December 1995

Fee Shifting and Predictability of Law

Keith N. Hylton

Follow this and additional works at: https://scholarship.kentlaw.iit.edu/cklawreview

Recommended Citation
Available at: https://scholarship.kentlaw.iit.edu/cklawreview/vol71/iss2/3
FEE SHIFTING AND PREDICTABILITY OF LAW

KEITH N. HYLTON*

I. INTRODUCTION

Lawyers are trained to distinguish between substance and procedure. The substantive law is comprised of standards, such as the Learned Hand formula of negligence, that are used to determine whether a violation of the law has occurred. Procedural rules, on the other hand, determine whether and under what conditions a party can bring suit or be joined in an ongoing suit, the conditions under which a decision may be appealed, the burden of proof, and the allocation of legal expenses.¹

Virtually no part of legal training involves the study of the influence of procedure on substantive law. This is odd, and even astonishing, because it should be obvious that procedural rules have an enormous influence on the development of legal doctrine. Would anyone believe that the substance of the law would not be affected by a procedural rule that had the effect of barring access to the courts to virtually all claims? Consider, for example, a rule requiring tort plaintiffs to pay ten times the legal expenses of the defendant. Would anyone think such a rule had no implications for the future development of tort doctrine?

Not only are procedural rules important because they set the terms on which parties may take advantage of the legal entitlements granted by substantive rules, but also procedural rules have effects on the margin. One can think of procedural rules as gates that determine the flow of claims into courts, which are constantly reshaping legal doctrine; open the gate a little wider, and the emerging doctrine takes a different shape.

* I thank Andy Chen, Richard Craswell, Tony D’Amato, John Donohue, Mark Grady, David Haddock, Gary Lawson, Peter Seigelman, and Janet Spragens for helpful suggestions.

¹ This list is not intended to be exhaustive. It should include the rules governing summary disposition of claims, the discovery process, class actions, and the jurisdiction of courts and probably should include rules governing the admissibility of evidence.

There is the further difficulty that the distinction between substance and procedure is hard to maintain in some areas of litigation. Consider, for example, contract law. If one assumes that courts generally attempt to enforce agreements, then it seems to follow that much of contract law, if not all, should be labeled "procedural." Indeed, one could make a plausible argument that the expression "contract law" is an oxymoron. I do not deal with this problem below.

427
The point of this Article is to examine the marginal influence of a certain set of procedural rules on the substance of law. Substance, however, can mean many things. To narrow the scope, I will focus on predictability of law. I will also discuss general, speculative implications for the substance of law. Because the term procedure can also cover an almost infinite variety of rules, I will focus on a subset: litigation cost allocation rules. I will consider the American rule, under which each party pays its own legal expenses, the British rule, under which the losing party pays the legal expenses of the winner, and the Pro-plaintiff rule, under which the defendant pays the legal expenses of the prevailing plaintiff. The American rule has been referred to as "no-way" fee shifting, the British as "two-way" shifting, and the Pro-plaintiff as "one-way" shifting.

This Article arrives at two conclusions. First, I reject J. Robert S. Prichard's claim that the British rule is superior to the American in its influence on predictability. Second, I argue that the best rule on predictability grounds is the Pro-plaintiff.

The claim that the British rule enhances predictability is more than a simple statement about the economics of the litigation process. For it provides a rigorous—indeed, mathematical—foundation for the

2. Alternatively, this Article focuses on uncertainty in law. Since the subject of legal uncertainty is quite broad, I will not attempt to cite all of the relevant works in this area. I have been strongly influenced by the treatment in the following: Sheldon Amos, The Science of Law (1877); Oliver Wendell Holmes, The Common Law (1881); Bruno Leoni, Freedom and the Law (1961). For more recent work in this area, see Anthony D'Amato, Legal Uncertainty, 71 Cal. L. Rev. 1 (1983); William M. Landes & Richard A. Posner, Legal Precedent: A Theoretical and Empirical Analysis, 19 J.L. & Econ. 249 (1976). There is a recent positivist branch of this literature, analyzing the influence of legal error on incentives to comply with the law and to litigate. See, e.g., Keith N. Hylton, Costly Litigation and Legal Error Under Negligence, 6 J.L. Econ. & Org. 433 (1990); Louis Kaplow, The Value of Accuracy in Adjudication: An Economic Analysis, 23 J. Legal Stud. 307 (1994).


4. One-way fee shifting in favor of the plaintiff has been defended in other papers, though on different grounds. This is the first article to argue that the Pro-plaintiff rule minimizes uncertainty in the law.

Previous defenses of the Pro-plaintiff rule have either focused on incentive effects or "justice" arguments. Perhaps the most impressive presentation of the case on incentive grounds appears in John Leubsdorf, Recovering Attorney Fees as Damages, 38 Rutgers L. Rev. 439 (1986). In an earlier article, I demonstrated that the Pro-plaintiff rule is superior to other rules in enhancing incentives to comply with the law and encouraging settlement. See Keith N. Hylton, Fee Shifting and Incentives to Comply with the Law, 46 Vand. L. Rev. 1069 (1993) [hereinafter Hylton, Fee Shifting].

broader thesis, advanced by Atiyah and Summers,⁵ that British law is more formal and American law more substantive. While this thesis may be true, its relation to the predictability argument is not as straightforward as Prichard's article suggests.

Since the treatment below is fairly technical in places, it may help to provide a summary here. Prichard argued that the British rule enhances predictability because of certain "selection effects" that both encourage claims firmly grounded in law and discourage claims weakly grounded in law. I demonstrate that the selection effects relied on by Prichard are considerably more complicated and that it is unclear whether greater predictability is a byproduct of the British rule. The second part of my thesis questions the premises of the predictability-formality linkage implicit in the work of Prichard and that of Atiyah and Summers. These analyses ignore an important type of predictability, which I call "dynamic." When one takes this into account, the case for the Pro-plaintiff rule seems strong.

This Article is set out as follows. Section II.A defines the concept of predictability relied on in this Article. Section II.B presents an overview of the literature on the influence of litigation on legal uncertainty. Part III presents an analysis of the influence of various litigation cost-allocation rules on the predictability of law. Part IV presents the normative case for the Pro-plaintiff rule.

II. PRELIMINARIES

A. Predictability of Law

I use the terms predictability and certainty interchangeably, but the sense in which I use these terms needs to be defined. Two types of predictability (or, alternatively, sources of uncertainty) can be identified. One is static, which refers to the ease with which the meaning of a legal rule can be comprehended. Statutory interpretation is concerned with resolving this type of uncertainty, but it is only one of several areas in which the static predictability problem arises. The classic example of the interpretation problem is Blackstone's description of a statute that forbade a layman to "lay hands" on a priest.⁶ The

---

⁵ P.S. Atiyah & Robert S. Summers, Form and Substance in Anglo-American Law: A Comparative Study of Legal Reasoning, Legal Theory and Legal Institutions (1987). One can think of Prichard's argument as an application of Occam's razor to that of Atiyah and Summers. While Atiyah and Summers rely on a large list of possible explanations for the differences between American and British law, Prichard is able to show that they can almost all be explained by the incentives created by financing and fee-shifting rules.

courts had held that the statute applied to an individual who hurt a priest with a weapon, even though a literal construction would not support this result. Blackstone defended this interpretation as consistent with the popular use of the term "lay hands," which he understood to mean "harm." Bentham later seized on the example as an illustration of the uselessness of Blackstone's approach to statutory construction. As part of a general argument that the only useful approach to interpretation is one that aims to understand the end of the statute, Bentham asserted that the purpose of the statute was to prevent local justice authorities from apprehending priests; otherwise, it would have made little sense to restrict its application to laymen.

The Blackstone-Bentham dispute continues today with little progress. The example is useful here as an illustration of the number of plausible interpretations of a relatively simple statute. In this example, there are at least three: (1) a layman cannot put his hands on a priest; (2) a layman cannot hurt a priest; and (3) a layman cannot arrest a priest. The different interpretations implicit in the dispute are not only plausible, but each is supported by an independent theory of interpretation.

Again, statutory interpretation is only one example of the static predictability problem. A slight variation of the problem arises with common-law adjudication, where rules are announced that must then be interpreted in later court decisions. The other large category of static predictability problems, which for convenience I will label "rule identification," involves determining the laws that apply to a given situation.

7. Id. at 59.
10. The first, a layman cannot put his hands on a priest, is supported by a literal approach. It may seem an obtuse way to interpret statutes, but there is a sensible argument, and it runs as follows. Courts do not have the time to figure out what legislators were thinking about when they wrote the law in a certain way. The best way to give them an incentive to be clear is to read the statutes literally. Then legislators will be on notice that a statute that seems silly if applied literally will be applied in a silly manner. Although modern textualists have a number of reasons for promoting a more-or-less literal approach to interpretation, the notion of using literalism or textualism as a method of controlling legislators is one of them. See Frank H. Easterbrook, The Role of Original Intent in Statutory Construction, 11 Harv. J.L. & Pub. Pol'y 59, 66 (1988). The argument for the second interpretation is provided in Blackstone and is intuitive. The third interpretation results from Bentham’s view that one should start with attempting to determine the aim or the primary end of the statute. For a recent critical view of both sides of the Blackstone-Bentham dispute, see Eskridge, supra note 8.
transaction or injury. A common method of examining law students is to present them with problems of this sort. They are given a transac-
tion and asked to state the laws that apply or given an injury and asked to determine the laws violated. Much of the everyday work of lawyers involves listening to the stories of would-be plaintiffs, deter-
moving whether laws apply, and converting the stories into claims with standing in court. This is the reverse of statutory construction, which involves starting with a statute and determining the instances in which it applies; rule identification involves starting with an injury or trans-
action and determining the laws that apply.

The second type of predictability is dynamic. I refer to the de-
gree to which parties who understand the meaning of a rule today can be sure that meaning will remain valid tomorrow. There may be no static uncertainty, but if the law is likely to be changed in the near future, to a form that is unpredictable, then the current rule cannot be relied upon in the formation of long-term contracts and plans. Simi-
larly, if conditions change in a way that gives the law, as it is framed today, a different meaning tomorrow, then it is unreliable as a basis for plans.

This definition of dynamic predictability must be distinguished from the notion of stare decisis, and particularly that of an absolute rule of stare decisis. The two concepts are not necessarily the same. Stare decisis is a sensible policy because there must be some perma-
nence or stability in legal interpretations. Otherwise, one creates in-
centives for parties to challenge every interpretation, and the result is the complete absence of law. An absolute rule of stare decisis may seem desirable as a way of providing for finality in the interpretation of a rule, and it may seem to be the surest way of establishing rules that can be relied upon in making plans. An absolute rule of stare decisis may not be consistent, however, with the goal of enhancing dynamic predictability. A law that has been rendered obsolete or useless as the result of a change in conditions may not be a solid basis on which to plan. Knowing that the rule serves no useful purpose, parties that are subject to it may ignore it, under the assumption that only an unreasonable person would attempt to force them to comply

11. The distinction between static and dynamic uncertainty, though expressed in different terms, is emphasized in LEONI, supra note 2, at 77-96.

12. One could say this is the theme of Calabresi’s recent study of common-law evolution. See GUIDO CALABRESI, A COMMON LAW FOR THE AGE OF STATUTES (1982). Because of legisla-
tive inertia, statutes tend to stay on the books long after they become obsolete. In response to this problem, Calabresi argues that courts should openly review and modify statutes that have become obsolete.
with an obsolete or useless law. And, if the rule is truly useless, parties who were never aware of it are unlikely to learn of its existence, except through the employment of attorneys. The existence of a significant number of useless, though stable, rules could very well be a worrisome source of legal risk, or "man-traps" in Bentham's words.

There is some tension between the notions of static and dynamic predictability. On the most general level, the tension is implied by the definitions. An effort to eliminate all sources of static uncertainty in a law would require the legislature to indicate precisely the boundary of the statute, which would require a number of precise descriptions of instances in which the statute applies. Unless stated otherwise, the implication would be that the omitted instances are those in which the statute did not apply, but such an approach is likely to reduce the degree to which the statute's application is predictable in the dynamic sense. It is unlikely that a legislature would foresee all of the instances in which the statute should apply. Disputes will arise later that seem obvious instances in which the statute should apply, but the legislature failed to foresee them. In addition, tastes and technology will change in unforeseeable ways over time, presenting new scenarios that seem to require application of the statute.

13. For example, a rule regulating the use of horse-drawn carriages might have served a good purpose ninety years ago, but be of little use today. If the rule remains on the books, however, it may provide credibility to an otherwise frivolous lawsuit or it may give discretion to justice authorities to prosecute certain individuals who unknowingly violate the statute. As a second example, consider the Connecticut statute banning contraceptive devices struck down by the Supreme Court in Griswold v. Connecticut, 381 U.S. 479, 485 (1965). Calabresi notes that by the time of the Court's decision the statute had become a legislative anachronism, inconsistent with practice, common law, and other statutes in Connecticut. See Calabresi, supra note 12, at 8. Suppose the rule were still valid law. How would it affect behavior? Would adherence to stare decisis enhance the certainty of law in Connecticut governing the use of contraceptives? I believe that the law would sit like an undiscovered land mine in a farmer's field, or for some litigants it would be the equivalent of the occasional piano falling out of a window. Given its practical obsolescence, only a handful of Connecticut lawyers would probably have some inkling of the existence of the statute, and discovery of the rule would be nothing less than the surprise of a lifetime to the defendant prosecuted for violating it.

14. Bentham, supra note 9, at 164. Bentham regarded the law governing the interpretation of deeds as a large source of unintuitive, bizarre, and in some cases useless rules that were nevertheless fairly stable. Id. at 163-64. Their contribution to predictability is summarized by Bentham as follows:

I speak of those . . . [rules] . . . according to which a disposition in itself acknowledged to be legal is not to be carried into effect for want of certain legal words: words which were it not for that express appointment, which men are never told of, and which a man must be a sorcerer to divine, would appear to carry quite another meaning than what is given them. Yes, I must again repeat it, which men are never told of: for it is not to call it, telling men in general, the burying here and there a few obscure notices in a lay-stall of legal gibberish, which no man rakes into that is not paid. When the danger spreads itself the kingdom over, I will not call it warning for a chance lawyer to get up in Westminster Hall and whisper to his fellows 'there are man-traps here'.

Id.
Tax and securities laws present some of the famous examples of the tension between dynamic and static forms of legal uncertainty, though the tension is general and not at all unique to these areas of law. Tax attorneys and specialists frequently call on legislatures to make the laws more predictable by specifying precisely the manners of tax minimization that are prohibited and not prohibited. Yet, every effort to clarify tax legislation produces counterefforts to avoid taxation while complying with the highly specific letter of the code.\textsuperscript{15} Another example of the trade-off between static and dynamic predictability is observed in the regulation of insider trading. Securities traders generally prefer to have the uncertainty surrounding the scope of the insider-trading prohibition reduced, and this could be accomplished by setting out a precise definition of the acts that fall under the label “insider trading.”\textsuperscript{16} Securities regulators are, however, reluctant to provide this level of specificity, precisely because it would generate new forms of insider trading that fall outside of the enumerated prohibitions.\textsuperscript{17}

Bruno Leoni presented an alternative perspective on the tension between dynamic and static predictability.\textsuperscript{18} Countering Bentham’s claim that statutory law was preferable to common law on static predictability grounds, Leoni argued that statutory law was inferior on dynamic predictability grounds.\textsuperscript{19} A system of statutory law presents a set of rules that can be changed by a shift of perhaps only one vote. A common-law system, on the other hand, was not subject to shifts in voting patterns and, because it was relatively immune to such shifts,

\textsuperscript{15} See generally, Bayless Manning, Hyperlexis and the Law of Conservation of Ambiguity: Thoughts on Section 385, 36 Tax Law. 9 (1982). Manning’s thesis, which has generated a healthy debate among tax scholars, is that efforts to clarify the United States Internal Revenue Code do not reduce the total amount of uncertainty, they merely shift uncertainty from one provision or clause to another. Id. at 11.

\textsuperscript{16} For a general overview of the difficulties of defining insider trading, see Robert C. Clark, Corporate Law 320-26, 347-56 (1986).

\textsuperscript{17} One peculiar aspect of these examples is that the tension between static and dynamic predictability in tax and securities laws is due in large part to technological change and to efforts on the part of the regulated groups to avoid compliance with the statutes. Put another way, there is a “cat and mouse” game observed in these areas that contributes mightily to the uncertainty problem. The “cat and mouse” game is not observed in all areas of law. For example, in tort law, there is relatively little effort on the part of regulated groups to structure their transactions in a way that avoids compliance (and at the same time conceals that avoidance). That does not mean that the tension does not exist in tort and other areas, however. The tension would surely be observed in tort law if courts attempted to provide a precise definition of negligence in some area of activity.

\textsuperscript{18} Leoni, supra note 2, at 77-111.

\textsuperscript{19} Id. at 97-113.
provided little incentive for individuals to adopt a self-interested reform campaign.

Predictable law, then, achieves a high degree of static and dynamic predictability (or alternatively minimizes static and dynamic uncertainty). Although there is no measuring system that permits us to say how predictable law should be in order to be satisfactory, the concept of predictability can be used to compare legal systems, or systems of procedural rules. Just as common- and statutory-law regimes can be ranked according to the predictability criterion—with, according to Leoni, statutory law probably performing best on static grounds, and common law on dynamic grounds—systems of procedural rules can also be ranked within the common-law regime.

B. The Influence of Procedure on Predictability: A Review of the Literature

Bentham said a great deal about the predictability of law and procedural rules, but it seems not to have occurred to him that a simple tinkering with the procedure might have done something to alleviate the problems he identified. Of course, Bentham's view is entirely consistent because the problems, as he saw them, were far too large to be solved by a simple modification of procedural rules. Unpredictability was an inherent feature of the common-law system in his view. The only solution was to replace the common law with a system in which statutes framed in simple words were administered by judges with very little input by lawyers.20 Implicit in this is the view that procedural rules have little influence, on the margin, on the predictability of law.

One has to push forward to Oliver Wendell Holmes to find the first statement of a theory of the influence of procedure on the predictability of law, and even here the theory has little to say about alternative procedural rules.21 The model presented by Holmes is important, however, because it encompasses important features of some recent contributions to the literature.

Holmes argued that, although the outcome of a case in court was uncertain, the law itself tended toward increasing predictability over time.22 This was true, in Holmes's view, in both the static and dy-

21. See Holmes, supra note 2, 111-29.
22. Id.
namic senses. In terms of static predictability, the law moved toward certainty along the following lines. The early cases dealt with extreme propositions and simple applications of the rule. Later cases involved increasingly difficult boundary issues. After they were settled in early cases, the extreme propositions and simple applications were no longer subjects of dispute. Any case involving a simple application would settle because the parties could predict the outcome. The only disputes that were litigated would be those in which the answer was unclear and difficult to predict on the basis of earlier decisions. As those cases were decided, however, less and less remained of the areas in which the law's application was unpredictable.

In the dynamic sense, the law moved toward predictability because changes in tastes and technology presented new issues that would have to be resolved by a jury. Thus, the system permitted the law to change in response to changing conditions, which enhances predictability because one expects the law to change in response to changing conditions. Because juries drew on the common sense, habits, and norms of the community, however, the new laws that were generated could not be called entirely unpredictable. Indeed, because the new laws were reasonable in light of the jurors' experiences, which were presumably the same as those of the litigants, one could say they were foreseeable.

Surprisingly, the jury plays an important role in Holmes's argument, and it is perhaps the only procedural feature in the discussion.

23. For example, an early case might ask whether it was negligence for a pedestrian to cross a railroad track without first looking in the direction from which the train usually arrives. A later case might ask whether it was negligence for a pedestrian to cross a railroad track, even though he looked in the usual direction, when he failed to notice that the tracks for incoming and outgoing trains had been switched.

24. It may help to consider a visual picture of this process: imagine an iceberg that is continually replaced so that it maintains its size. The "revealed common law," which consists of the stock of new court decisions, is the visible portion of the iceberg above water. The litigation process resolves uncertainty by exposing new parts of the iceberg to the visible surface. This process does not lead to a reduction in the total amount of uncertainty because the iceberg, by assumption, is infinitely large. Each decision, however, reduces uncertainty by resolving some dispute concerning the contour of the iceberg below the water surface.

Suppose the iceberg is of finite size? Then eventually all contours of its body are exposed to the surface, at which point there is no more uncertainty. As a description of the litigation process, this image is consistent with the notion, attributed to Langdell, that the law consists of a fixed set of basic maxims or principles and that by studying case law one can come to an understanding of the entire set of maxims upon which all legal decisions are based. See Thomas C. Grey, Langdell's Orthodoxy, 45 U. Pitt. L. Rev. 1 (1983). In this view, new decisions reduce uncertainty only to the extent that they provide a clearer view of the underlying maxims. Of course, it is unlikely that new decisions make a great contribution to the resolution of uncertainty, under this theory, because one could presumably gain a clear view of the underlying maxims by studying old decisions.
The jury, however, is not necessary for the contribution to dynamic predictability that Holmes attributes to it. Dynamic predictability under Holmes's model requires that the court draw on the experiences of the community when confronting new applications of the law arising with changes in technology or tastes. That function could be carried out by a judge, provided enough evidence is presented. Indeed, in order for the jury's role to be constructive in Holmes's theory, quite a lot of power must be exercised by the judge in determining which questions should be given to the jury to decide. For if every question were given to the jury, it would be unlikely that any rules would develop with the permanency and consistency that one expects of law.  

In any event, the key implication of Holmes's model is that the law tends toward predictability as a result of the litigation process. Easy cases are decided early in the evolution of a legal doctrine, establishing the rules that the vast majority of people comply with on a daily basis. Later, the easy cases are seldom the subject of legal disputes, because the parties, aware of the established rule, settle their disputes quickly. The difficult questions of application are decided as they arise, in a manner that takes expectations, habits, and norms of the community into account.

The model described by Holmes has been a central part of economic analyses of the predictability of law. For example, Landes's and Posner's analysis, which argues that law moves toward increasing predictability, is in important respects identical to the model articulated by Holmes. The analysis of George Priest and Benjamin Klein, probably the most important recent contribution to the study of the influence of legal procedure on predictability of law, also shares many important features with Holmes's model.

According to the Priest-Klein model, only the disputes whose outcome in court is most uncertain are pursued to the point of a judgment. The disputes whose outcome in court is predictable are settled. This has two implications. On an empirical level, one should expect

25. The system in which all questions are assigned to the jury is equivalent to one in which decisions have very little or no precedential value. Such a system can be thought of as a "Markov process" in which alternative rules replace each other randomly. For an examination of such a process, see Robert Cooter & Lewis Kornhauser, Can Litigation Improve the Law Without the Help of Judges?, 9 J. LEGAL STUD. 139, 145-50 (1980) (demonstrating, among other things, that if each of several alternative rules is litigated with positive probability, no single rule will prevail in the long run).


the win rate of plaintiffs at trial to be roughly fifty percent, since the disputes pursued to the point of judgment are as uncertain as coin tosses. Second, on a more general level, the Priest-Klein analysis implies that the vast bulk of law is unaffected by ordinary litigation. The established rules affect behavior and influence the terms of settled disputes, but are seldom the subject of new litigation. Hence, a social scientist who claimed that a recent string of seemingly pro-plaintiff or pro-defendant court decisions reflected a fundamental shift in legal doctrine could easily be wrong. It is this second implication that was almost entirely anticipated by Holmes’s analysis.

Although Priest and Klein provided empirical support for the first implication, it is controversial because the empirical support is questionable and, more importantly, it suggests that the law may not move toward increasing predictability as a result of litigation. There is considerable overlap between these criticisms of the Priest-Klein empirical hypothesis (the fifty percent win-rate rule), but I will discuss them separately.

The empirical support is questionable because later researchers have shown that the fifty percent win-rate prediction does not hold with respect to many subject matters of litigation. Eisenberg has shown, for example, that medical malpractice and products liability disputes consistently exhibit plaintiff win rates well below fifty percent. A study by Gross and Syverud suggests that as one disaggregates subject areas, for example, from torts to tort-malpractice and so on, one finds less support for the fifty percent rule.

More generally, the fifty percent rule is troubling because it contradicts Holmes’s proposition that litigation tends to push the law toward greater predictability. If the fifty percent rule is correct, then the law is a random walk. One step to the left is equally likely to be followed by a step to the right or another step to the left. Of course, if the law is a random walk, where does it go? The answer is nowhere. A basic result of probability theory is that over the long term the expected deviation of a random walk from a straight path is zero.

30. Interestingly, in an article published four years before the Priest-Klein study, George Priest discusses the random-walk implication and expresses a pessimistic view of theories of the content of common law. See George L. Priest, Selective Characteristics of Litigation, 9 J. LEGAL STUD. 399 (1980).
31. Why? Recall that the expected frequency of “heads” resulting from the toss of an equally balanced coin is fifty percent. The random walk is like a coin toss in which “heads =
Thus, the Priest-Klein fifty percent rule implies that the law does not move at all over the long run. If it moves toward greater fairness or greater efficiency, it does so by chance and is likely to move away if one waits long enough. It also follows that the law does not move as the result of litigation toward increasing predictability. Certainly, static predictability is unlikely to result if the Priest-Klein rule is valid. An interpretive approach to rules that follows a random path is unlikely to enhance the clarity of a rule. Dynamic predictability is also unlikely because, as tastes and technology change, the litigation process fails to respond to these changes in a manner that is consistent with either existing rules or expectations.

Moreover, it follows from the Priest-Klein fifty percent rule that procedural rules have relatively little influence on the predictability of law. Under the hypothesis, disputes are litigated because of equally balanced two-sided uncertainty. A change in procedural rules might change the type of case filed, but the influence of litigation on predictability would remain the same. Thus, surprisingly, the Priest-Klein model provides a rigorous version of Bentham’s view of the common-law process.

Note that I am referring to the implications of the Priest-Klein fifty percent rule. The more general hypothesis that cases whose outcomes are predictable tend to settle more frequently than those with unpredictable outcomes is not inconsistent with Holmes’s model; indeed, it is a central claim of Holmes’s model. The empirical evidence does not refute this prediction. In addition, one could argue that the Priest-Klein model assumes a world in which the legal standard is fixed, i.e., where there is no legal change. Hence, it is perfectly consis-

---

32. See, e.g., Keith N. Hylton, Asymmetric Information and the Selection of Disputes for Litigation, 21 J. LEGAL STUD. 187, 188 (1993) [hereinafter Hylton, Asymmetric Information].

33. I am referring to the fifty percent rule in particular. To be more precise, the Priest-Klein fifty percent rule implies that the predictability of law is not affected by minor changes in procedural rules, such as a change in the fee-shifting rule. In addition, it implies that, if the law is crystal clear at date zero, it will meander randomly, but in the long term return to its firm rule. On the other hand, if the law is unclear at date zero, the law will meander randomly and fail to move in the direction of greater clarity. This implication is strange in light of Priest’s earlier claim that common law tends to generate economically efficient rules. See George L. Priest, The Common Law Process and the Selection of Efficient Rules, 6 J. LEGAL STUD. 65, 65-66 (1977).

34. In the area of employment discrimination, the predictable cases tend to settle, while the unpredictable ones are litigated. This holds even though win rates are much less than fifty percent in this area. Peter Siegelman & John J. Donohue III, The Selection of Employment Discrimination Disputes for Litigation: Using Business Cycle Effects to Test the Priest/Klein Hypothesis, 24 J. LEGAL STUD. 427 (1995).
tent that the model implies the path taken by new decisions is a random walk. But this argument essentially limits the fifty percent rule to a very small set of disputes, which is surely not the interpretation suggested by the Priest-Klein article. One could also argue that the Priest-Klein model allows for plaintiff win rates that deviate significantly from fifty percent when the parties have asymmetric stakes.35 There are, however, so few areas of litigation in which fifty percent win rates are observed;36 one would have to say, to remain consistent with their model, that almost all litigation involves parties with asymmetric stakes. But that is also hard to square with the empirical evidence.37

The foregoing suggests that the predictability thesis would be in poor shape if the fifty percent rule of Priest and Klein were valid. The empirical evidence suggests, however, that it is not valid, as a general rule. In related work, I have developed a model that suggests that where there is one-sided uncertainty, or asymmetric information, win rates will diverge significantly from the fifty percent prediction.38 The theory is consistent with the low plaintiff win rates observed in medical malpractice and products liability litigation, which are areas in which defendants generally have an advantage in predicting the outcome of a trial. The Priest-Klein result is observed as a special case in which two-sided uncertainty is present.

To this point I have argued that the early literature has implications for the influence of procedural rules on predictability. The first direct examination of this question was carried out by J. Robert S. Prichard. Taking advantage of the contributions of Landes,39 Posner,40 Gould,41 and especially Shavell42 Prichard argued that the

35. This was part of their original argument. See Priest & Klein, supra note 27, at 24-29.
36. See Eisenberg, supra note 28, at 357-58.
37. See Hylton, Asymmetric Information, supra note 32, at 201-02.
American rule tended to encourage novel legal claims while the British rule discouraged such claims. As a result, the British system seemed to be characterized by a higher degree of predictability, manifested by stricter observance of precedents, fewer dissenting opinions, and a greater commitment to law as a system of rules. Prichard did not distinguish between static and dynamic predictability, but he seemed to think that both types of predictability were greater under the British procedural system. Although several features contributed to this difference, Prichard emphasized the different cost-allocation rules and the rules governing the financing of lawsuits.

Prichard took no clear position on whether the law is pushed in the direction of greater predictability as a result of litigation. The general thrust of the argument, however, suggests a belief that litigation is capable of enhancing predictability. But, as I have suggested, there are several questions that remain to be addressed. In the remainder of this Article, I will re-examine the influence of cost-shifting rules on the predictability of law and follow with implications for the influence of procedure on substance.

III. INCENTIVE EFFECTS OF ALTERNATIVE COST ALLOCATION RULES AND IMPLICATIONS FOR PREDICTABILITY

Prichard’s examination of the influence of cost-allocation rules on predictability of law relies on Shavell’s analysis of the incentives created by the American and British rules. In this section, I re-examine the implications of Shavell’s work and incorporate those of recent contributions.

A. Litigation Incentive Effects

Cost-allocation rules influence the substance of law through two channels: compliance and selection effects. By compliance effects, I mean effects on incentives of injurers to comply with the law. They are important because an allocation rule that leads to greater compliance reduces the number of injuries that could give rise to legal disputes and affects the characteristics of the typical dispute. This, in turn, affects the issues that will be litigated. For example, if almost all

43. Prichard, supra note 3, at 452 n.2.
44. Id. at 453.
45. Id. at 455.
injurers are complying with the law, then the typical dispute is likely to involve weak or novel legal claims. Conversely, if few injurers are complying with the law, then the pool of injuries that could give rise to disputes will include a large percentage of strong, "core-violation" claims. The mix of novel and core-violation cases flowing through the courts is likely to affect the substance of legal doctrine.47

By selection effects, I refer to influences on the incentives to file suit and to settle. Prichard's analysis was concerned largely with the selection effects of alternative cost-allocation rules. A cost-allocation rule that has the effect of encouraging novel or relatively weak claims is likely to have an impact on the predictability of law and its substance.

Since I am reconsidering the implications of Shavell's analysis, it will help to introduce the terms he used, as well as a few new terms. Assume the probability of an injury to a victim occurring is $p$, if an injurer fails to take care, and $q$, if the injurer takes care. Let the loss to the injured victim be $J$. Let the litigation costs of the plaintiff be $C_p$, and the costs for the defendant be $C_d$. Let $P_d$ be the defendant's estimate of the likelihood of a verdict in favor of the plaintiff, and let $P_p$ be the plaintiff's estimate of the likelihood of a verdict in his favor.

1. Compliance Effects

Recent contributions have demonstrated that cost-allocation rules affect incentives to comply with the law.48 The allocation rule affects the expected liability of an injurer. An allocation rule that forces a potential defendant to pay higher costs increases expected liability. Assuming the injurer is rational, he is more likely to take care or to comply with the relevant law after an increase in expected liability.

The compliance effects of a change in the cost-allocation rule can be divided into two components. One reflects new incentives, for the injurer, created by changing expected liability. The other reflects changes in the victim's incentive to bring suit.

a. Liability Effect

To see the effect of alternative cost-allocation rules on incentives to comply, let us compare the American and British rules. For sim-

47. Hylton, Asymmetric Information, supra note 32, at 204-05.
plicity, assume (1) that every injured victim sues and (2) that the court operates without error, so that the careful defendant expects to win while the careless defendant expects to lose.

Under the American rule, each party bears its own litigation expenses. If the court operates without error, so that the defendant is held liable only if he failed to take care, then under the American rule the expected liability of an injurer who failed to take care is \( pJ + pCd \). The expected liability of an injurer who takes care is \( qCd \). Thus, the increase in expected liability that results from failing to take care is \( pJ + (p-q)Cd \).

Under the British rule, the expected liability of an injurer who fails to take care is \( pJ + pCp + pCd \), and in the case of an injurer who takes care, the expected liability is \( 0 \). Thus, under the British rule, the increase in expected liability that results from failing to take care is the full expected liability of a careless injurer. Because \( (p-q)Cd < p(Cp+Cd) \), the following result may be stated.

**Result 1:** The incentive to take care is greater under the British than under the American rule.\(^49\)

A numerical example may make this discussion clearer. Suppose the probability of an injury to a victim occurring is .75 if an injurer fails to take care and .25 if the injurer takes care. Suppose the loss to the victim is \( \$2 \). Let the litigation costs of the plaintiff and the defendant be \( \$1 \) each. If the court operates without error, under the American rule the expected liability of an injurer who failed to take care is 
\[
(.75)(\$2) + (.75)(\$1) = \$2.25.
\]
The expected liability of an injurer who takes care is 
\[
(.25)(\$0) + (.25)(\$1) = \$0.25.
\]
Thus, the increase in expected liability that results from failing to take care is \( \$2 \). If the cost of taking care is greater than \( \$2 \), the rational injurer will not take care, and conversely.

Under the British rule, the expected liability of an injurer who fails to take care is 
\[
(.75)(\$2) + (.75)(\$2) = \$3.00
\]
and in the case of an injurer who takes care, the expected liability is \( \$0 \). Thus, the increase in expected liability that results from failing to take care is \( \$3 \). Because the increase in liability is greater under the British rule, one should observe greater compliance under it than under the American rule. All injurers for whom the cost of taking

---

\(^49\) Using a similar argument, it is easily demonstrated that the Pro-plaintiff rule is equivalent to the British in this analysis.
care is less than $2.00 take care under the American rule, while all injurers for whom the cost of taking care is less than $3.00 take care under the British rule. It is easy to show that the Pro-plaintiff rule is equivalent to the British rule in this example.

b. Compliance-Selection Feedback

The examples above suggest that the British and Pro-plaintiff rules generate greater compliance than the American rule. The general level of compliance affects incentives to bring suit. If everyone complies with the law, then potential plaintiffs will know that the probability of a victory is small and will therefore be less likely to sue. A decline in the likelihood of a lawsuit, however, reduces the deterrent effect of liability.

An analysis that focuses on injurer incentives is incomplete because it fails to consider the full impact of a change in the cost-allocation rule. Consider, for example, a shift from the American to the British rule. The shift increases expected liability and so leads in the first round to an increase in the level of care, but an increase in care across the board reduces the incentive to bring suit, which dampens the increase in care. The net result is that the level of care does not increase to the full extent suggested by an analysis that focuses only on the injurer's expected liability. Indeed, the level of care may fall in the event that the victim's incentive to sue is substantially lessened.

Result 2: The difference between caretaking incentives under the American rule and under the British rule is dampened by their respective influences on the incentive to bring suit.

The outcome in which compliance incentives are totally undone by the reduction in incentives to sue is not observed in the case of a shift from the American to the Pro-plaintiff rule. Although any increase in compliance reduces incentives to sue, the Pro-plaintiff rule is always at least as encouraging of lawsuits as is the American rule. The British rule is distinguishable because the victim who brings a low probability claim under it faces the (correspondingly) high probability of having to pay the defendant's legal expenses.

2. Selection Effects

I now examine selection effects, which can be subdivided into filing and settlement effects. "Filing effects" includes influences on the plaintiff's incentive to bring a lawsuit. "Settlement effects" includes influences on the parties' incentives to settle a lawsuit that has been filed.

a. Filing Effect

A plaintiff files a lawsuit if the expected award exceeds the expected loss. It follows that under the American rule, the plaintiff will file a lawsuit if the expected value of the claim, \( P_pJ - C_p \), is positive. What gains or losses would a plaintiff under the American rule experience if he were put under the British rule? The plaintiff would gain the value of his litigation expenses multiplied by his perception of the probability of plaintiff victory in court. The plaintiff would lose the value of the defendant's litigation expenses multiplied by his perception of the probability that the plaintiff loses in court. Thus, the plaintiff is more likely to sue under the British rule if he experiences a net gain in the expected value of the lawsuit by being shifted from the American to the British rule, which occurs if \( P_pC_p - (1-P_p)C_d \) is positive. It follows that suit is more likely to be brought under the British rule if \( P_p > C_d/(C_d+C_p) \) or, in words, if the plaintiff's estimate of the likelihood of winning is greater than the defendant's share of the total litigation expense.\(^5^2\) This has two implications:

Result 3: The British rule provides more encouragement to optimistic plaintiffs than does the American rule.

Result 4: Defendants who can credibly commit to large litigation expenses are less likely to be sued under the British than under the American rule.

b. Settlement Effect

The second type of selection effect concerns incentives to litigate rather than settle a dispute. The incentive to litigate depends on the perceived stakes (\( P_pJ \) for the plaintiff and \( P_dJ \) for the defendant) and the total costs of litigating (\( C_d+C_p \)). If the difference in perceived stakes is larger than the total costs, the parties are likely to litigate, because both think the payoff from litigating far exceeds that from settling. The converse holds if the difference in perceived stakes is

\(^{5^2}\) Shavell, supra note 42, at 59.
less than the total costs. By shifting expenses from the winning to the losing party, the British rule increases the perceived stakes and therefore encourages litigation. The basic result is as follows:

**Result 5:** The incentive to litigate rather than settle a dispute is greater under the British than under the American rule.

The Pro-plaintiff and American rules can be subjected to the same analysis. The incentive to bring suit is greater under the Pro-plaintiff rule than under either the American or British rule. Moreover, unlike the American rule, and like the British rule, the Pro-plaintiff rule gives the greatest amount of encouragement to high probability claims. Unlike the British rule, and like the American rule, the Pro-plaintiff rule does not present a special disincentive to low probability claims.

The incentive to litigate rather than settle is greater under the Pro-plaintiff rule than under the American rule. The incentive to litigate is, however, less under the Pro-plaintiff rule than under the British rule. The reason is that the Pro-plaintiff rule enhances the perceived stakes, but not as much as does the British rule.

**B. Implications for Predictability of Law: Re-examination of the British and American Rules**

1. Predictability

Prichard relied on Result 3 in framing the hypothesis that the British rule generated a greater degree of predictability. The greater predictability hypothesis seems plausible because optimistic plaintiffs will tend to be those whose claims are solidly grounded in existing law. Plaintiffs whose claims are weakly grounded in law are relatively discouraged by the British rule.

Thus, Result 3 suggests that the British rule leads to a more stable body of law. The stability results because the claims that are brought tend to be firmly grounded in the law, while novel and other claims weakly grounded in law are discouraged. The law tends to reproduce itself under this system, with relatively few attempted invasions of the established order by new legal claims.

This analysis, however, which focuses on the incentive to bring suit, is incomplete. The low probability claims discouraged by the British rule include novel legal arguments and well-grounded legal claims in which the evidence of violation is weak. Thus, one implication of this analysis is that strong legal claims based on weak evidence are disfavored by the British rule. Many claims of fraud, misrepresen-
tation, or discrimination are likely to fall in this category. The claim of violation may be solidly grounded in law in many of these disputes, and yet it will often be that the facts are ambiguous and permit competing inferences in even the strongest cases in this class.

Result 4 introduces another new consideration: in areas of litigation in which defendants typically devote large sums to their defense, plaintiffs will be relatively reluctant to sue under the British rule. The reason is that the plaintiff who litigates takes the risk that he will have to pay an enormous legal bill for the opposing party. Simply threatening such a charge may be enough to discourage some plaintiffs from filing.

This suggests that wealthy or large corporate defendants who commit themselves to high-priced law firms will tend to gain a greater degree of immunity from suit under the British rule. The immunity results not from the substantive law, but from the penalty that may be imposed on an unsuccessful plaintiff.

Result 4 has important implications for the predictability thesis, for it suggests that some of the claims well grounded in law will actually be discouraged by the British rule—specifically, those against defendants who have committed themselves to large litigation expenses. Put another way, the British rule enables some defendants to strategically preclude lawsuits by committing themselves to high litigation expenses. This suggests that the law governing defendants in this class will tend to be less predictable.

What about Results 1 and 2? They show that compliance is enhanced by the British rule (in comparison to the American), but that the compliance effect is probably not great.53 If the compliance effect were substantial, then the results would provide additional support to the claim that the British system tends to discourage low probability claims. It would suggest that the equilibrium is not due entirely to costs; low probability claims are deterred in part because people assume correctly that others are complying with the law. As Result 2 suggests, however, the compliance effect is probably not an important difference between the American and British rules. The compliance effect becomes an important part of the analysis when we examine the Pro-plaintiff rule in a later section of this Article.

Result 5 undercuts Prichard's claim of greater stability under the British rule. Although novel claims are less likely to be filed under the British rule, those claims that are filed are more likely to be liti-

gated, rather than settled. If courts occasionally err, the result is that the existing law will be occasionally muddied by a holding that cannot be reconciled with it.\textsuperscript{54} Thus, although novel claims are seldom filed, established law is at greater risk of being made to appear less certain under the British rule.\textsuperscript{55}

The implications of Result 5 are somewhat broader when one takes into account informational disparities among litigants. In some instances, either the defendant or the plaintiff may have an informational advantage in litigation, which permits the advantaged party to make a more accurate prediction of the outcome of a trial. One can categorize areas of litigation as follows: (1) where both plaintiff and defendant are uninformed; (2) where both plaintiff and defendant are informed; (3) where the plaintiff is informed and the defendant uninformed; and (4) where the defendant is informed and the plaintiff uninformed.\textsuperscript{56}

When both parties are informed, we should expect settlement. If there is no disagreement as to the outcome of a trial, both parties will realize that a trial would only result in a wasteful expenditure of resources. When neither party is informed, we have the case of equally-balanced, two-sided uncertainty assumed in the Priest-Klein litigation model. The interesting cases are those in which one of the parties enjoys an advantage. If the defendant enjoys an advantage, the guilty defendants will tend to settle their disputes, generating a high percentage of weak claims in the pool of litigated disputes.\textsuperscript{57} Conversely, when plaintiffs are advantaged, weak plaintiffs tend to settle their disputes, generating a relatively high percentage of strong claims in the pool of litigated disputes.\textsuperscript{58}

The upshot is that the predictability implications of the British and American rules depend to some extent on the presence of informational disparity. In the Priest-Klein scenario of equally-balanced, two-sided uncertainty, the American and British rules will generate equal proportions of strong and weak claims pursued to a judgment.

\textsuperscript{54} And this returns us to Cooter and Kornhauser's "Markov process" in which alternative rules replace each other periodically. See Cooter & Kornhauser, \textit{supra} note 25, at 145-50.

\textsuperscript{55} It may be helpful to consider an example. Suppose the probability that a case does not settle is .8 under the British rule and .4 under the American rule, which is suggested by Result 5. Suppose the rate of erroneous reversals of established law is .15 under the American rule and .1 under the British. Then even though the British rule has a lower error rate in this example, the frequency with which a filed claim leads to an erroneous reversal will be greater under the British rule (.06 under the American, .08 under the British).

\textsuperscript{56} Hylton, \textit{Asymmetric Information}, \textit{supra} note 32, at 188-89.

\textsuperscript{57} \textit{Id.} at 189-90.

\textsuperscript{58} \textit{Id.} at 199.
The British rule, as Result 5 indicates, will simply produce more litigation.

If, on the other hand, the defendant has an informational advantage, the British rule will exacerbate the tendency of the litigation process to screen out relatively strong claims. The reason is that with two-way fee shifting, guilty defendants will have a strong incentive to settle and uninformed plaintiffs facing innocent defendants will have a strong incentive to litigate. Why? The informed defendant who is guilty has a strong incentive to settle under the British rule because he knows that a verdict against him will result in liability for both the injury and the litigation expenses of plaintiff. The result is that a larger percentage of guilty defendants settle. Why not all guilty defendants? An outcome in which all guilty defendants settle could not sustain itself as an equilibrium. If all guilty defendants settled, plaintiffs' attorneys would know that the defendants who refused to settle were innocent and would drop their cases. But then guilty defendants would realize that by hanging in longer, mimicking the actions of innocent defendants, they could avoid all liability whatsoever. The likely result is a "pooling" equilibrium in which guilty defendants are indifferent between the options of accepting and rejecting the plaintiff's settlement demand. Introduction of the British rule increases the cost of rejecting, so a larger percentage of guilty defendants accept. As more guilty defendants exit (by settling), however, the signalling value of rejection increases, making it more valuable to the remaining guilties to reject the plaintiff's settlement demand. Equilibrium is restored when the guilty defendant is indifferent between accepting and rejecting, and, since this requires that rejection have a higher signalling value, it follows that the pool of litigating defendants must be "cleaner" in the sense of having fewer guilty defendants.59

Finally, suppose the plaintiff has the informational advantage. Then, for reasons similar to those just stated, the British rule will exacerbate the tendency of the litigation process to screen out relatively weak claims.

These considerations suggest that the tendency of the British rule to encourage relatively strong claims is not uniform across all subject matters of litigation. Pick an area in which defendants are likely to have an informational advantage. In that area, the British rule will generate a higher frequency of litigation of those claims that are filed (and not dropped) and will also tend to encourage the weakest of

59. For further details, see id. at 187-205.
those claims to go to court (as opposed to settlement). This, of course, is quite inconsistent with the thesis that the British rule encourages predictability and stability of law.

The foregoing suggests that Prichard’s preliminary conclusion that law is more predictable under the British system is difficult to support on the basis of theoretical speculation. Though Prichard never distinguished between static and dynamic predictability, his arguments concern the former. It is not at all clear, however, that the British system enhances static predictability. The British system enables defendants who commit to high expenses to virtually bar some high probability claims, which weakens static predictability. In addition, the British rule encourages litigation rather than settlement of claims that are filed, which implies courts will tend to revisit established doctrines at a higher frequency under the British than under the American rule. The process of constantly revisiting established doctrines enhances the risk that some will be overturned occasionally as the result of error. Finally, in those areas of litigation in which defendants have an informational advantage, the British rule will tend to encourage litigation of relatively weak claims.

In terms of dynamic predictability, it is unlikely that the British rule is preferable to the American. Technology and tastes change over time, and predictability therefore requires a system that is capable of reshaping itself to meet the expectations generated by new conditions. The process of reshaping is enhanced by the court’s consideration of novel claims. A procedural system, such as the British, that discourages such claims is not well suited to this task.

2. Substance

The implications of the American and British rules for substantive differences in the law were largely explained by Prichard. The previous section suggests, however, that Occam’s razor could have been applied with even greater force than in Prichard’s discussion: it may be possible to attribute many of the substantive law differences identified by Prichard to the different incentives created by the British and American cost-allocation rules.60 In addition, the incentives created by the cost-allocation rules have some substantive implications that are inconsistent with Prichard’s thesis.

60. Prichard relied on rules governing fee shifting, the financing of litigation, and the bringing of class action suits as determinants of the differences in substantive law in America and Britain. I am suggesting here that many of the differences described by Prichard could be attributed to the fee-shifting rules alone.
Prichard attributes certain differences in the British and American law to the greater ease with which group litigation can be brought in America. But many of the substantive law differences Prichard attributes to class action and similar procedural rules (e.g., derivative suits) could be explained by the different incentives created by the British and American cost-allocation rules.

Consider, for example, Prichard's observation that American corporate law is more detailed and "developed" than British corporate law. He attributes this largely to the incentives provided for derivative suits in America. The relative underdevelopment of British corporate law could also be explained, however, by the different cost-allocation rules. The British rule discourages low probability claims, including those well grounded in law but supported by ambiguous evidence. The negligence and fraud claims typically involved in shareholder derivative actions in America are often based on ambiguous evidence. Indeed, the problem is inherent in the legal definition of a corporation. It is a standard form contract or a series of short, standard form contracts connected with a number of complicated implicit agreements. The definitions of diligence and negligence are seldom, if ever, explicitly stated. A negligence claim against the directors and officers, therefore, will probably involve ambiguous evidence. They will generally be perceived to be low probability claims and thus less attractive under the British rule.

A second explanation attributable to the cost-allocation rules is that the British rule enables corporations to strategically preclude claims by committing themselves to high litigation expenses. If corporations act in this manner, then only plaintiffs with the strongest claims will file suit against them. The result is a legal system in which the law governing such corporations seems underdeveloped in comparison to American law.

The incentive to strategically preclude lawsuits may also explain the observation that British law firms have not been subjected to the kind of competitive cost-cutting pressures applied by large corporate clients in America. The greater survival in Britain of practices, such as the sharing of profits on a relatively equal basis, lock-step increasing of compensation, and committing to hierarchy based on seniority

61. Prichard, supra note 3, at 463 (quoting L.C.B. Gower, Some Contrasts Between British and American Corporation Law, 69 HARV. L. REV. 1369 (1956)).
62. This is consistent with anecdotal reports that American tort litigation "dwarfs" that in Britain. See "On Trial," A Survey of the Legal Profession, ECONOMIST, July 18, 1992, at 11.
63. Id. at 5-6.
within the firm,\textsuperscript{64} may be a reflection not of culture, but of the willingness of corporate clients to pay a luxury tax. Clients are willing to pay the tax because the act of committing themselves to expensive law firms has a deterrent effect among potential plaintiffs.

To see the new substantive implications suggested by this analysis, consider medical malpractice litigation. It is an area in which the negligence rule has been in place for years. There is little novel, at the level of theory, about a medical malpractice claim. The difficult and novel questions are factual, and "fact-law" issues are generated by the appearance of new technology.\textsuperscript{65} Thus, many claims that are well grounded in negligence doctrine are likely to be discouraged under the British rule because the evidence of a violation is not easily established. This prediction is consistent with the Snyder and Hughes study of Florida medical malpractice claims, which found a higher percentage of cases dropped under the British rule.\textsuperscript{66}

For those cases that are filed and not dropped, the question is whether the British rule tends to encourage the litigation of relatively strong claims. This is unlikely because medical malpractice is an area in which defendants have an informational advantage in litigation. Thus, medical malpractice litigation under the British rule will tend to involve a higher percentage of relatively weak claims. Whether the typical claim is weaker than that brought under the American rule is unclear. The British rule results in a higher percentage of weak claims not being filed, or being filed and dropped. The American rule tends to encourage relatively strong claims into court, but these claims are drawn out of a pool that is generally weaker than would be observed under the British rule.

The likely result is that medical malpractice claims litigated under the British rule and under the American rule are of roughly the same merit. If this is so, the major impact of the British rule is that it discourages low probability claims. In the area of medical malpractice, however, these claims will tend to be those based on new technology. The major implication, then, is that medical malpractice litigation under the British rule will put a damper on the process, described by Mark Grady,\textsuperscript{67} in which new technology gives rise to ever more stringent negligence tests.

\textsuperscript{64} Id.
\textsuperscript{66} Edward A. Snyder & James W. Hughes, 6 J.L. ECON. & ORG. 345 (1990).
\textsuperscript{67} Grady, supra note 65.
IV. **Toward a Normative Theory of Procedural Rules**

**A. Pro-plaintiff Rule and Predictability**

I have tried to demonstrate that the British and American cost-allocation rules are not very different in their implications for the predictability of law. In this section, I will argue that the ideal rule on predictability grounds is the Pro-plaintiff. The Pro-plaintiff rule incorporates the desirable features of the British and American rules while avoiding the undesirable.

The important characteristics of the Pro-plaintiff rule are as follows. First, like the British rule, the incentive to litigate under the Pro-plaintiff rule increases substantially as the merit of the claim increases. Second, like the American rule, the Pro-plaintiff rule does not introduce a special disincentive to low probability claimants. Third, and most important, the Pro-plaintiff rule goes further than any other in removing costs as an obstacle to plaintiffs with valid claims. Fourth, the Pro-plaintiff rule does not encourage litigation over the settlement of filed claims as strongly as does the British rule.

There is a fifth characteristic that deserves separate treatment: the compliance effect. I noted in the previous section that compliance incentives generated by the British and American rules were similar. The reason is that although the British rule increases the liability of injurers, it reduces the incentive of victims to bring suit. The Pro-plaintiff rule, however, increases the expected liability of injurers without reducing the incentive to file a valid claim and, therefore, has a substantial impact on compliance.

The connection between the compliance effect and predictability is not obvious in the analysis of the preceding section, but there is a straightforward intuitive explanation. A regime of high compliance enhances predictability as follows. In a high compliance regime, actors comply with the laws and also the norms, conventions, and expec-

---

68. To see this, consider two claims, alike except for the probability of victory by the plaintiff. Suppose the cost of litigating is $1, and the expected judgment is $2. Suppose the plaintiff's probability of victory is .8 in one case and .9 in the other. For the lower merit plaintiff, the expected gain from bringing suit is

\[(.8)(2) - (.2)(1) = 1.6 - .2 = 1.40.\]

For the high merit plaintiff the expected gain from suit is

\[(.9)(2) - (.1)(1) = 1.8 - .1 = 1.70.\]

Under the American rule, the expected gain for the lower merit plaintiff is $.60, and for the higher merit plaintiff $.80. Thus, the Pro-plaintiff rule gives greater encouragement to high merit claims than does the American rule.

69. It should be noted that the Pro-plaintiff rule does imply a greater incentive to litigate a filed claim than does the American rule. My point is that the incentive to litigate rather than to settle is weaker than under the British rule.
tations on which the laws are based. The greater the level of compliance, the firmer and more stable are the underlying expectations.

This might seem to be an argument for massive amounts of litigation, but that need not follow. A regime of high compliance may minimize litigation. First, by reducing the number of injuries that give rise to disputes, encouraging compliance reduces litigation. Second, the greater the level of compliance, the greater the expectation on the part of potential litigants that, in unclear cases, the defendant actually complied with the law. This leads plaintiffs to settle their disputes rather than hold out for large damage judgments.

As I noted earlier, the existence of informational asymmetry among litigants can further complicate efforts to link the merit of the typical claim to the type of fee-shifting rule. Recall that when the defendant is informed and the plaintiff is uninformed, the British rule, when compared to the American rule, may have the effect of encouraging weaker claims into court. The same charge might be leveled against the Pro-plaintiff rule. The important differences, however, are that the Pro-plaintiff rule does not encourage as much litigation as does the British rule, and, while the American and British rules are fairly similar in their influences on compliance, the Pro-plaintiff rule has a substantially bigger impact on compliance than either the American or British rules. These differences suggest that the Pro-plaintiff rule is unlikely to present the risk to predictability that the British rule presents in this area.

The first, second, and fourth characteristics of the Pro-plaintiff rule are important for static predictability purposes. Because the incentive to litigate increases with the merit of the claim, the Pro-plaintiff rule, like the British rule, supplies the greatest encouragement to claims that are firmly grounded in the law. One might argue that it is also desirable that the rule discourage low probability claims, but this is not clearly desirable. Many low probability claims may be well

70. In other words, I must distinguish my argument from the "public goods" theory of litigation advanced perhaps most forcefully in Owen M. Fiss, Against Settlement, 93 YALE L.J. 1073 (1984). The public goods theory holds that litigation is desirable because it results in clear statements of the law, which is a general benefit. I am not calling for more litigation. The public goods theory of litigation is not central to the theory of this Article, and I have rejected it in earlier work. See Hylton, Fee Shifting, supra note 4, at 1075-76.

71. Thus, it does not follow that the Pro-plaintiff rule will lead to a large surge in filings, stressing the capacity of courts to make informed decisions. If compliance increases, and plaintiffs assume that a higher percentage of defendants actually complied with the law, the number of complaints might fall, and the frequency of settlements might increase.

grounded in law, while lacking in supporting evidence. The object of common law in this instance is to develop rules specifying the standard of proof and adequacy of evidence. The decisions may be interpreted by some as dealing only with the standard of proof, but they can equally well be viewed as determining substantive law.\textsuperscript{73} In light of this function, procedural rules should be designed to encourage claims of this sort. It enhances static predictability to know what the standard of proof requires with respect to these claims.

Still, one might argue that the British rule is superior because at least some of the low probability claims that are discouraged by it are nonmeritorious. These claims, if successful, can only lead to confusion over the meaning of a rule.

Consider, for example, the interpretation of a statute, such as the Sherman Act. The British rule, by encouraging claims well grounded in law and discouraging others, would tend to produce a stable and consistent interpretation. The Pro-plaintiff rule, which is actually applied, leads to a rapidly evolving law at the cost of stability and consistency. One benefit of the Pro-plaintiff rule, however, is that it encourages low probability claims that establish evidentiary standards under the Sherman Act. One large class of claims falling in this category are the conscious parallelism charges involved in the famous \textit{Interstate Circuit v. United States}\textsuperscript{74} and \textit{Theater Enterprises v. Paramount Film Distrib. Corp.}\textsuperscript{75} decisions. The conscious parallelism cases involve conspiracy claims based on circumstantial evidence. Because they are necessarily low probability claims, it is unlikely that they would be brought under a British fee-shifting rule. It is beyond dispute that consideration of these claims provides an important set of doctrines connected to the conspiracy section of the Sherman Act.

Still, one might argue that a decision such as \textit{Interstate Circuit}, which lays out several types of circumstantial evidence that a court should consider in examining the merit of a conscious parallelism conspiracy claim,\textsuperscript{76} harms the predictability of law by setting up complicated, fact-dependent tests. Specifically, the Court in \textit{Interstate Circuit} discussed the extent to which the parties charged with conspiracy shared information, the complexity of their parallel actions, and the extent to which the evidence suggested that at least some of the par-

\textsuperscript{73} \textit{Holmes, supra} note 2, at 120-21.
\textsuperscript{74} 306 U.S. 208 (1939).
\textsuperscript{75} 346 U.S. 537 (1954).
\textsuperscript{76} \textit{Interstate Circuit}, 306 U.S. at 225-30.
ties were motivated by purely independent interests. Surely, one could contend that it cannot enhance predictability for a court to be charged with the responsibility of sifting through the facts to resolve these questions.

If one views predictability in the static sense alone, the answer to this question is probably negative: complicated, fact-dependent legal tests tend to reduce predictability. Correspondingly, bright-line rules enhance predictability. The problem is that bright-line rules are inflexible in the face of changing conditions. The factors announced in *Interstate Circuit* are lines of inquiry that follow naturally from the theory of the statute. To apply Section 1 of the Sherman Act in a manner consistent with the underlying theory would seem to require consideration of the *Interstate Circuit* factors. An approach that adheres to the theory of the statute is likely to reach sensible answers when applied to novel fact settings, which enhances dynamic predictability. In other words, bright-line or “per se” rules often enhance static predictability while weakening dynamic predictability.

Further, complexity alone cannot be relied on as a reason for rejecting law on predictability grounds. A large number of conventions involve complex rules that would be difficult to spell out in full. Institutions, such as families and businesses, develop bodies of common law that are as complicated and fact dependent as that generated by courts. And yet, the common law within these institutions, such as the “law of the shop” in the employment context, is often quite predictable to the parties involved. Contracts are also complicated in many settings, and require difficult fact-dependent assessments. But this does not imply that the parties cannot understand their agreement.

The discussion of *Interstate Circuit* suggests the existence of an overlap between sources of static and dynamic uncertainty. Decisions establishing the standard of proof under a statute address problems of static and dynamic uncertainty at the same time. They reduce static uncertainty by contributing to the understanding of the statute’s scope of application. They reduce dynamic uncertainty because they estab-

77. *Id.* at 226-27.

78. There are counterexamples. A bright-line rule that driving on the left side of the street is negligence per se will hardly raise the kind of dynamic predictability problems as the more expansive per se doctrines of antitrust. Thus, bright-line rules that are limited to a specific action are easier to defend on predictability grounds. I do not see the existence of per se rules in tort law, for example, as inconsistent with the argument in the text.

lish rules covering new conditions, unforeseen at the time the statute was drafted.

Thus, it is unlikely that the British rule is superior to the Pro-plaintiff rule on static uncertainty grounds. The British rule is arguably superior as a procedural rule designed to throw light on the established interpretation of a given statute. Another class of static uncertainty problems, referred to earlier as rule identification, is, however, probably best solved through the use of the Pro-plaintiff rule.

Though it is not clear whether the Pro-plaintiff rule is superior to the British on static grounds, this should be true as far as dynamic predictability is concerned. The reason is that the Pro-plaintiff rule goes further than any other rule in removing costs as an obstacle to valid claims. If law is going to evolve in a manner that responds to changing conditions, then it is important that there be as few obstacles as possible in the way of valid claims.

One might argue that the Pro-plaintiff rule also encourages nuisance or nonmeritorious claims. The weaknesses in this argument are two. First, it is the responsibility of courts to distinguish meritorious and nonmeritorious claims and to sanction bad faith litigants. The best way to discourage nonmeritorious suits is to enhance the power of courts to make these distinctions and penalize bad-faith litigants, not to make it difficult to bring any claim regardless of merit. Second, because the Pro-plaintiff rule makes a greater contribution to predictability than the American rule, it should be easier over time for courts to distinguish meritorious from nonmeritorious claims. In other words, the goal of enhancing predictability deserves a higher priority than the discouragement of meritless claims.

B. The Point of Procedure

The preceding discussion assumes that rapid evolution is a desirable characteristic of law, a point that remains to be justified. It is beyond the scope of this Article to provide a detailed justification. A few comments are all I can provide here.

The assumption underlying this project is that law establishes a set of rules that makes an implicit contract among actors. It should be clear that it is easier for actors to comply with the rules when they are the same or at least close to the rules they would have adopted themselves in a setting of low transaction costs. Take, for example, the negligence rule. It is a simple solution to the problem of controlling accident costs. It is unlikely that parties in a zero transaction cost set-
ting would choose strict liability because, in addition to enormous administrative problems, it would fail to adequately regulate incentives to take care. Thus, the negligence rule suggests itself as a solution parties probably would choose in a low transaction cost setting.

Since law must operate largely without the help of officials or judges, it is desirable that it correspond closely with terms the contract parties would choose themselves. Law that diverges significantly from this norm takes the form of a set of rules that would seem inappropriate to the parties. Put another way, law that is dislodged and unconnected to the expectations and norms generated in ordinary intercourse would be a constant source of surprise, like Caligula's edicts.

It follows that dynamic predictability is the most important criterion. Procedural rules that foster dynamic predictability have the effect of tying law closely to expectations. These expectations change with underlying conditions and with the introduction of new technology. Procedural rules that enhance dynamic predictability enable the law to respond to shifts in expectations.

Thus, the underlying normative theory of procedure may be stated as follows: procedural rules make up the cords connecting law to the set of expectations and norms generated in social interaction and the cost-allocation rules form an important part of these cords. Rules such as the American and British rules put a substantial barrier in front of claims that are well grounded in expectations, but not in established law. The result is that the law drifts further from the underlying set of expectations—in the direction of Caligula's edicts. The Pro-plaintiff rule minimizes the barrier between claims that are well grounded in expectations, but not so in law. By enabling the largest percentage of such claims, the rule maximizes the likelihood that law will modify itself in light of underlying expectations.

As I have suggested, this goal may require the sacrifice of some degree of static predictability. The desirableness of static predictability, however, has been exaggerated. A statute whose scope of application is entirely clear is of little use if it is likely to be rendered obsolete.

80. HOLMES, supra note 2, at 90-95 (discussing causation and strict liability).
81. See John P. Brown, Toward an Economic Theory of Liability, 2 J. LEGAL STUD. 323, 338-41, 347 (1973); see also Steven Shavell, Strict Liability Versus Negligence, 9 J. LEGAL STUD. 1, 7 (1980).
82. BLACKSTONE, supra note 6, at 46 ("[I]t is incumbent on the promulgators to [state the law] in the most public and perspicuous manner; not like Caligula, who . . . wrote his laws in a very small character, and hung them up upon high pillars, the more effectually to ensnare the people.")
by a change in conditions. Equivalently, clarity is not useful if it also limits a rule's scope of application to a set of measure zero. The most important predictability characteristic is one that would allow ordinary people to identify the rules that apply to their transactions.

The evolution of antitrust law may be taken as a case in favor of the Pro-plaintiff rule. I contend that the most important characteristic of a statute such as the Sherman Act is the ease with which firms can determine whether they are in compliance. This is enhanced by bringing the law in line with norms of conduct that operate in the market. The federal common law connected with the Sherman Act has essentially moved in this direction over time. The Supreme Court's first effort to interpret Section 1 of the Sherman Act, *United States v. Trans-Missouri Freight Ass'n*,\(^{83}\) resulted in the literal reading that the statute outlawed all combinations in restraint of trade.\(^{84}\) This reading was in large part superseded by the Court's adoption of the Rule of Reason in 1911.\(^{85}\) Under the Rule of Reason, only those combinations that *unreasonably* restrained trade would be found in violation of the Act. It is probably fair to say that the evolution of antitrust law reveals a steady movement toward the adoption of rule of reason tests in all areas of litigation.

The increasing reliance on reasonableness tests in antitrust law is a movement away from rigid, formal legal tests, based largely on the words of the statute, toward flexible rules that incorporate the expectations of actors. The formal tests are embodied in the per se rules, which hold certain practices unlawful irrespective of reasonableness justifications. Although Congress has attempted several times to compel courts to apply per se rules to antitrust violations, the most famous attempt culminating in the Clayton Act of 1914,\(^{86}\) the trend over time in court decisions has been to move away from such rules. For example, the law governing tying arrangements, specifically prohibited by

---

83. 166 U.S. 290 (1897).
84. *Id.* at 328.
Section 3 of the Clayton Act, has moved generally in the direction of a rule of reason standard.\textsuperscript{87}

While it seems to be the general consensus that this trend is the result of changes in the preferences of judges, this theory is unpersuasive. It is hard to believe that the set of judges interpreting the Sherman Act could have become an increasingly anti-interventionist body over the last one hundred years. The trial selection process is a more plausible explanation, and a crucial part of that selection process is the Pro-plaintiff rule. By placing a minimal burden on plaintiffs whose claims, though weak in terms of established law, are justifiable on economic reasonableness grounds, the Pro-plaintiff rule enables courts to develop a case law that gives careful consideration to the reasonableness arguments of the parties. This case law, in turn, permits courts to make distinctions between claims that are grounded in the form, but not in the substance of antitrust law.

\textbf{V. Conclusion}

Most analyses of legal uncertainty are concerned with certainty in the static sense, i.e., in determining the scope of a rule’s application at a given moment. Another source of uncertainty, dynamic, however, arises from the possibility that the law may change over time or fail to be consistent with expectations. Procedural systems and rules within procedural systems may be ranked in terms of their influence on these types of uncertainty.

This Article ranks the American ("no-way" fee shifting), British ("two-way"), and Pro-plaintiff ("one-way") litigation cost-allocation rules in terms of their influence on uncertainty or, equivalently, predictability of law. I conclude that the British and American rules are relatively similar in their influences on static uncertainty, with perhaps a slight advantage for the British rule. The American rule, however, is superior as a device for minimizing dynamic uncertainty. The best rule overall on predictability grounds is the Pro-plaintiff rule.
