of the class of future plaintiffs, the present value of A's future costs (S^a) will decline compared to an out-of-court settlement (state 0) which would involve no legal determination of the dispute and therefore no change in λ.[75] Conversely, if B is successful at trial, λ and p will fall and A's future costs will increase compared to a settlement. Thus, by going to trial litigants are gambling not only on whether they will win and if so how large the amount of damages will be but also on changes in the amount of future damage and avoidance costs brought about by the judgment. The magnitude of the gamble on future costs will depend on the importance of the current decision relative to the existing stock of precedents and on the responsiveness of future costs to a change in precedents.[76]

---

74. Of course, not every litigated judgment creates a precedent. Jury verdicts, and indeed most trial-court judgments, are not published and therefore have little or no precedential significance. Most precedents are generated at the appellate level. This would make no difference if all judgments were appealed, for then we could simply treat the trial and appeal as a single litigation. But not all judgments are appealed. In general, one would expect the appeal rate to be positively related to the stakes in the case and the parties' future stakes (i.e., their interest in the precedential significance of the case). A more elaborate model than the one we use here would recognize the distinction between trial and appeal and would focus on precedent-creating litigation rather than simply litigation as we do. We do not think, however, that the more elaborate model would alter our conclusions materially. It might actually reinforce them by strengthening the positive relationship in our model between future stakes and (precedent-creating) litigation.

75. Initially, we assume that if someone is a plaintiff (defendant) today, he will also be a plaintiff (defendant) in the future. If not, there would be a conflict between a litigant's present and future interest. We remove this restriction later on.

76. Notice that a necessary (and we assume sufficient) condition for A to press his legal claim against B is that A's net expected gain from going to trial exceeds zero. If $G^a < 0$, the gains to A from suing B will be less than the costs of legal action against B and hence A would drop his suit. Unless explicitly stated to the contrary, we assume that $G^a > 0$.

B's expected wealth if a trial takes place is

$$\bar{w}^b = w^b - r^b - px - (pS_1^b + (1 - p)S_2^b) \tag{7}$$

and if A drops his claim is

$$w_o^b = w^b - S_o^b. \tag{8}$$

B's net expected gain (or loss) from a trial compared to A's dropping his claim is

$$G^b = -px - r^b + [S_o^b - (pS_1^b + (1 - p)S_2^b)] \tag{9}$$

where

$$S_1^b > S_o^b > S_2^b . \tag{10}$$

The ordering in (10) is reversed compared to that of (6) because if B wins (state 2) the precedents shift in B's favor and his future costs fall relative to state 0, whereas if B loses (state 2) the precedents shift in A's favor and B's costs rise relative to state 0.

A well-known result of the economic analysis of litigation is that the decision to go to trial or settle out-of-court depends on whether

$$\pi = G^a + G^b \gtrless 0. \tag{11}$$

If $\pi > 0$ (e.g., A expects to gain 10 but B expects to lose 7), B's maximum settlement offer will be less than A's expected gain and a trial will occur. Alternatively, if $\pi < 0$ (e.g., A expects to gain 10 but B expects to lose 15), B's maximum settlement offer (=15) will exceed A's expected trial gain and both parties will be better off with any out-of-court settlement greater than 10 but less than 15.[77]

We are interested in how the decision to go to trial is affected by precedent, since this will reveal how the private decisions of the parties to lawsuits affect the creation of legal rules through the litigation process.

---

77. In both examples, the costs of bargaining to a settlement are assumed to be negligible.

But we approach this question differently from Professor Rubin who as noted earlier, assumes that if A loses (and hence the inefficient liability rule prevails) there is no change in λ but that if A wins (and the efficient rule prevails) the precedents shift in favor of plaintiffs so that in the future the probability that defendants will be liable will exceed .5. Under Rubin's approach, for example, if p equalled .1 but A nevertheless was successful at trial, the earlier precedents--which overwhelmingly favored defendants--would be sharply devalued and in the future p would be > .5 But while the efficient rule would thus replace the inefficient one if A won, the inefficient rule would not be strengthened if he lost. As a result of this asymmetry, when both parties have approximately equal stakes in the future there is never an incentive to litigate an efficient liability rule, but there is an incentive to litigate an inefficient rule anytime the expected present value of the gain from moving to an efficient rule is greater than the litigation costs of the two parties.[78]

---

78. Rubin's model implies that if precedents initially favored B, $S_o^a = S_2^a$ and $S_o^b = S_2^b$ because if B wins (state 2) future costs will not increase for A or decrease for B relative to the no trial situation. The trial-settlement condition is then

$$\pi = -r^a - r^b + p[(S_o^a + S_o^b) - (S_1^a + S_1^b)] \lesseqgtr 0.$$

The term in brackets is positive (assuming symmetrical future stakes) because B is the cheaper cost avoider. Suppose initially that B is not likely to be liable ($p < .5$) but π is positive. Although A is likely to lose (which doesn't alter p in Rubin's analysis), eventually an A litigant will win, shifting liability to B. But since B is now the cheaper cost avoider, if A were to litigate again there would be no gain in efficiency and hence $S_1^a = S_0^a$ and $S_1^b = S_o^b$. Therefore

$$\pi = -r^a \quad r^b + (1 - p)\ [(S_o^a + S_o^b) - (S_2^a + S_2^b)]\ < 0$$

because if B were to win in a trial (state 2), A's and B's costs would be higher than if the case were settled. Thus, once the inefficient rule is overturned in Rubin's model, there is no longer an incentive to litigate any further.

Judge-made law, however, does not change as drastically as Rubin's analysis implies. A single decision favoring an efficient outcome is unlikely to eradicate a hundred prior opinions supporting an inefficient one.[79] Gradual or incremental change is the dominant form of change in a decentralized system of judge-made law. Our analysis, in contrast to Rubin's, assumes that current decisions generate marginal and symmetric changes in precedents[80] -- i.e., an inefficient rule becomes more or less durable depending on whether the outcome of the current trial is or is not in conformity with it.

2. Efficiency Implications. We can rewrite the decision to go to a trial or settle out of court as

$$\pi = -(r^a + r^b) + [S_o^a - (pS_1^a + (1 - p)S_2^a)] + [S_o^b - (pS_1^b + (1 - p)S_2^b)] \gtrless 0, \quad (12)$$

noting that $S_1^a < S_o^a < S_2^a$, $S_1^b > S_o^b > S_2^b$, and for all participants in j that $S_1 < S_o < S_2$. (The latter follows from the assumption that in the aggregate it is less costly to assign liability to B.) These inequalities are illustrated in Figure I, where p_o equals the initial probability of B's liability and p_1 and p_2 ($< p_1$) the subsequent probabilities conditional on the outcome of the current dispute.

---

79. This is especially true, of course, if the decision is at the trial level. Many unappealed trial-court decisions have little, and some zero, precedential value (many trial-court decisions, especially in jury trials, are not even reported). We assume Rubin's analysis--like our own in this part of the paper--is implicitly limited to decisions at the appellate level.

80. Some empirical support for this assumption can be found in our paper on legal precedents, supra note 68. Utilizing citations in federal appellate and U.S. Supreme Court judicial opinions to prior decisions, we estimated a relatively low depreciation rate (4-5 precent) for legal precedent. See id. at 279.

Let

$$\Delta S^a = S^a - S^a > 0 \qquad \Delta S^b = S^b_o - S^b_1 < 0$$
$$\delta^a \Delta S^a = S^a_o - S^a_2 < 0 \qquad \delta^b \Delta S^b = S^b_o - S^b_2 > 0 \tag{13}$$

ΔS^a and ΔS^b denote changes in future costs when B's liability increases, and $\delta \Delta S^a$ and $\delta^b \Delta S^b$ changes in future costs when B's liability decreases. Since δ^a and δ^b are likely to be approximately of the same magnitude, the exposition is simplified if we assume that $\delta^a = \delta^b = \delta$.[81] Since $\delta < 0$, the choice between a trial or settlement now depends on whether

$$\pi = -(r^a + r^b) + (p + (1 - p)\delta) \quad (\Delta S^a + \Delta S^b) \gtrless 0. \tag{14}$$

We can derive the following implications from (14):

(1) If the on-going interests in activity j of both parties are negligible, $\Delta S^a \cong 0$, $S^b \cong 0$, and $\pi < 0$; hence there will be an out-of-court settlement. This is the well-known result that if both parties agree on the expected outcome of a trial (and future costs are ignored), a trial is equivalent to a gamble with a negative expected value, so that risk-neutral and risk-averse persons will always prefer to settle. Since in this example the expected assignment of liability has no effect on the incentive to litigate, efficient and inefficient precedents would have the same survival value.

---

81. When A wins and liability marginally shifts to B, part (but not all) of A's reduction in future costs will be shifted to B. As a first approximation, therefore, we can write

$$\Delta S^b / \Delta S^a \cong -k$$

where $0 < k < 1$. By analogous reasoning, if B wins, part of B's gain involves a transfer of costs to A such that

$$\delta^b \Delta S^b / \delta^a \Delta S^a \cong -k.$$

This, of course, implies that $\delta^a \cong \delta^b$.

(2) Suppose that both parties have substantial and approximately equal future stakes in activity j and that $\delta = -1$.[82] The limiting case would be where activity j is specific to A and B and hence $S^a + S^b = S$. In general we would expect

$$S_1^a + S_1^b < S_o^a + S_o^b < S_2^a + S_2^b, \qquad (15)$$

because if A wins (state 1), liability will shift towards B, the cheaper cost avoider, and in the aggregate $S_1 < S_o$, whereas if B wins (state 2) liability will shift towards A, the more expensive cost avoider, and $S_2 > S_o$. Assuming the parties have sufficient and symmetrical stakes, their combined future costs ($S_1^a + S_1^b$, $S_o^a + S_o^b$, and $S_2^a + S_2^b$) will move in the same direction as the aggregate future costs (S_1, S_o and S_2) and (15) will hold (see Figure I). The ordering in (15) implies that:

$$\Delta S^a + \Delta S^b > 0. \qquad (16)$$

Since $\delta = -1$, π is necessarily negative if $p < .5$. Alternatively, if $p > .5$, π is more likely to be positive the greater is p, the smaller the r's, and the larger the combined gain ($\Delta S^a + \Delta S^b$) of the two parties from the marginal shift in liability to the cheaper cost avoider. This suggests that, given strong and symmetric stakes:

(a) the disputes most likely to go to trial are those where, on balance, the existing precedents already favor an efficient outcome (i.e., the probability that B is liable is greater than .5); and

(b) the disputes least likely to go to trial are those where the existing precedents favor an inefficient allocation ($p < .5$).

---

82. $\delta = -1$ implies that the precedential significance of winning and of losing a trial is approximately the same, and hence the corresponding decrease and increase in future costs would be of equal magnitude. The latter implication always holds for the case of linear cost functions, but holds only for marginal changes in costs in the case of nonlinear cost functions.

Over time, (a) and (b) imply a tendency towards an increase in the average efficiency of legal precedents since efficient precedents become more so while there is no obvious change in inefficient precedents.[83] If all legal rules tended to favor efficiency (p's > .5), then the degree of efficiency would be strengthened over time provided the combined savings in costs weighted by 2p - 1 exceeded the costs of going to trial.

This result has an interesting interpretation. When the legal rules that govern particular disputes are uncertain, disputes will never be litigated if an inefficient outcome is more likely than an efficient one ($p < .5$.)[84] Litigation will (assuming the parties have roughly symmetrical future stakes) be confined to disputes in which the existing legal rules on balance favor efficient outcomes ($p > .5$). Moreover, this process is cumulative because, ohter things being equal (e.g., in the absence of legislative intervention), litigation will tend to occur in those areas of law where the rules are becoming progressively more efficient.

This analysis suggests that inefficient rules will lie dormant but at the same time will affect behavior--i.e., people will be guided by these rules in their allocation of resources to damage avoidance but will not litigate when disputes arise. This surprising result, however, is probably an artifact of our assumption that the parties always agree on p. In fact, different estimates by disputants of the likely outcome of litigation are probably a very important, indeed dominant, cause of litigation. And such

---

83. Given $p > .5$, the stock of precedents would become continually more favorable to A until in the limit p approached 1. Although we have not done so here, we could build in the assumption that as p rises $S^a + S^b$ falls at a decreasing rate which would tend to more than offset the gain from going to trial as p rises. Eventually, a long-run equilibrium would be reached at $p < 1$.

84. This assumes that parties agree on p and x. If the parties disagree on these variables, then litigation will arise even though the parties have no stakes in the future and the existing legal rules favor inefficiency. We take up these points later.

differences are probably a negative function of the age of the relevant precedents, since it is more difficult to infer from an old precedent how the court will decide a current dispute than from a recent precedent. If this is correct, the very dormancy of an area of legal disputes will raise the litigation rate in that area. This is, indeed, a direct implication of our analysis elsewhere of how precedents are created in a system in which the author of a precedent is not directly compensated for his efforts.[85]

To summarize, in the case of strong and symmetric future stakes, we agree with Rubin--though we reach his result by a slightly different route--that the common law system of rule creation is biased in favor of efficiency not necessarily because of any systematic judicial preference for efficient outcomes but as a function of the sample of cases that are likely to be litigated in a system where the decision to sue or litigate and the investment in litigation are private. Moreover, this conclusion holds whether plaintiffs or defendants are the cheaper cost avoiders and also where parties do not know whether they will be plaintiffs or defendants in future disputes.[86]

---

85. See William M. Landes & Richard A. Posner, supra note 68, at 271-72.

86. (i) If plaintiffs rather than defendants were the cheaper cost avoiders, so that resource allocation would be improved were defendants not liable, the disputes most likely to go to trial would be those where precedents now favored defendants (i.e., $p < .5$). For, when plaintiffs are the cheaper cost avoiders

$$S_1 > S_o > S_2$$

and

$$S_1^a + S_1^b > S_o^a + S_o^b > S_2^a + S_2^b$$

then

$$\Delta S^a + \Delta S^b < 0.$$

Assuming that $\delta = -1$, a necessary condition (see (14)) for a trial is $p < .5$ or that precedents favor B, the higher-cost avoider. Since defendants will on average win more trials than they lose, precedents will gradually shift over time in favor of defendants, leading to a more efficient allocation of resources. In contrast, if inefficient precedents predominate ($p > .5$),

(3) But now suppose the parties have asymmetrical stakes in future costs. A natural way to define asymmetry is to order the parties' combined costs according to which party has the greater stake in S. Thus, if A had the greater future stake $S_1^a + S_1^b < S_o^a + S_o^b < S_2^a + S_2^b$, while if B had the greater future stake, $S_1^a + S_1^b > S_o^a + S_o^b > S_2^a + S_2^b$. The former implies that $\Delta S^a + \Delta S^b > 0$ and the latter that $\Delta S^a + \Delta S^b < 0$. In terms of the trial-settlement choice (equation (14)), a necessary condition for a trial is that the odds favor the party with the greater stake. If that were A, a trial would occur only if $p > .5$, and if B, only if $p < .5$. Since by assumption B is the cheaper cost avoider, the precedents will shift in favor of or against efficient liability rules depending on whether A or B has the greater stake. In the latter case (i.e., B has the greater stake) we have an example of a tendency to strengthen an inefficient legal rule, but the tendency is probably weak. A is unlikely to press his claim (since G^a is likely to be < 0) and

---

86. (cont'd) disputes will not be litigated. Therefore inefficient rules will persist but will not become more entrenched over time.

(ii) Imagine that both parties continue to have substantial future stakes in activity j, but neither knows whether he will be a plaintiff or a defendant in the future. For example, suppose A is suing B for patent infringement suit because both are heavily involved in inventive activity. Or, in our earlier example of rescue at sea, both parties may be engaged in shipping and equally likely to be either a salvor or a victim of distress in the future. (A more common example would be the contract litigation of a business firm. Since the firm both buys inputs and sells outputs, it is performer-payee with regard to some of its contracts and payor with regard to others. However, this is an example of a preexisting contractual relationship between the parties and is therefore considered in subpart D infra.) If we denote by q the probability that A or B will be a future plaintiff, and assume $\delta = -1$, then (14) becomes

$$\pi = -(r^a + r^b) + (2p - 1)\, 2(q\Delta S^a + (1 - q)\Delta S^b) \gtrless 0$$

If each party is equally likely to be a plaintiff or defendant, $q = .5$ and the above is identical to (14). Thus, our analysis of the relative survival properties of efficient and inefficient precedents depends only on both parties' having significant future interests in activity j and not on each party's ability to identify the position (plaintiff or defendant) he will occupy in the future.

therefore a trial will not take place.[87]

(4) The most restrictive assumption of our model is that both parties have future stakes. At first glance, it might seem that this assumption would be satisfied in only a limited subset of cases, mainly involving business firms and government agencies. The individual accident victim, for example, whose probability of being involved in litigation over future accidents is very small, would not have a significant stake in the precedential significance of the decision determining liability for the accident. This is true, and yet the legal system does contain various devices for bringing future stakes to bear in current litigation. One is the amicus curiae (friend of the court) brief, which enables an individual or other entity who anticipates that the decision in a case to which he is not a party will be a precedent affecting his own activities to participate in the litigation (though in a highly circumscribed way). Also, trade and other associations (e.g., the NAACP), public-interest law firms, and other organizations of firms or individuals having future stakes in precedent can sometimes participate directly or indirectly (through class actions, "test cases," etc.) in litigation in which the nominal parties do not themselves have significant future stakes.

The problem discussed here is an aspect of the general externality problem in private precedent production examined in Part I of this paper. The social benefits of precedent are not limited to the parties to the case--indeed, if those parties have no interest in future disputes for which the

---

87. It follows from equation (5) that G^a falls and is more likely to be negative as p decreases. Notice, however, that if A were the cheaper cost avoider and had the greater stake, litigation would be more likely to occur when $p > .5$. Since G^a is also more likely to be positive, litigation will be more likely when precedents on balance favor efficient legal rules.

decision in their case might constitute a precedent they derive zero private benefits from helping to create a precedent. It is only if they have such a future interest, or if others who do are somehow represented in the litigation, that the social benefits of precedent can be privately appropriated.

D. Contractual Relationship Between the Parties

When it is feasible to contract around a legal rule, the parties will do so if the costs are less than those of the legal rule.[88] Let S^m denote the costs of a voluntary or market assignment of liability where S^m equals the present value of the sum of contracting costs, future damages, and avoidance costs. Figure II illustrates the relationship between $S^a + S^b$ and S^m.

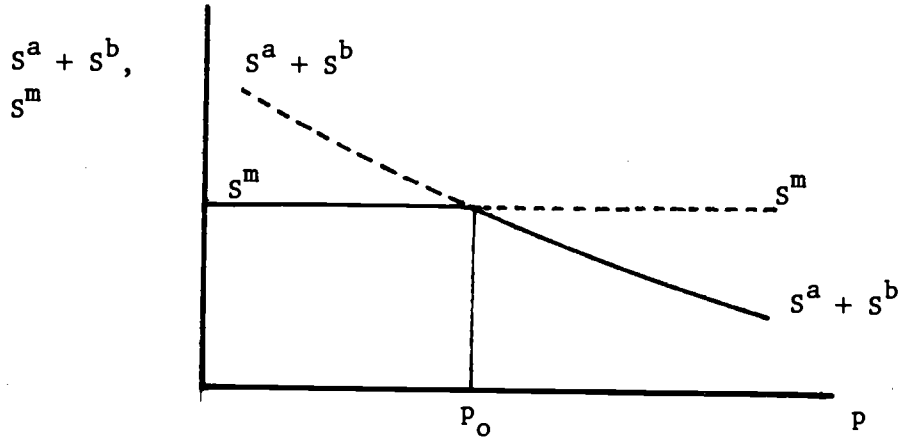


Figure II

Notice that S^m is constant and independent of the probability of B's liability under the legal rule. If $p < p_o$ in Figure II, the parties would choose to contract around the legal rule, assigning liability to B, the cheaper cost avoider, because $S^m < S^a + S^b$. In contrast, if $p > p_o$, the contracting solution would be more costly and the parties would prefer the

88. We assume that the agreement between the parties is enforceable. If not, the analysis of subpart C would apply here even though a contractual relationship exists between the two parties.

legal rule even though the probability of B's liability is less than unity. In the latter instance, the costs of contracting, which would have enabled the parties to assign liability to B with a probability near unity, are greater than the gains associated with the increased certainty of B's future liability.

Suppose initially that the existing legal precedents assign liability to B with a probability p_o, resulting in joint future costs of $S_o^a + S_o^b$. If A were to go to trial and lose (state 2), the precedents would shift marginally in B's favor but A and B's future costs would not change because they would opt for the market or contract alternative where $S^m = S_o^a + S_o^b$.[89] However, if A wins and liability shifts towards B (state 1) both parties gain in the aggregate since $S_1^a + S_1^b < S_o^a + S_o^b$. Thus, the truncation of the cost function at p_o implies that $S_o^a = S_2^a$ and $S_o^b = S_2^b$, and hence the trial-settlement condition becomes

$$\pi = -(r^a + r^b) + p_o\ (\Delta S^a + \Delta S^b) \gtrless 0 \qquad (17)$$

The possibility of contracting around the legal liability rule has some interesting implications regarding the trial-settlement decision and the evolution of legal rules:

(1) Since π is greater when the parties are able to contract around the legal rule (compare equation (17) to (14)), the likelihood of a trial is also greater.[90] This does not necessarily imply a greater tendency for legal rules to become more efficient over time. That depends critically on

---

89. Although there is no change in aggregate costs, the distribution of costs between A and B may differ depending on whether the parties choose the market or legal assignment of liability. We assume, however, that the distribution of future costs is not affected by this choice.

90. π is greater in (17) than (14) by the amount $(\Delta S^a + \Delta S^b)\ (1-p_o)$ assuming $\delta = -1$ in (14).

the value of p_o (p_o being defined in general as the probability where $S^m = S^a + S^b$). If $p_o > .5$ and the parties go to trial, A will win on average and the precedents will shift in favor of A. As a result either:

(a) a more efficient legal rule will replace the market assignment of liability (assuming at p_o or slightly less than p_o, the parties were contracting around the legal rule and incurring future costs of S^m); or

(b) the existing legal rule will tend to become more efficient over time (assuming at p_o or slightly greater than p_o, the parties were utilizing the legal liability rule and incurring future costs of $S^a_o + S^b_o$).

In contrast, if $p_o < .5$, the parties may still go to trial when contracting is available. If so, the legal rule will tend to become less efficient (at least in the neighborhood of p_o), but will not affect resource allocation or future costs because A and B will find it cheaper to contract around the less efficient legal rule and incur costs of S^m. This result is equivalent to the noncontracting solution (subpart C) because π was negative when $p < .5$, and hence there was no tendency for the legal rule to become more inefficient over time. In sum, the option of contracting tends to improve the efficiency of legal rules when $p_o > .5$ but has no effect on these rules when $p_o < .5$.

(2) When the stakes are asymmetrical, the option to contract plays a crucial role. We showed earlier that in the absence of this option a dispute would tend to go to trial if A had the greater future stake and $p > .5$ or if B had the greater future stake $p < .5$. In the latter case, an inefficient rule could become more so over time. But not if contracting is possible, for then, although B would be able to shift the precedents in his favor by going to trial, he would derive no gain from this shift. Since B is in a contractual relationship with A, competition among B's will generate the least costly method of assigning liability. And at $p_o < .5$, S^m is the least costly

solution. For example, imagine that B's product is sometimes defective and accidents to A can result. Although both A and B can spend resources to reduce the frequency of these accidents, if B is the cheaper cost avoider, the full costs of B's product will be lower when B assumes liability than under a legal rule that assigns liability to B with a probability less than .5. Thus, if contracting is available, there is no longer a tendency for the legal rule of liability to evolve in favor of the party with the larger future stake when the rule is less efficient than the market alternative.

(3) Suppose the private cost of obtaining and disseminating information declines, lowering the cost of using the market to contract around legal rules. For a particular legal rule, S^m would fall and p_o shift to the right (say to p_o') in Figure II, with two effects on the overall efficiency of legal rules. First, it would now be cheaper for parties to substitute the market for the legal assignment of liability in those instances where the legal rules assigned liability to B with a probability between p_o and p_o'. Second, if $p_o' > .5$, the incentive of some parties to go to trial who would not have done so before would increase. (The reason is that the potential increase in joint future costs vanishes since the parties have the option, if B wins, of contracting around the more costly legal rule.) Since $p_o' > .5$, the precedents would shift gradually in A's favor, reducing future costs and increasing the efficiency of the particular legal rule. Put differently, there is a complementarity between market and legal efficiency--the greater the efficiency of the market, the greater the tendency for a subset of legal rules to become more efficient.

(4) Our analysis of contracting has been limited to the case where the

probability of B's being held liable is in the neighborhood of p_o.[91] If instead the existing precedents generated a p significantly greater than p_o, the market or contract alternative would not be relevant because the future costs of this alternative would be substantially greater than that of the legal liability rule. In this case the analysis would be identical to the noncontractual model of subpart C. On the other hand, if the existing precedents yielded a p significantly lower than p_o, the legal rule would not be a relevant alternative. In that case, $\Delta S^a + \Delta S^b = 0$, π would be less than 0, and the parties would settle out of court.

E. Litigation Expenditures

Suppose each party determines his litigation expenditures r^a and r^b by maximizing his net gain from going to trial (G^a and G^b), taking as given the litigation expenditures of the other party. Rewriting the first-order conditions yields

$$-\frac{\partial p/\partial r^a}{\partial p/\partial r^b} = \frac{x + (S_1^b - S_2^b)}{x + (S_2^a - S_1^a)} . \qquad (18)$$

In equilibrium, therefore, the ratio of A to B's marginal product of litigation expenditures will be equal to the ratio of B to A's gain from winning

---

91. We say "neighborhood" because if p were slightly less than p_o it would be possible, if A went to trial and won, to move to the declining part of the $S^a + S^b$ curve in Figure II. Similarly, if p were slightly greater than p_o, the movement, if A lost, would be to the S^m curve.

the trial.[92] Since A and B are able to purchase equivalent or equally efficient inputs of legal services at similar costs, we assume that $\partial p/\partial r^a$ and $\partial p/\partial r^b$ are equal but of opposite sign when $r^a = r^b$. Thus, if A and B gain equal amounts from winning the trial (e.g., no future costs and A gains x if state 1 occurs while B does not lose [i.e., gains] x if state 2 occurs), (18) will hold only if A and B's expenditures are equal. Moreover, this result is independent of the extent to which precedents initially favor A relative to B. Put differently, we expect parties with comparable stakes in the litigation to spend similar amounts on the case in question.[93] Alternatively, if A's savings in future costs from winning were greater than B's savings (i.e., $S_2^a - S_1^a > S_1^b - S_2^b$), (18) would hold only if $\partial p/\partial r^a < -\partial p/\partial r^b$, which requires, from the assumption of diminishing marginal products, that A spend more than B. This implies that the probability of A's winning and B's being held liable would be greater compared to our earlier analysis (parts C and D) where the

---

92. Note that $\partial p/\partial r^a > 0$ and $\partial p/\partial r^b < 0$. We also assume that $\partial^2 p/\partial r^{a^2} < 0$, $\partial^2 p/\partial r^{b^2} > 0$, and $\partial^2 p/\partial r^a \partial r^b = 0$. The first two conditions assure that the levels of r^a and r^b that satisfy the first-order conditions yield maximum values of G^a and G^b respectively. The third condition simplifies the exposition by eliminating the following type of behavior. A initially picks a level of r^a assuming a particular value of r^b. B's expenditures, however, differ from A's expectations. This leads A to adjust his expenditures, leading B to adjust, leading A to adjust again, and so forth. By assuming that the cross-partial derivative is zero, A's marginal product depends only on A's expenditures and B's marginal physical product only on B's expenditures, and thus the optimal level of each party's expenditures is independent of the other's expenditures.

93. Deviations from risk neutrality, however, would alter this result because then the initial wealth levels of the parties would influence litigation expenditures via the effect of wealth on the derivative of expected utility with respect to litigation expenditures.

litigation expenditures of the two parties were assumed to be equal.[94]

The explicit introduction of litigation expenditures leads to some small changes in our earlier analysis.

1. Symmetrical Stakes. Assuming that B is the cheaper cost avoider and that the parties have symmetrical and sufficient stakes in the future (i.e., their combined future costs move in the same direction as aggregate future costs), a shift in liability towards B will produce greater future costs savings to A than losses to B (i.e., $S_2^a - S_1^a > S_1^b - S_2^b$)[95] so that the right-hand-side of equation (18) will be less than one. In equilibrium, therefore, $r^a > r^b$ and the resulting p will be greater than that of the analysis in parts C and D where r^a was assumed to be equal to r^b. Our earlier analysis also showed that a necessary condition for a trial to occur was that $p > .5$ (i.e., that, on balance, the precedents favored the more efficient outcome of holding B liable). This conclusion still holds but with a slight modification. Earlier we measured p at the point where $r^a = r^b$. Now, however, if p were slightly less than .5 (assuming $r^a = r^b$) because the precedents were slightly in favor

---

94. This analysis ignores, of course, the strong incentive of the parties, by cooperating, jointly to reduce their litigation expenses. For example, if A and B are initially spending r_0^a and r_0^b, then by reducing their expenditures to, say, $.5r_0^a$ and $.5r_0^b$, p will be unchanged but both A and B will have a greater expected wealth. A reduction in expenditures may result from an agreement between the parties to limit the number of issues litigated, or from a stipulation not to dispute certain facts relevant to the litigation. These devices are in fact used, but not so frequently as to require us to abandon the analysis in the text. A further point is that even if both parties stipulate to certain facts or agree to limit the number of issues, r^a and r^b may not fall if the parties concentrate greater litigation resources on the remaining issues.

95. From equation (15) we have $S_1^a + S_1^b < S_2^a + S_2^b$ which implies $S_2^a - S_1^a > S_1^b - S_2^b$.

of B, A's additional expenditures on litigation (holding B's constant) could push the probability above .5. Thus, when litigation expenditures are endogenous, some legal rules that initially favored an inefficient outcome (at $r^a = r^b$) would now be litigated; and given that $p > .5$ when parties are spending optimally on litigation, we would observe a movement over time towards the development of precedents favoring efficiency.

2. Asymmetrical Stakes. When litigation expenditures are endogenous, the tendency discussed earlier of inefficient legal rule to become even more inefficient when B has the greater future stake is reinforced because B invests more in litigation than in the earlier analysis, thereby increasing his probability of winning the current trial and the likelihood that future precedents will move in his favor.[96]

Consider, for example, the case of railroad crossing accidents. Assume that the railroad is both the cheaper cost avoider (B) and the party with the greater future stakes. If p were somewhat greater than .5, the railroad might by investing sufficient amounts in litigation be able to reduce it below .5. In this event the railroad would be eager for a trial and over time the precedents would tend to build up in its favor. On the other hand, if the rules were highly efficient to begin with, so that the probability of the railroad's being held liable were significantly greater than .5, optimal litigation expenditures by the railroad would not be sufficient to lower the

---

96. Our prior analysis of a contractual relationship between the parties is largely unaffected by treating litigation expenditures as a decision variable. In the case of symmetrical stakes, p will rise in response to an increase in litigation expenditures of A relative to B, which will tend to increase the value of π in equation (17) compared to the case of $r^a = r^b$. This in turn will increase the likelihood of a trial and the development of more efficient legal rules. If B has the greater future stake, there would be no incentive for B to expend litigation resources to shift the legal rule in B's favor (to the detriment of efficiency) because competition among B's would lead the parties to assign liability in the cheapest way either by adhering to the legal rule or by contracting around it.

probability below .5. The railroad would choose to settle with the plaintiff and the legal rule would remain highly efficient. To be sure, the railroad would obtain better terms in the settlement by the implicit threat of outspending the plaintiff in a trial because the minimum amount the plaintiff will accept to settle will decline as p falls. In sum, whether the legal rule will evolve in favor of the dominant party will depend on the initial efficiency of the rule and on the responsiveness of the expected trial outcome to the litigation expenditures of the party who has the greater future stakes.

Since, insofar as judicial outcomes are influenced by the relative investments of the parties in litigating, and these investments in turn by the relative stakes of the parties in the precedents created by those outcomes, the tendency of the common law process to generate efficient rules is weaker in areas where the parties' stakes are asymmetrical than in areas where they are symmetrical, an interesting empirical question is whether, for example, the rules of tort law relating to railroad crossing accidents are less efficient than those relating to collision between ships; we would predict a greater tendency toward efficient rules in the latter area.[97] As a parallel example, consider a government agency like the Federal Trade Commission that has a greater interest in precedent than most of its opponents. The optimal strategy for the FTC might be to "pick on" small firms having relatively small present (and perhaps no future) stakes in the litigation.[98] The FTC's optimal expenditure of resources in such a case would be very high relative to the defendant's, so that if the case should happen to be tried (obviously, the

---

97. A similar, though not explicitly economic, argument appears in Marc Galanter, Why the "Haves" Come Out Ahead: Speculations on the Limits of Legal Change, 8 Law & Soc. Rev. 96 (1974).

98. See Richard A. Posner, The Behavior of Administrative Agencies, 1 J. Legal Studies 305, 311 (1972).

probability of a trial rather than settlement would be small because of the small stakes of the defendant) the chances of a favorable outcome for the FTC would be enhanced. In this way, insofar as a systematic tendency for the FTC to invest more heavily than its opponents would influence the evolution of legal rules in its favor, we could hardly expect FTC law to be moving in the direction of greater efficiency.

F. The Priest Model

Both Rubin's analysis and our criticism and expansion of his analysis assume that at least one of the parties has a significant interest in the precedential significance of the decision determining liability for the accident. Professor Priest, in contrast, attempts to show that the common law will tend to become more efficient over time even though neither party has an interest in the precedential significance of the decision.

In the absence of such an interest, disputes go to trial only if the parties have divergent estimates of the expected trial outcome, not identical estimates as assumed in Rubin's and in our model. For example, assuming that A and B agree on x (the damages) and that future stakes are zero, the parties will go to trial if

$$\pi = x(p^a - p^b) - (r^a + r^b) > 0 \qquad (19)$$

where p^a and p^b are A and B's estimates of the probability of B's liability in the event of a trial. Mutual optimism ($p^a > p^b$), therefore, is a necessary condition for $\pi > 0$ and, thus, for a trial to take place. Observe that the greater the current stakes (x) are, assuming $p^a > p^b$, the more likely it is that $x(p^a - p^b)$ will exceed the parties' combined litigation costs and hence the more likely a trial is.[99] The positive relationship between the size of

99. This result must be qualified because litigation expenditures are positively related to x; for example, A's optimal expenditures require $(\partial p^a/\partial r^a)x - 1 = 0$ which implies that as x increases so will r^a. We assume here that litigation expenditures increase at a proportionately slower rate than x, so that π will increase as x increases.

x and the incentive to go to trial is a critical component of Priest's analysis. For example, consider the effect on x of an inefficient legal rule. Since such a rule assigns liability to the higher-cost avoider, fewer resources are allocated to accident avoidance and there are more accidents and more damages per accident.[100] Hence the less efficient the legal rules is, the more likely a trial is. Given this result and Priest's assumption that in each period the judiciary announces a constant proportion (=a) of efficient urles, Priest shows that the proportion of efficient rules will tend to rise over time and that the equilibrium proportion will be greater, the greater the value of a and the greater the difference in litigation rates between inefficient and efficient legal rules.[101]

Although Priest's analysis is formally correct, it contains, we believe, a conceptual error. Decision according to precedent is ignored in Priest's model. In each period, judges produce efficient and inefficient decisions in the proportions of a and 1-a respectively, regardless of the type of cases before them and the degree to which the body of precedent built up from prior cases favors an efficient or inefficient outcome. Thus, the probability that an efficient rule will be announced in a particular case today is independent of how similar cases were decided in earlier periods. Although one can speak of a body of legal rules in the aggregate, the central feature of a legal system that decides according to rules or precedents is missing from Priest's model: in Priest's model, prior decisions provide no information or guide to the likely outcome of the current dispute.

We assume, instead, that the likely outcome (p) of a trial depends on the extent to which precedents favor A relative to B (λ in equation (1)) and on the

---

100. We assume that expenditures on accident avoidance reduce both the probability of an accident and the severity of the accident if it occurs.

101. Priest's conclusion also requires the assumption that in the base period the proportion of efficient legal rules is less than or equal to a.

litigation expenditures of the two parties.[102] Ignoring the determination of optimal litigation expenditures (assume that $r^a = r^b$), if the precedents favor B (i.e., $\lambda < 1$) the probability will be less than .5 that B, the cheaper cost avoider, will be held liable and hence less than .5 that the outcome of the trial will be a judicial opinion in the direction of greater efficiency. Alternatively, if the precedents favor A (i.e., $\lambda > 1$), the probability of B's liability will be greater than .5, and now the outcome of the trial will shift the precedents toward greater efficiency.

Suppose there are a large number of current disputes, some of which may go to trial, categorizable into n classes ($i = 1, \ldots, n$). Assume that within each class there is a set of precedents, some favoring efficiency and others inefficiency, applicable only to disputes in that class, so that to each class of disputes we can assign a particular λ_i. Let the n classes now be separated into two subclasses, I and II. In I precedents tend to favor an inefficient liability rule; i.e., $\lambda_i < 1$ for all classes of disputes in I. In II precedents tend to favor an efficient liability rule; i.e., $\lambda_i > 1$ for all classes in II. Since most class I disputes that go to trial will result in a victory for B, inefficient precedents will tend to accumulate relative to efficient ones in this class and hence the liability rules governing class I activities will tend to become even less efficient over time. On the other hand, class II disputes will result in more decisions for A than for B and thus the body of precedent affecting class II activities will become more efficient. However, because the trial rate is higher in class I than II (because of the higher average x in I), legal rules

---

102. One might object to our analysis because we (like Priest) impose a mechanical decision rule on judges -- that judges follow precedent and do not have or express any preferences for efficient or inefficient rules. But this objection is superficial. Adherence to precedent is consistent with judges having preferences and following them, provided that in the aggregate these preferences do not change, or only change slowly, over time. If this condition is satisfied, as it generally is in our legal system, there will be a strong positive correlation between past and current decisions and we will tend to observe adherence to precedent.

in class I will become more inefficient than the rules in II become efficient. As a consequence, the average efficiency of legal rules will tend to decline over time. This result is the opposite of Priest's conclusion that legal rules tend to become more efficient over time.

Figure III illustrates our analysis. Suppose initially that we have a uniform frequency distribution of λ's, ranging from 0 to 1, with a mean of .5. The mean λ of class I disputes is .25 and the mean of class II disputes is .75. Over time the mean of class I will fall while the mean of class II will rise. This is illustrated by the new frequency

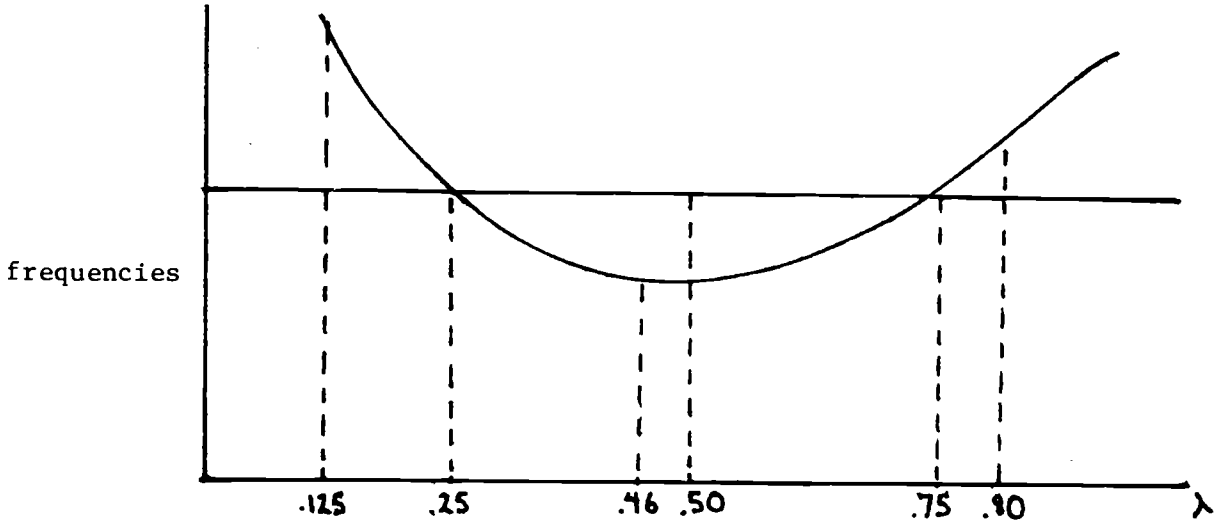


Figure III

distribution in which the means are now .125 and .80 for class I and II disputes respectively. The new overall mean, however, is .46, indicating an overall reduction in average efficiency of liability rules.

It is implicit in our analysis (as Priest's comments on a previous draft of this paper helped us to realize) that once the role of divergent estimates of the litigation outcome are incorporated into our model, the effect is to generate strong tendencies to push any λ in the neighborhood of 1 (i.e., where

the weight of precedent is equally balanced between the efficient and the inefficient rule) toward either infinity or zero. Intermediate points involve uncertainty as to how a court applying the precedents will decide the current case, and this uncertainty generates divergent estimates by the parties and so increases the likelihood of litigation. Assuming (and admittedly we go outside our model in doing so) that appellate courts have an incentive to reduce conflict among precedents and product a harmonious, consistent rule,[103] the litigation resulting from uncertainty generated by inconsistent precedents will tend to eliminate that inconsistency by overruling, limiting, reinterpreting, or disregarding either the precedents favoring A or the precedents favoring B. If the initial balance of precedents favors B (the inefficient result), then chances are that the litigation process will eliminate the precedents favoring A. The tendency of litigation to move an initially efficient rule in the direction of $\lambda = \infty$ will be less pronounced, because with lower damages (smaller x) there will be less litigation. A challenge to future research is to build a formal model incorporating this insight.

CONCLUSION

This paper has examined the question whether adjudication can be viewed as a private good, i.e., one whose optimal level will be generated in a free market. Part I focused on private courts, noting their limitations as institutions for dispute resolution and rule creation but also stressing the important role that the private court, in its various manifestations, has played both historically and today. Part II discussed a recent literature which has argued that the rules generated in the public court system, in areas of the law where the parties to

---

103. This incentive might be quite weak in a system where judges are paid out of litigant fees, as suggested in Part I. However, in a system of salaried judges, an incentive to produce a consistent rule can be derived from an influence-maximization model of judicial behavior that we have employed elsewhere (see William M. Landes & Richard A. Posner, supra note 68, at 272-73), since the effect of a rule in directing activity is clearly greater the greater the probability that it, and not its opposite, will actually be applied in the event a dispute arises and is litigated.

litigation are private individuals or firms and the rules of law are judge-made, are the efficient products of purely private inputs. Our analysis suggests that this literature has overstated the tendency of a common law system to produce efficient rules, although areas can be identified where such a tendency can indeed be predicted on economic grounds.

Viewed as a contribution to the emergent literature on the positive economic theory of law, our finding that the public courts do not automatically generate efficient rules is disappointing, since it leaves unexplained the mechanisms by which such rules emerge as they seem to have done in a number of the areas of Anglo-American judge-made law.

However, our other major finding, that the practices and law governing private adjudication appear to be strongly influenced by economic considerations and explicable in economic terms, is evidence that economic theory has a major role to play in explaining fundamental features of the legal system.