

SUIT, SETTLEMENT, AND TRIAL: A THEORETICAL ANALYSIS UNDER ALTERNATIVE METHODS FOR THE ALLOCATION OF LEGAL COSTS

STEVEN SHAVELL*
Harvard Law School

WILL a party who believes that he has a legally admissible claim for money damages decide to bring suit? If so, will he subsequently settle with the opposing party or will he go ahead to trial? These questions are considered in this article under four methods for allocating legal costs,¹ namely, under the *American system*, whereby each side bears its own costs; under the "indemnity" or *British system*, whereby the losing side bears all costs; under the *system favoring the plaintiff*, whereby the plaintiff pays only his own costs if he loses and nothing otherwise; and under the *system favoring the defendant*, whereby the defendant pays only his own costs if he loses and nothing otherwise.²

* Assistant Professor, Harvard Law School. I wish to thank Mark Freeman and David Skalak for able research assistance; Abram Chayes, Bruce Greenwald, Richard Markus, Thomas Rowe, Frank Sander, and David Shapiro for helpful comments; and the National Science Foundation (grant no. SES-8014208) for financial support.

¹ The analysis to be presented of suit, settlement, and trial and of the methods for allocating costs builds on earlier work in William Landes, *An Economic Analysis of the Courts*, 14 J. Law & Econ. 61 (1971), a theoretical and empirical study focusing on settlement versus trial in criminal cases; in John Gould, *The Economics of Legal Conflicts*, 2 J. Legal Stud. 279 (1973), a formal theoretical investigation of settlement versus trial under the American system in civil cases; and in Richard Posner, *Economic Analysis of Law* 434-440, 445-453 (2d ed. 1977), a theoretically oriented and quite general discussion of issues pertinent to settlement versus trial under the American and under the British system. The analysis is also informed by Philip J. Mause, *Winner Takes All: A Re-Examination of the Indemnity System*, 55 Ia. L. Rev. 26 (1969), a key reference on litigation under the British and under the American system.

The contribution made here lies principally in the distinction drawn between the question concerning suit and the subsequent question concerning settlement versus trial. However, the generality of the analysis—the consideration of the four methods of allocating costs, the allowance for parties' differing beliefs not only as to the likelihood of prevailing but also as to the judgment (which is taken to be continuously variable)—may also be of interest.

² The American system is the prevailing method for allocating legal costs in the United

[*Journal of Legal Studies*, vol. XI (January 1982)]

© 1982 by The University of Chicago. 0047-2530/82/1101-0003\$01.50

To examine in a systematic and careful way the two questions under the methods for allocating legal costs, it will be necessary to study a simple, stylized model of the litigation process. Analysis of this model in the text is informal, with a series of numerical examples illustrating the main points of interest, and so should be accessible to the widest audience; a formal statement of the model and proofs of results are contained in an Appendix. Following the analysis, two brief illustrations are considered and comments are made on the relative social desirability of the methods for allocating legal costs.

I. THE MODEL³

The parties are assumed to view suit and settlement or litigation solely as a financial matter and to face the situation shown in figure 1. As indicated, the plaintiff⁴ must first decide whether to bring suit.⁵ He is assumed to be able to bring suit at no cost⁶ and to decide to do so if and only if he would be willing to go to trial.⁷ Of course, whether he would be willing to go to trial depends on an ex ante evaluation of his chance of prevailing, on the probable magnitude of a judgment, and on the legal costs of going to trial and the method by which they are to be allocated.

If the plaintiff does decide to bring suit, it is assumed that he and the defendant will reach a settlement if and only if there exists some settle-

States, although exceptions to its use are made in certain jurisdictions for certain categories of case, whereas the British system is the usual rule determining the bearing of legal costs in the United Kingdom. For details, see Mause, *supra* note 1, and references cited therein. The systems favoring the plaintiff or the defendant are thus departures from the norm, but they are sometimes employed. (Examples of systems favoring plaintiffs exist under federal law, for instance, under Title VII of the Civil Rights Act as interpreted in *Christiansburg Garment Co. v. EEOC*, 434 U.S. 412, 415-21 [1978]. An example of the other type of system may also be noted—the state of Florida has recently adopted the system favoring the defendant in medical malpractice cases. Ch. 80-67, 1980 Fla. Laws 224.)

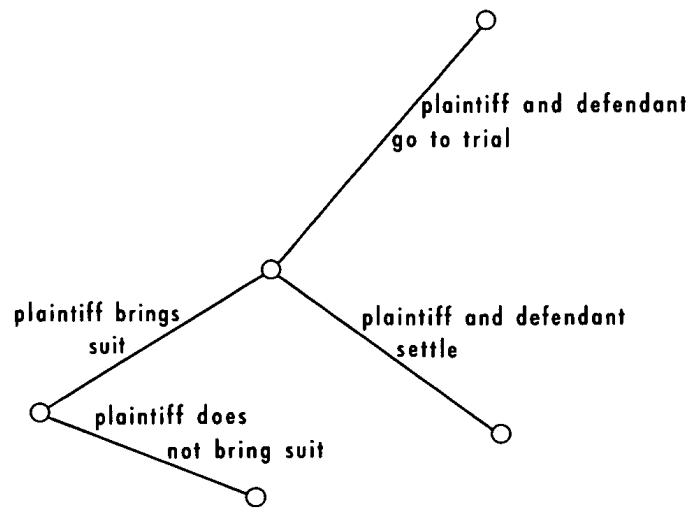
³ The usual cautionary remark should be made about this model, that the aim is not for descriptive accuracy, but rather for sufficient simplicity and analytic tractability as to provide a generally useful tool for thought and guide for the design of empirical research.

⁴ For simplicity, the party contemplating suit will be referred to as the "plaintiff" even though he may decide against bringing suit.

⁵ From fig. 1, it is clear that "bringing suit" is interpreted as any action which results ultimately in settlement or in trial. Thus, we would consider a mere threat to initiate formal proceedings as a suit if the threat resulted in a settlement.

⁶ This assumption is inessential; all that will matter to our results is that going to trial involves legal costs in addition to any connected with bringing suit.

⁷ Thus, it is implicitly assumed that when the plaintiff would not be willing to go ahead to trial, he would not bring suit merely in the hope of extracting a settlement from the defendant. Note that this assumption would be plausible if, in the latter case, the plaintiff believed that the defendant knew the plaintiff's position, for then the defendant might not in fact offer anything in settlement.



SEQUENCE OF ACTIONS

FIG. 1

ment amount that both he and the defendant would prefer to going to trial.⁸ (The settlement process is itself assumed to be costless.)⁹ Thus, whether settlement will occur depends on both the plaintiff's and the defendant's *ex ante* evaluations.

These evaluations will be analyzed under two different assumptions. According to the first, a party is assumed to be *risk neutral*. Under risk neutrality a party evaluates an uncertain prospect by its *expected value*, that is, by discounting possible outcomes by their probabilities.¹⁰ If the plaintiff thinks the likelihood of obtaining a judgment for \$10,000 is 25 percent, then we would say that the expected value to him of the judgment is \$2,500 (i.e., 25 percent $\times$ \$10,000), or, more simply, that his expected judgment is \$2,500.¹¹ By definition, a risk-neutral party would be indiffer-

⁸ The existence of a mutually attractive settlement amount is of course a necessary condition for there to be a settlement. Thus, the force of the assumption is that the existence of such an amount is also a sufficient condition for settlement, that problems of bargaining would not defeat settlement when it is a possibility.

⁹ The sole importance of this assumption is that it means that settlement is less costly than litigation.

¹⁰ For an introductory treatment of these and other aspects of decision theory, see Howard Raiffa, *Decision Analysis* (1968).

¹¹ The \$2,500 figure may be thought of as the average amount the plaintiff would gain were he to go to trial repeatedly under very similar circumstances. The reader should be careful to keep this interpretation and terminology in mind and not, say, to think of the \$10,000 figure as the expected judgment.

ent between receiving (or paying) \$2,500 for sure and receiving (or paying) \$10,000 with probability 25 percent; he cares only about expected value, not about the degree of uncertainty. In contrast a *risk-averse* party is one for whom uncertainty itself is undesirable;¹² that is, he cares not only about expected value but also about uncertainty. Thus, a risk-averse party would prefer receiving (or paying) \$2,500 for sure to receiving (or paying) \$10,000 with probability 25 percent, and this to receiving (or paying) \$100,000 with probability 2.5 percent.¹³ Most of the analysis that follows will focus on the case where both parties are risk neutral. This case is emphasized only because it allows an especially simple characterization to be made of certain underlying factors determining suit and settlement or trial. The case where one or both of the parties is risk averse is not viewed as less important and would in fact be likely to be of significance when a party has little wealth or, more generally, when the amount at stake is large in relation to his assets.¹⁴

II. SUIT

Given the assumptions of the model and, for now, that of risk neutrality, it is straightforward to identify and to compare the circumstances under which a plaintiff will bring suit under the four systems of allocating legal costs. The first statement to be made is simply that *under the American system, the plaintiff will bring suit if and only if his expected judgment would be at least as large as his legal costs*.¹⁵ To illustrate, consider

Example 1: The plaintiff's legal costs from a trial would be \$1,000, and he believes that if he prevails he would obtain a judgment for \$10,000. Thus, if he thinks the likelihood of prevailing is, say, 75 percent, he will bring suit, for then his expected judgment from a trial would be \$7,500 (i.e., 75 percent $\times$ \$10,000), which

¹² Another possibility is that a party is risk loving and finds uncertainty inherently desirable. Although this possibility is not analyzed, it will be clear how it would alter our conclusions.

¹³ Exactly how a party weighs expected value against uncertainty depends on his *degree of risk aversion*. See Raiffa, *supra* note 10; and John W. Pratt, Risk Aversion in the Small and in the Large, 32 *Econometrica* 122 (1964).

¹⁴ It is plausible and often assumed that individuals display a risk-averse attitude in respect to the possibility of large fluctuations in their wealth, and especially to the possibility of losses that would leave them in a financially precarious position. (For example, the hypothesis of risk aversion is used to explain the purchase of insurance.)

¹⁵ Let p denote the plaintiff's estimate of his chance of prevailing, w his estimate of his judgment, and x his legal costs were he to go to trial. (This notation and the assumption that there is a single estimated value, rather than a range of values, will be used without further comment in the present footnote and in other footnotes. The Appendix considers a more general case.) Then under the American system he would bring suit if $pw > x$. See Proposition 1a and eq. (1) of the Appendix (which, as noted before, allows for the plaintiff to contemplate a range of possible magnitudes of the judgment should he prevail).

would exceed his legal costs of \$1,000. However, if he believes the chance of success is only, say, 5 percent, he will not bring suit, since his expected judgment would be \$500, which would be less than his legal costs.

Similarly, under the British system, the plaintiff will bring suit if and only if his expected judgment would be at least as large as his expected legal costs—that is, the total legal costs discounted by his probability of losing at trial.¹⁶ Consider now

Example 2: The plaintiff's legal costs and beliefs concerning a judgment are as in the previous example, and the defendant's legal costs from a trial would be \$1,500.¹⁷ Consequently, if the plaintiff thinks the probability of prevailing is 75 percent, his expected judgment of \$7,500 would exceed his expected legal costs of \$625 (i.e., 25 percent $\times$ \$2,500), so that he would bring suit. But if he thinks the probability is 5 percent, his expected judgment of \$500 would be below his expected legal costs of \$2,375 (i.e., 95 percent $\times$ \$2,500), so that he would not bring suit.

Comparing the two systems, it is apparent that *the frequency of suit will be greater under the British system when the plaintiff believes the likelihood of prevailing is sufficiently high—above a “critical” level—and the frequency will be greater under the American system when the likelihood is below the critical level.*¹⁸ This is so because when the plaintiff is relatively optimistic about prevailing, his expected legal costs will be relatively low under the British system—he will be thinking about the possibility of not having to pay any such costs—whereas under the American system he must bear his own costs with certainty. Thus he will be likely to find suit a more attractive prospect under the British system. But when the plaintiff is not optimistic, converse reasoning explains why he would

¹⁶ Let y denote the defendant's legal costs were he to go to trial. Then the plaintiff would bring suit if $pw > (1 - p)(x + y)$. See more generally proposition 1a and eq. (2) of the Appendix.

¹⁷ The plaintiff's and the defendant's legal costs will also be assumed to be \$1,000 and \$1,500, respectively, in later examples.

¹⁸ Specifically, the frequency of suit will be higher under the British system if $p > y/(x + y)$ and higher under the American system if $p < y/(x + y)$. If $p = y/(x + y)$, the plaintiff will sue under the British system if and only if he would under the American system. Thus, “sufficiently high” means exceeding $y/(x + y)$, etc. See more generally proposition 1a and eq. (5) of the Appendix. It should also be observed that the plaintiff's assessment of his chance of prevailing, and thus the comparison of frequencies of suit, may depend on the relevant legal rule or area of law. For instance, one would expect that a party who suffers a loss would generally be more optimistic about his chance of prevailing under a strict liability approach than he would under a fault-based approach. Thus, one might expect the ratio of suits under the British system to suits under the American system to be higher under a strict liability approach than under a negligence approach. See the discussion in the concluding section.

be expected to sue more often under the American system.¹⁹ The following example explicitly demonstrates this result.

Example 3: If the plaintiff believes that his chance of prevailing is at the "high" level of 75 percent, then it was observed that under the British system his expected legal costs would be \$625. Since this is less than the \$1,000 in costs that he would bear under the American system, the plaintiff would be more likely to bring suit under the British system; specifically, if his expected judgment would exceed \$625 but would be less than \$1,000, he would sue under the British system but not under the American.²⁰ However, if instead the plaintiff thinks his chance of prevailing is at the "low" level of 5 percent, his expected legal costs from going to trial under the British system would be, as was noted, \$2,375, so that he would be more likely to bring suit under the American system; whenever his expected judgment would exceed \$1,000 but not \$2,375, he would sue only under the American system. The *critical probability* of prevailing—above which the plaintiff would find suit more attractive under the British system and below which under the American system—turns out in this example to be 60 percent, for at 60 percent, the expected legal costs under the British system would be \$1,000 (i.e., 40 percent $\times$ \$2,500), which equals what they would be under the American system.²¹

Turning now to the other two systems, it is clear that *under the system favoring the plaintiff, he will bring suit if and only if his expected judgment at trial would be at least as large as his expected legal costs—his own legal costs discounted by his probability of losing. And under the system favoring the defendant, the plaintiff will bring suit if and only if his expected judgment would be at least as large as his expected legal costs—now equal to his own legal costs, plus those of the defendant discounted by the probability of losing.*²² Moreover, with regard to a comparison of

¹⁹ Because so-called nuisance suits typically have a low probability of succeeding, this conclusion is in line with the notion that such suits are more likely brought under the American than the British system. See, for example, Mause, *supra* note 1.

²⁰ Suppose, for instance, that the plaintiff believes that if he prevails he will receive a judgment for \$1,200. Then his expected judgment would be \$900.

²¹ The critical probability equals $y/(x + y)$; see note 18 *supra* and proposition 1a and eq. (5) of the Appendix. In this example the plaintiff's cost x was \$1,000 and the defendant's y was \$1,500, so $y/(x + y) = \$1,500/\$2,500 = 60$ percent. If the defendant's costs were not \$1,500 but \$1,000, and thus equal to the plaintiff's, then the critical probability would be 50 percent, since $y/(x + y) = \$1,000/\$2,000$; and were the defendant's costs \$500, the critical probability would be 33 percent, since $y/(x + y) = \$500/\$1,500$. That the critical probability falls when the defendant's legal costs fall simply reflects the fact that the smaller the defendant's costs, the more attractive the British system becomes to the plaintiff.

²² We will refrain from illustrating this in the text, but it may be helpful to provide one calculation here: If the plaintiff believes his chance of prevailing is 75 percent, then under the system favoring himself his expected legal costs would be \$250 (i.e., 25 percent $\times$ \$1,000), while under the system favoring the defendant they would be \$1,375 (i.e., \$1,000 + 25 percent $\times$ \$1,500), so the plaintiff would sue under the former system if his expected judgment exceeds \$250 and under the latter if it exceeds \$1,375. More generally, the plaintiff will bring suit under the system favoring himself when $pw > (1 - p)x$, and under the system favoring the defendant when $pw > x + (1 - p)y$. See eqq. (3) and (4) and proposition 1a.

frequencies of suit under the four systems, it is readily shown that *the frequency of suit will be greatest under the system favoring the plaintiff and least under the system favoring the defendant*.²³ Why this should be the case is obvious. On the one hand, the plaintiff finds the system favoring himself most attractive—he never bears legal costs greater than under the other systems and may bear a lesser cost. On the other hand, the plaintiff finds the system favoring the defendant least attractive—the plaintiff then never bears legal costs lower than under the other systems, and may bear a greater cost.

How is what has been said to this point affected if the plaintiff is assumed to be risk averse rather than risk neutral? Plainly, *the general effect of the plaintiff's risk aversion is to reduce the likelihood of suit*, for going to trial involves uncertainty, which by definition the risk-averse plaintiff but not the risk-neutral plaintiff finds disadvantageous.²⁴ To see this effect illustrated under the American system, consider

Example 4: Given that his legal costs would be \$1,000 and the prospect of receiving a judgment for \$10,000, a risk-neutral plaintiff would bring suit under the American system if his probability of prevailing exceeds 10 percent. To see that a risk-averse plaintiff would not bring suit as often, assume, for instance, that the plaintiff's initial wealth is \$5,000 and that his "utility of wealth" is the square root of wealth.²⁵ Then it is indeed so that the plaintiff might not sue when his risk-neutral counterpart would. If the risk-averse plaintiff believes the probability of prevailing is, say, 11 percent, he would choose not to sue: if he does not sue, his utility of wealth is 70.71 (i.e., $\sqrt{5,000}$); if he goes to trial his "expected utility" is less, it is 69.30 (i.e., $11 \text{ percent} \times \sqrt{14,000} + 89 \text{ percent} \times \sqrt{4,000}$ —the utility of his wealth if he wins discounted by the probability of winning plus the utility of his

²³ See proposition 1a and fig. 2.

²⁴ See proposition 1b. The reader should recall that risk aversion becomes an important factor whenever a party is poor or when the amount at stake or legal fees are large in relation to his assets. However, if such a party and his counsel employ a contingent fee arrangement, counsel's willingness and ability to bear risk will be relevant, since the party will not himself bear the risk of paying legal fees if he does not prevail; thus, the effect of the party's risk aversion will be diminished.

²⁵ To perform calculations illustrating the effect of risk aversion, it is necessary to make explicit assumptions about the plaintiff's utility of wealth and then to determine his expected utility (to be explained shortly in the text) in the various situations under consideration. (Some readers may not find such calculations instructive and may omit this and subsequent examples of the effect of risk aversion without loss of continuity.) The importance of the particular assumption that the utility of wealth is its square root is only that it implies that the increment in utility from an extra dollar declines as the absolute level of wealth increases, that is, so-called declining marginal utility of wealth. This in turn is the logical equivalent to an assumption of risk aversion. The intuition behind the relationship between declining marginal utility of wealth and the existence of risk aversion is as follows. If the marginal utility of wealth declines as the absolute level of wealth increases, then the loss of an amount reduces utility more than the gain of an equal amount would increase utility. This notion that losses "hurt" more than equivalent gains "help" comports with the notion that an individual would prefer his present position to taking a risk (unless of course the prospect of gains was sufficiently attractive to outweigh the possibility of loss).

wealth if he loses discounted by that probability).²⁶ In fact, for the particular risk-averse plaintiff described here to bring suit, his likelihood of success must exceed approximately 13½ percent.²⁷

Were examples of the effect of risk aversion under the other systems as well to be considered (which would be tedious), it would become evident that the reduction in the likelihood of suit would not be uniform because the variability in the plaintiff's position as between prevailing or not differs among the systems: the variability is least under the American system, where it equals the magnitude of the judgment;²⁸ it is larger under the system favoring the plaintiff, where it equals the judgment plus the plaintiff's legal costs, and also under the system favoring the defendant, where it equals the judgment plus the defendant's costs;²⁹ and the variability is greatest under the British system, where it equals the judgment plus the plaintiff's and the defendant's legal costs.³⁰ *Thus, we should expect the effect of risk aversion—the reduction in the frequency of suit—to be strongest under the British system, less under the systems favoring the plaintiff and the defendant, and least under the American system.*³¹ However, the qualitative nature of the comparison of frequency of suit under the systems is essentially unaltered.³²

²⁶ An individual's expected utility may be interpreted as the utility he would enjoy on average were he to engage in the risky activity repeatedly.

²⁷ Were the plaintiff more risk averse, that is, if the proportional rate of decline of his marginal utility of wealth were greater, then the likelihood of success would have to be higher than 13 1/2 percent before he would bring suit.

²⁸ In example 4, the plaintiff's wealth under the American system is $\$5,000 - \$1,000 = \$4,000$ if he loses and $\$5,000 - \$1,000 + \$10,000 = \$14,000$ if he prevails, the difference being the \$10,000 judgment.

²⁹ In example 4 the plaintiff's wealth under the system favoring himself is $\$5,000 - \$1,000 = \$4,000$ if he loses and $\$5,000 + \$10,000 = \$15,000$ if he prevails, so the difference is \$11,000, the judgment plus his own legal costs. If the defendant's legal costs are \$1,500, the plaintiff's wealth under the system favoring the defendant is $\$5,000 - \$1,000 - \$1,500 = \$2,500$ if he loses and $\$5,000 - \$1,000 + \$10,000 = \$14,000$ if he does not, so the difference is \$11,500, the judgment plus the defendant's legal costs.

³⁰ In example 4 the plaintiff's wealth under the British system is $\$5,000 - \$1,000 - \$1,500 = \$2,500$ if he loses and $\$5,000 + \$10,000 = \$15,000$ if he prevails, so the difference is \$12,500, the judgment plus his and the defendant's legal costs.

³¹ It is difficult to find simple conditions under which this is always true; see note 60, *infra*. Nevertheless, it is the most plausible conclusion and should be borne in mind. For instance, if plaintiffs are poor and very risk averse, then the discouragement of "meritorious" suits might be greater under the British system than under the American system. (But as remarked in note 24, *supra*, contingent fee arrangements provide an offsetting factor.)

³² The frequency of suit is still greatest under the system favoring the plaintiff and least under that favoring the defendant. Also, the relationship between the American and British systems still depends on the plaintiff's estimate of the probability of prevailing; see note 60, *infra*.

III. SETTLEMENT VERSUS TRIAL

Assuming that the plaintiff has found it worthwhile to bring suit, the question remains as to whether he and the defendant will settle or go to trial. It is important to remember that this question is *conditional* upon the plaintiff's having already brought suit. To determine the *absolute* rates of settlement or litigation under the different systems for allocating legal costs generally requires answering both the conditional question and the prior question about the frequency of suit.³³ The two questions are treated separately here solely for analytical convenience and for clarity. To answer the conditional question, we will use the supposition of the model (see above) that there will be a settlement if and only if there is a settlement amount that both the plaintiff and the defendant would prefer to going to trial. As in the previous section, the case where both parties are risk neutral is considered in detail, followed by a brief discussion of the case where one or both parties are risk averse.

*Under the American system,*³⁴ *there will be a trial if and only if the plaintiff's estimate of the expected judgment exceeds the defendant's estimate by at least the sum of their legal costs.*³⁵ The simple logic behind this result is that because the plaintiff and the defendant will save the sum of their legal costs by settling, the only factor that could lead to a trial is that the plaintiff's expectations as to the likelihood of success or the judgment that could be obtained are more optimistic than the defendant's.³⁶ This is what would make the plaintiff find unsatisfactory what the defendant would be willing to offer in settlement; and, clearly, the plaintiff's feeling pessimistic relative to the defendant could only reinforce the

³³ When, for instance, it is said below that given suit (and among other assumptions, that the parties are risk neutral), trial would be more likely under the British system than under the American system, it will not be possible to infer that the absolute number of trials would be higher under the British system, for it might be that the volume of suits would be lower under the British system. See the discussion in the concluding section.

³⁴ The analysis presented here of settlement versus litigation under the American system is essentially the same as that first done by Landes, by Gould, and by Posner, *supra* note 1.

³⁵ Let q denote the defendant's estimate of the plaintiff's chance of prevailing, w' , his estimate of the judgment that he would have to pay if the plaintiff prevailed, and let p , w , x , and y be as defined in notes 15 and 16, *supra*. Then there will be a trial if and only if $pw - qw' \geq x + y$. This is shown in proposition 2a (see eq. [11]), which employs more or less the following argument (which also underlies example 5): A settlement will be possible whenever the most the defendant would be willing to pay exceeds the minimum amount that the plaintiff would find acceptable. But the most the defendant would pay is the expected judgment plus his legal costs, that is, $qw' + y$; similarly, the least the plaintiff would accept is $pw - x$. Thus, there will be a settlement when $qw' + y > pw - x$, or equivalently, when $pw - qw' < x + y$.

³⁶ This comports with the notion that what leads to litigation is uncertainty as to the law or as to the facts. Without such uncertainty, the plaintiff's beliefs about his chances or possible judgment could not differ from the defendant's.

incentive to settle on account of potential savings in legal costs. (As will be seen below, this conclusion about settlement vs. litigation is a general one, and will apply under the other cost-allocation systems as well.)

Example 5: In this and the other examples set forth in this section, it will be assumed that the defendant believes that the plaintiff will prevail with probability 50 percent and that both the plaintiff and the defendant believe that the judgment the plaintiff would obtain if he prevailed would be for \$10,000. What will vary in this and later examples is only the assumption concerning the plaintiff's belief about the likelihood of prevailing.

Suppose first that the plaintiff's estimate of the likelihood of prevailing is identical to the defendant's. Then the plaintiff's expected judgment would equal the defendant's estimate of it; thus, there will not be a trial. To see this result explicitly, observe that the plaintiff's expected net gain from going to trial would be \$4,000—his expected judgment of \$5,000 (i.e., 50 percent $\times$ \$10,000) less the \$1,000 in legal costs—and he would therefore be willing to make a settlement for an amount exceeding \$4,000. Similarly, the defendant's expected loss from going to trial would be \$6,500—his expected judgment of \$5,000 plus \$1,500 in legal costs—so he would be willing to make a settlement for an amount below \$6,500. Consequently, both the plaintiff and the defendant would be willing to make a settlement for an amount in the range between \$4,000 and \$6,500. Where in that range the settlement would fall would depend on factors (bargaining power, negotiating skill, etc.) lying outside the scope of present concern.

Now suppose that the plaintiff is somewhat more optimistic about his situation, believing that the likelihood of prevailing is 60 percent. Then his expected judgment would be \$6,000, which exceeds the defendant's expected judgment by only \$1,000, an amount less than the \$2,500 joint legal costs, so that there would still be a settlement according to the result under consideration. To see why this result obtains, observe that the plaintiff's expected net gain from going to trial would now be \$5,000 (i.e., \$6,000 – \$1,000), meaning that the range of potential settlements is reduced to that between \$5,000 and \$6,500, but is not eliminated.

However, the possibilities for settlement would be eliminated if the plaintiff's likelihood of prevailing is sufficiently high. If the likelihood were 80 percent, then the plaintiff's expected judgment of \$8,000 would exceed the defendant's by \$3,000, which is more than the sum of their legal costs, so that there would be a trial. To see why, note that in this instance the plaintiff's expected net gain from going to trial would be \$7,000, which is more than the \$6,500 maximum amount the defendant would be willing to pay.

*Under the British system, there will be a trial if and only if the plaintiff's estimate of the expected judgment exceeds the defendant's estimate by at least the sum of their expected legal costs.*³⁷ This result follows from much the same reasoning that applied above concerning the American system, the only difference being that here the condition determining whether there will be a trial involves the sum of the plaintiff's and defendant's expected legal costs. The reason for the difference is of course that

³⁷ There will be a trial when $pw - qw' \geq (1 - p)(x + y) + q(x + y)$; see more generally proposition 2a and eq. (12).

what motivates settlement is what the parties perceive that they will save in legal costs; and what each party perceives himself as saving under the British system is not his own legal costs but rather his expected legal costs, that is, his likelihood of loss times the costs of both parties. The following example illustrates this.

Example 6: If the plaintiff's estimate of the likelihood of prevailing were 50 percent, equal to the defendant's, then since there would be no difference between the estimates of the expected judgment, there would be a settlement. Because the plaintiff's expected net gain from going to trial would be \$3,750 (i.e., \$5,000 less the expected legal costs of 50 percent $\times$ \$2,500) and the defendant's expected loss from going to trial would be \$6,250 (i.e., \$5,000 + 50 percent $\times$ \$2,500), the amount of the settlement would fall between these two figures.

If the plaintiff thought the likelihood of success were 60 percent, then the difference between his estimate of the expected judgment and the defendant's would be \$1,000, which is less than his expected legal costs of \$1,000 (i.e., 40 percent $\times$ \$2,500) plus the defendant's legal costs of \$1,250. Thus, there would be a settlement; in fact, it would be in the range between the plaintiff's expected net gain of \$5,000 (i.e., \$6,000 - \$1,000) were he to go to trial, and the defendant's expected loss of \$6,250.

But if the plaintiff believed the probability of success were 80 percent, then the \$3,000 difference that would exist between his estimate of the expected judgment and the defendant's would exceed his expected legal costs of \$500 (i.e., 20 percent $\times$ \$2,500) plus the defendant's \$1,250; thus, there would be a trial. In this case, the least that the plaintiff would accept would be his \$7,500 (i.e., \$8,000 - \$500) expected net gain were he to go to trial, which is larger than the highest amount the defendant would be willing to pay, \$6,250, his expected loss were he to go to trial.

To compare the likelihood of trial under British and American systems, *assume that the plaintiff and defendant agree about the magnitude of a possible judgment, so that their opinions can diverge only in respect to the probability of prevailing at trial.*³⁸ Then, conditional on suit having been brought, the likelihood of trial under the British system will be greater than under the American system.³⁹ In essence, the reason for this result is

³⁸ While this assumption reflects the notion that the parties to a suit will probably disagree more in their estimates of the likelihood of prevailing than in their estimates of the size of a judgment, the assumption is clearly inappropriate in many plausible situations. The text comments on the implication of relaxing the assumption, and the Appendix considers the most general type of divergence in beliefs between the plaintiff and the defendant.

³⁹ Something like this point was, to my knowledge, first made by Mause, *supra* note 1, who advanced its plausibility employing a numerical example; Posner, *supra* note 1, first proved it under the assumption made here in the text (that $w = w'$). A more precise statement of the point is that there will be a trial under the British system but not under the American system if $pw - qw$ exceeds $(1 - p)(x + y)$ and is less than $x + y$; and if there is a trial under the American system, there must necessarily be a trial under the British system. See, more generally, proposition 2a (which expresses the comparison in a slightly different way, in terms of the relationship between p and q). It bears repeating that the result under consideration, like all results in this section, concerns the likelihood of trial *given* suit; in

that for litigation to be a possibility in the first place, the plaintiff's estimate of the expected judgment must exceed the defendant's, which implies that the plaintiff's estimate of the probability of prevailing exceeds the defendant's (by our assumption, they agree about the magnitude of a possible judgment). But when this is true, the British system tends to reduce the sum of expected legal costs⁴⁰ and thus tends to make a trial more likely.

Example 7: Initially, this example illustrates that if the plaintiff's estimate of the probability of prevailing exceeds the defendant's, then the sum of expected legal costs would be smaller under the British system than the \$2,500 sum under the American system. As noted before, under the British system, the defendant's expected costs would be \$1,250. Thus, if the plaintiff's estimate of the likelihood of prevailing is, say, 55 percent, his expected legal costs would be \$1,125 (i.e., 45 percent $\times$ \$2,500) and the sum of expected costs would be \$2,375 (i.e., \$1,250 + \$1,125); similarly, if the plaintiff's estimate of the likelihood were 70 percent, then, his expected costs would be \$750 (i.e., 30 percent $\times$ \$2,500) and the sum of expected costs would be just \$2,000. Since the sum of expected costs is lower under the British system, it is possible for there to be litigation under that system but not under the American system (yet impossible for there to be litigation under the American system but not under the British system). This would occur whenever the difference between the plaintiff's and the defendant's estimates of the expected judgment exceeds the sum of expected costs under the British system but not the \$2,500 sum of costs under the American system. For instance, if the plaintiff's estimate of the probability of prevailing were 72 percent, the difference between the estimates of the expected judgment would be \$2,200, which exceeds the expected sum of costs of \$1,950 (i.e., 28 percent $\times$ \$2,500 + \$1,250) but not \$2,500; hence there would be litigation under the British system but room for settlement under the American system.⁴¹

It should be remarked, however, that if the assumption that the plaintiff and the defendant agree about the magnitude of the judgment is relaxed, then the conclusion just illustrated may be reversed; it is possible that there would be litigation under the American system but not under the British system. This would occur when the plaintiff is pessimistic relative

particular, it does *not* mean that the number of suits would be higher under the British system. See note 33, *supra*.

⁴⁰ This general point is clearly illustrated by the extreme case where the plaintiff believes he is virtually certain to win and the defendant thinks the plaintiff is virtually certain to lose. Then, under the British system, each party would believe that he would not have to bear legal costs, so the sum of expected legal costs would be negligible.

⁴¹ Specifically, under the American system, there will be a settlement in the range between the plaintiff's expected net gain of \$6,200 (i.e., \$7,200 - \$1,000) and the defendant's expected loss of \$6,500. Under the British system, by contrast, the plaintiff's expected net gain of \$6,500 (i.e., \$7,200 - 28 percent $\times$ \$2,500) exceeds the defendant's expected loss of \$6,250. In fact, it may be verified that (given our other assumptions) whenever the plaintiff's estimate of prevailing falls between 70 percent and 75 percent, there would be litigation under the British system and settlement under the American system.

to the defendant about the probability of prevailing, yet believes that the judgment would be much higher should he prevail than does the defendant. In this type of situation, since the plaintiff is the party who is relatively pessimistic about the likelihood of prevailing, the British system tends to increase the perceived joint legal costs of the parties⁴² and therefore makes it more likely that they will settle.⁴³

Consider now the systems favoring the plaintiff or the defendant. Under these systems the general way of expressing the condition determining whether there is settlement or trial is what it was before, that is, that there will be a trial if and only if the plaintiff's estimate of the expected judgment exceeds the defendant's estimate by at least the sum of their expected legal costs. But, of course, since the parties' expected legal costs will be different under the two systems, the frequencies of settlement and litigation will also differ.⁴⁴ More precisely, *assuming that the plaintiff and the defendant agree about the size of the potential judgment, it can be shown that conditional on suit having been brought, the likelihood of trial under either the system favoring the plaintiff or the defendant is greater than that under the American system and less than that under the British system.*⁴⁵ There will be a greater likelihood of litigation under the two systems in question than under the American system because (when the plaintiff's estimate of the chances of prevailing exceeds the defendant's) the joint expected legal costs tend to be lower under them than the joint costs under the American system.⁴⁶ And there will be a lesser likelihood

⁴² Think of the situation converse to that described in note 40, *supra*: Suppose the plaintiff thinks he is very unlikely to prevail and the defendant thinks the plaintiff is very likely to prevail. Then, under the British system, each party would believe that he would probably have to pay the entire legal costs, so the expected sum of costs would be high (approaching twice the actual sum of costs).

⁴³ Consider the following example (and see proposition 2a). The defendant's beliefs are as in the examples of this section; he believes that he will lose with probability 50 percent and pay a judgment of \$10,000. But the plaintiff's beliefs are different; he thinks that his chance of prevailing is only 40 percent, yet that if he succeeds at trial, he will receive a judgment for \$19,000. Thus, under the American system, the plaintiff's expected net gain from going to trial would be \$6,600 (i.e., 40 percent $\times$ \$19,000 - \$1,000) and the defendant's expected loss would be \$6,500, so there would be no room for settlement. Under the British system, however, the plaintiff's expected net gain from going to trial would be \$6,100 (i.e., 40 percent $\times$ \$19,000 - 60 percent $\times$ \$2,500) and the defendant's expected loss would be \$6,250, so there would be a settlement.

⁴⁴ There will be litigation under the system favoring the plaintiff when $pw - qw' \geq (1 - p)x + (qx + y)$, and under the system favoring the defendant when $pw - qw' \geq [x + (1 - p)y] + qy$. See generally proposition 2a and eqq. (13) and (14).

⁴⁵ See generally proposition 2a.

⁴⁶ To see why, suppose that the plaintiff thinks he is almost sure to prevail and that the defendant believes that the plaintiff is almost sure to lose. Then under the system favoring the plaintiff, the plaintiff will believe that he will not have to bear any legal costs and the defendant will think that he will have to bear only his own, so that the sum of expected legal

of litigation under the two systems than under the British system for analogous reasons. (For a comparison of the conditional likelihood of trial under the system favoring the plaintiff and under that favoring the defendant, the reader is referred to the Appendix.)⁴⁷

With regard to the effect of the parties' attitudes toward risk on their behavior, it is clear that *risk-aversion tends to promote settlement*; by settling, the risk inherent in going to trial is avoided. The differential effect of risk aversion may be shown most easily under the American system.

Example 8: If the plaintiff and the defendant are risk neutral, then under the American system there will be a trial whenever the plaintiff's estimate of the likelihood of suit is at least 75 percent, for at 75 percent the difference between his and the defendant's estimate of the expected judgment would be \$2,500, which equals the sum of legal costs. However, if one or both of the parties are risk averse, there will be less litigation. Assume, say, that the plaintiff is risk averse, and, as in example 4, that he has initial wealth of \$5,000 and utility of wealth given by its square root. Then there would be a settlement if the plaintiff's probability of prevailing were 76 percent. Consider, for instance, whether the amount \$6,400 would be a mutually satisfactory settlement. The defendant would be willing to pay it in settlement, as it is less than his expected cost of going to trial of \$6,500 (i.e., 50 percent $\times$ \$10,000 + \$1,500). And the risk-averse plaintiff would be willing to accept it even though his expected gain from going to trial would be \$6,600 (i.e., 76 percent $\times$ \$10,000 - \$1,000) because the expected utility of going to trial would be 105.10 (i.e., 76 percent $\times \sqrt{14,000}$ + 24 percent $\times \sqrt{4,000}$), which is less than the utility of the settlement, namely, 105.77 (i.e., $\sqrt{5,000} + 6,400$).⁴⁸

The effect of risk aversion under the American system, as illustrated in this example, is different from its effect under the other systems, for, as was observed in the last section, the risk inherent in going to trial is least under the American system and greatest under the British system. Accordingly, *the effect of risk aversion—the promotion of settlement—should be greatest under the British system, intermediate under the sys-*

costs will be approximately equal to the defendant's costs alone. Similarly, under the system favoring the defendant, the sum of expected legal costs will be approximately equal to the plaintiff's costs. In either case, then, the sum of expected costs is less than the sum of costs under the American system, namely, the sum of the plaintiff's and the defendant's legal costs.

⁴⁷ See proposition 2a and fig. 3.

⁴⁸ In fact, there would be settlement so long as the plaintiff's estimate of the chance of prevailing were less than 79.88 percent. If $p = 79.88$ percent, then the plaintiff's expected utility were he to go to trial of $p \sqrt{14,000} + (1 - p) \sqrt{4,000}$ would just equal his utility of $\sqrt{11,500}$ (were the defendant to make his maximum offer of \$6,500). If the plaintiff were more (less) risk averse or if the defendant were risk averse, this threshold probability would be higher (lower).

*tems favoring the plaintiff and the defendant, and least under the American system.*⁴⁹

IV. CONCLUDING DISCUSSION

a) Although the model considered here abstracted from a variety of issues pertinent to litigation under the different methods for allocating legal costs, the results obtained should be of use for purposes of explanation or of prediction.⁵⁰ To illustrate, we will ask what would be expected to follow from adoption of the British system for the allocation of costs of litigation arising in two areas, that of automobile accidents and of accidents caused by "ultrahazardous" activities.

Consider first the situation in regard to automobile accidents. If in these accidents a harmed party's likelihood of prevailing—of proving that the other party was at fault⁵¹—is generally "low," then the number of suits (or threats to bring suit) should fall. As was explained in the analysis, because losing at trial is more costly under the British system than under the American, suit is less attractive to a plaintiff under the former system when his chance of winning is below a critical probability, this probability depending on the relative magnitudes of the plaintiff's and the defendant's legal costs and on the plaintiff's attitude toward risk. When, for instance, both sides' costs would be about equal and the plaintiff can be assumed to behave in an approximately risk-neutral manner, the analysis implies that any likelihood of prevailing which is less than 50 percent would be low enough to make suit less attractive, and therefore less probable, under the British system. And when the plaintiff can be assumed to behave in a risk-averse way, then the analysis suggests that suit would become even

⁴⁹ As was true in respect to suit, it is difficult to find appealing assumptions under which this is always the case, but intuition suggests it is the typical result. Thus, for instance, it could well be that if one or the other of the parties is very risk averse there would be a settlement under the British system and a trial under the American system, reversing the previous result that the likelihood of a trial is greater under the British system than under the American system.

⁵⁰ I am well aware of many of these excluded issues, for example, of those related to use of contingent fee arrangements, or to the fact that legal costs are not fixed but variable (and are likely to be influenced by the method for allocating costs) or to the possibility that counsel's incentives may differ from his client's, or to the nonfinancial aspects of legal outcomes. However, as was indicated at the outset, it was necessary to exclude all but a few issues to gain an understanding through the use of theory of our subject, for it is a complicated one. If the theoretical investigation has been successful, it will allow the analyst studying a particular area and striving for realism to determine the importance and role of the excluded factors more easily than would otherwise have been possible.

⁵¹ Consider a state where some form of the negligence rule determines liability (rather than no-fault legislation).

less attractive under the British system, since under that system the variability in the possible outcomes is increased. However, because the importance of the factor of plaintiffs' risk aversion is limited by the extent to which individuals own automobile liability insurance and their insurers (who presumably act in a risk-neutral way) determine whether to bring suit, the assumption of risk neutrality may be the appropriate one to keep in mind. With regard to the effect of use of the British system on the number of trials as opposed to the number of suits, recall that the likelihood of a trial conditional on suit having been brought increases under the British system if the litigants are taken to be risk neutral. But a decline in the number of suits could, and in the case under discussion, plausibly would, outweigh the effect of an increase in the conditional likelihood of trial, thereby resulting in a lower absolute number of trials. Thus, the prediction is that adoption of the British system would reduce the number of suits and the number of trials.

The situation in respect to accidents caused by ultrahazardous activities is different. In accidents of this type, one suspects that a harmed party's likelihood of prevailing would usually be high, for the outcome at trial would generally be based on the doctrine of strict liability.⁵² Accordingly, adoption of the British system would be expected to make suit more attractive to plaintiffs and consequently to increase its frequency. Additionally, as noted above, the likelihood of trial conditional on suit would also rise if parties are assumed to behave in a risk-neutral way. However, this assumption may often be inappropriate; plaintiffs' risk aversion would seem a more important consideration in the present context than in that of automobile accidents, for insurers probably play a lesser role in the present context. And if plaintiffs' (or defendants') risk aversion is a sufficiently important consideration, the conditional likelihood of trial would fall under the British system due to the greater variability of financial outcomes. Nevertheless, one would suppose that even if the conditional likelihood of trial did fall, the effect of the increase in the number of suits would be determinative. In summary, then, use of the British system would be likely to result in an increase in the number of suits and, more tentatively, in the number of trials as well.

The differences between these two examples point up what was no doubt evident from the analysis—that the consequences of use of alternative systems for allocating legal costs depend very much on context; that a variety of factors need to be taken into account before the theoretical results can be employed to predict or to explain; yet it is hoped that the examples indicate also the feasibility of doing this.

⁵² See, for example, William L. Prosser, *Law of Torts*, 505-516 (4th ed. 1971).

b) The questions addressed in this article were purely descriptive; there was no inquiry made into the social desirability of the methods for the allocation of legal costs. While to have engaged in such an inquiry would have involved us in a task beyond the scope of a single article, we can sketch here the main elements regarding the two issues that would have been of greatest relevance and then can illustrate our remarks by reconsidering briefly our examples concerning the use of the British versus the American systems.

The first issue involves the determination of the socially desirable level of litigation.⁵³ This very general issue is of interest because, as has been the whole point of our analysis, the allocation of legal costs affects the level of litigation; thus, if we wish to know whether the allocation of legal costs should be changed, we must decide whether the current level of litigation should be changed. It will be convenient for us to regard the socially appropriate level of litigation as reflecting certain social benefits and social costs. For simplicity, let us assume the social benefits of litigation to inhere only in its deterrent and compensatory roles, and the social costs to be comprised of the legal costs borne by plaintiffs and by defendants plus those public administrative expenses associated with the operation of the courts.

On this view, litigation will be worthwhile to society as long as its deterrent and compensatory value exceeds total legal costs plus public administrative expenses. By contrast, litigation will be worthwhile to potential plaintiffs when its benefits to them, namely, their expected judgments or settlements, exceed its costs to them, their expected legal expenses. Thus, it is apparent that the private benefit and cost calculation may diverge from the social, and in any direction.⁵⁴ Accordingly there can be no general presumption about the relation between the actual and the desirable levels of litigation, or, therefore, about the relative appeal of the methods for allocating legal costs. But of course conclusions can be drawn in particular areas of litigation. (Moreover, in any area of litigation

⁵³ The point of view described in this paragraph is developed in Steven Shavell, *The Social versus the Private Incentive to Bring Suit*, forthcoming in *J. Legal Stud.*

⁵⁴ Specifically, in respect to benefits it is clearly possible for the social benefits to outweigh the private—perhaps the deterrent effect of liability would be especially large in relation to plaintiffs' gains—and clearly possible also for the private benefits to exceed the social benefits—perhaps plaintiffs' gains are large in relation to the deterrent effect of liability and no special need for compensation of plaintiffs through use of the legal system exists. On the other hand, in respect to costs, it appears that the social costs would be likely to outweigh the private, for total legal costs plus public administrative expenses exceed plaintiffs' expected legal costs under any method for the allocation of legal costs. In other words, the divergence between social and private costs suggests a tendency toward excessive litigation, and that between social and private benefits either a reinforcing or contrary tendency, the consequence being that the net effect is indeterminate.

there will be opportunity for certain plaintiffs to bring "nuisance" suits, and to deter such suits, it is evident that the British system or the system favoring the defendant would be desirable.)⁵⁵

Let us now briefly note the second issue that seems necessary to consider for normative analysis. This issue concerns the sense of fairness with which a method for the allocation of legal costs is likely to be regarded.⁵⁶ When a party prevails at trial, it is often said that it is not fair that he should have to bear the "loss" of his legal costs, and thus it is said that the British system is fair and the American system unfair (the other two systems falling in between). One presumes that this claim about fairness is associated with the idea that an outcome at trial establishes the moral superiority of the prevailing party or the blameworthiness of the losing party. Hence, the weight accorded by the analyst to the notion of fairness should probably depend on whether the parties would be expected to view the decision at trial as morally significant.

Given the consideration of fairness and the previous issue concerning the socially desirable level of litigation, how might we evaluate the American and British systems for allocating legal costs in our two examples? With regard to automobile accidents, suppose that it is believed that the volume of litigation is too large under the American system. Suppose, for instance, that it is felt that the social benefits of litigation are low—that the deterrent effect of tort liability is negligible (the primary determinants of driving behavior being criminal liability and the desire not to be injured); that the compensatory role of tort liability is limited (private and social insurance providing satisfactory alternative means of compensation); and that legal costs and public administrative expenses are significant. Then, in view of plaintiffs' often substantial gains from bringing suit, the actual volume of litigation could indeed be thought to exceed the desirable level. If this were the belief and if, as suggested above, adoption of the British system would result in a decline in the number of suits and trials, then use of that system would enhance social welfare relative to the goal of achieving an appropriate volume of litigation. However, the de-

⁵⁵ See note 19, *supra*. The deterrence of nuisance suits should be seen as a logically distinct goal from that of altering the overall rate of litigation. Presumably, we would always wish to deter nuisance suits—because they are suits with rather special characteristics that make them identifiable as subversive of the purpose of the legal process. But most suits (or potential suits) do not have the characteristics of nuisance suits, and as we have explained, there is no general presumption about whether we would wish to increase or reduce the overall level of this usual type of suit.

⁵⁶ Although the reader may regard inclusion of the goal of adherence to a sense of fairness as "noneconomic" and thus out of place in an analysis using the methodology of microeconomic theory, that would be a mistake. If individuals care about an aspect of the legal process, that should enter into a normative calculus (or, a descriptive model) just as should, say, the number of accidents or litigation costs; indeed, from the viewpoint of formal theory, there is no distinction among any of these variables.

gree to which use of the British system would enhance welfare in virtue of its fairness might be thought small, assuming that insurers would typically be the litigants and bear the legal costs. Furthermore, even were insurers not the litigants, the relevance of the aspect of fairness is made problematic by the fact that automobile accidents are frequently not obviously due entirely to the fault of any one party.

In respect to accidents caused by ultrahazardous activities, suppose that, as in the area of automobile accidents, the amount of litigation is felt to be sufficient.⁵⁷ Under this assumption adoption of the British system would appear to affect litigation undesirably because it would lead to an increase in its level. But, of course, consideration of fairness would work in favor of the British system and might be given more weight than in the area of automobile accidents, assuming that insurers have a lesser role and that because of the very nature of ultrahazardous activities and the use of a strict liability approach, prevailing plaintiffs would often feel that they are strongly in the right.

From these two examples and our general discussion, the importance of context is again apparent, and it is clear that context matters not only because the effects of the systems on the litigation rate may depend on the area of litigation, but also because the social desirability of altering the underlying litigation rate and the weight to be given the consideration of fairness may depend on the area of litigation.

APPENDIX

As the model has been described in the text, it will be sufficient here to define notation and very briefly state assumptions before presenting the formal analysis. Let

- p = plaintiff's probability of prevailing at trial, where $0 < p < 1$;
- w = magnitude of trial judgment, where $w \in [a, b]$ and $b > a > 0$;
- $F(\cdot)$ = plaintiff's cumulative probability distribution function (c.d.f.) over w , conditional on prevailing (F assumed nondegenerate);
- x = plaintiff's legal costs (before any reallocation of such costs) of going to trial, where $x \geq 0$;
- u = plaintiff's initial wealth;
- $U(\cdot)$ = plaintiff's (von Neumann-Morgenstern) utility function (in wealth);
- q = defendant's probability that *plaintiff* will prevail; where $0 < q < 1$;
- $G(\cdot)$ = defendant's c.d.f. over w , conditional on plaintiff prevailing (G assumed nondegenerate);
- y = defendant's legal costs of going to trial, where $y \geq 0$;
- v = defendant's initial wealth;
- $V(\cdot)$ = defendant's (von Neumann-Morgenstern) utility function (in wealth); and
- s = amount (if any) paid by defendant to plaintiff in settlement, where $s > 0$.

⁵⁷ The reader is reminded that this is only an illustrative hypothesis.

Several points should be noted. (i) The plaintiff's and the defendant's (subjective) beliefs are allowed to differ both as to who will prevail and as to the trial judgment. (ii) The legal costs x and y are taken as fixed. Also, it is assumed that they are borne only if there is a trial; otherwise (if there is no suit or if there is a suit and a settlement) neither side bears any legal costs. (iii) When the parties are presumed to be risk neutral, their utility functions will be taken to equal wealth (and their levels of initial wealth will be ignored) without loss of generality.

By the plaintiff's bringing suit is meant that he threatens the defendant with going to trial and that either a trial actually occurs or else the defendant settles with the plaintiff. This was illustrated in figure 1 as the plaintiff's choosing the action "bring suit" in the first step of a two-step game. It is assumed that the plaintiff will bring suit if and only if going to trial would increase or leave equal⁵⁸ his expected utility. This assumption is in accord with that of "perfect equilibrium" of game theory, under which players only make threats that they would later be willing to carry through.⁵⁹ In addition, the assumption is clearly equivalent to an assumption that the defendant would not settle if the plaintiff's expected utility from going to trial were less than that of not going to trial.

PROPOSITION 1: (a) A risk-neutral plaintiff will bring suit if and only if his estimate of the expected value of the trial judgment exceeds his estimate of the expected legal costs he would bear. Consequently, under the American system, he will bring suit precisely when

$$p \int_a^b w dF(w) \geq x; \quad (1)$$

under the British system, he will bring suit when

$$p \int_a^b w dF(w) \geq (1 - p)(x + y); \quad (2)$$

under the system favoring the plaintiff, he will bring suit when

$$p \int_a^b w dF(w) \geq (1 - p)x; \quad (3)$$

and under the system favoring the defendant, he will bring suit when

$$p \int_a^b w dF(w) \geq x + (1 - p)y. \quad (4)$$

It follows that the number of suits will be greatest under the system favoring the plaintiff and least under the system favoring the defendant. However, a comparison between the American and British systems depends on the plaintiff's beliefs:

⁵⁸ The assumption that the plaintiff will bring suit when the expected utility of going to trial equals the utility of not doing so is of course a convention, and we will use without comment a similar convention when we analyze settlement.

⁵⁹ See, for example, Reinhard Selten, Reexamination of the Perfectness Concept of Equilibrium Points in Extensive Games, 4 Int'l J. Game Theory 25 (1975).

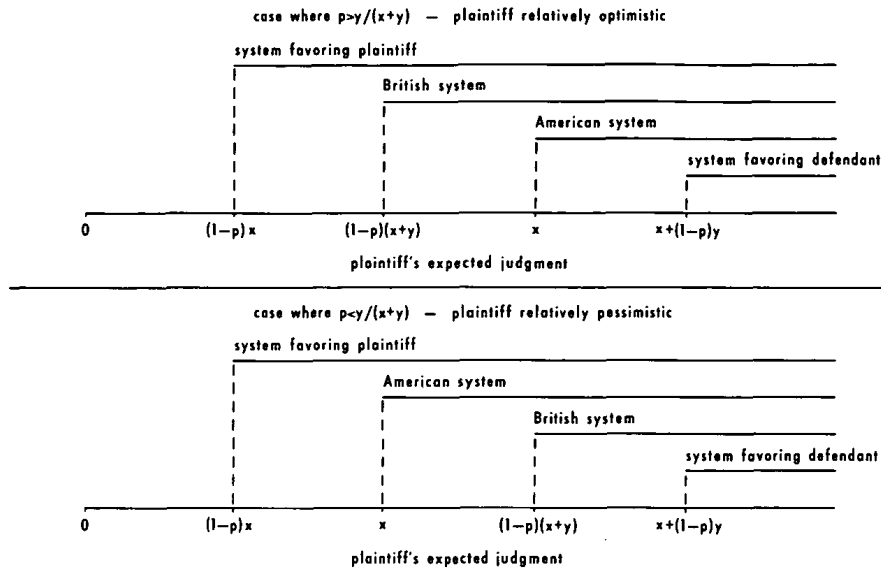


FIG. 2
REGION OF SUIT
(as a function of the size of plaintiff's expected judgment)

There will be more suits under the British system if the plaintiff is sufficiently optimistic about prevailing, and specifically if

$$p > y/(x + y); \quad (5)$$

but otherwise (if the inequality [5] is reversed) there will be more suits under the American system. (Figure 2 summarizes the propensity to sue under the different systems.)

b) A risk-averse plaintiff will bring suit strictly less often than a risk-neutral plaintiff.⁶⁰

PROOF: With regard to part a, formulas (1)–(4) follow immediately from our assumptions and the definitions of the various systems for payment of legal costs.

Since $(1 - p)x$ is less than x , $(1 - p)(x + y)$, and also $x + (1 - p)y$, clearly (3) may be satisfied when (1) or (2) or (4) is not, and if (1) or (2) or (4) is satisfied, (3) must be. Thus there are more suits under the system favoring the plaintiff than

⁶⁰ It can also be shown true of a risk-averse plaintiff that (i) the number of suits will be greatest under the system favoring him and (ii) least under the system favoring the defendant. With regard to the frequency of suit under the American and the British systems, one can show a partial analogue to (iii) of proposition 1a: given x , y , and F , then if p is sufficiently high, the plaintiff will sue under the British system, whereas he might not under the American system. However, there is no necessary effect of the plaintiff's risk aversion on the relative frequency of suit under the American versus British systems: Examples can be constructed in which if the plaintiff is risk averse, the ratio of suits under the American system to suits under the British system is higher than if the plaintiff is risk neutral; examples can also be constructed in which the ratio is lower.

under the other three systems. Since $x + (1 - p)y$ exceeds x , $(1 - p)(x + y)$, and $(1 - p)x$, it is also clear that there are fewer suits under the system favoring the defendant than under the other systems. With regard to the American and British systems, note that $(1 - p)(x + y) \geq x$ if and only if $y/(x + y) \geq p$. Thus if and only if (5) holds is it true that (1) implies (2) but (2) does not imply (1); that is, there are more suits under the British system.

Part *b* will be shown for the American system. (The argument under the other systems is virtually the same.) It will first be shown that if a risk-averse plaintiff brings suit, so would a risk-neutral plaintiff. Now if a risk-averse plaintiff brings suit, then

$$p \int_a^b U(u + w - x) dF(w) + (1 - p)U(u - x) > U(u). \quad (6)$$

But since U is strictly concave, Jensen's inequality⁶¹ gives

$$U\left(u + p \int_a^b w dF(w) - x\right) > p \int_a^b U(u + w - x) dF(w) + (1 - p)U(u - x). \quad (7)$$

Hence

$$U\left(u + p \int_a^b w dF(w) - x\right) > U(u),$$

so that

$$p \int_a^b w dF(w) > x,$$

which, by (1), means that a risk-neutral plaintiff would also bring suit. To complete the proof of *b*, an example will be given in which a risk-neutral plaintiff brings suit but a risk-averse plaintiff would not. Consider an x and F such that

$$\int_a^b w dF(w) > x, \quad (8)$$

and define probabilities $p(\epsilon)$ for small nonnegative ϵ by

$$p(\epsilon) \int_a^b w dF(w) - x = \epsilon. \quad (9)$$

⁶¹ See, for example, Morris De Groot, *Optimal Statistical Decisions* 97 (1970).

(Eq. [8] assures that $p(\epsilon) > 0$.) Using (7) and (9), we have

$$p(0) \int_a^b U(u + w - x) dF(w) + [1 - p(0)] U(u - x) < U(u), \quad (10)$$

so that a risk-averse plaintiff would not bring suit. Since (10) is a strict inequality and the left-hand side is continuous in the probability of prevailing, the risk-averse plaintiff would still be discouraged from bringing suit for $p(\epsilon)$ for small $\epsilon > 0$. But by (9) and (1), a risk-neutral plaintiff would bring suit for such probabilities. Q.E.D.

It is assumed that if the plaintiff brings suit, there will be a settlement if and only if there exists some settlement amount which would result in a strictly higher utility for each party than his expected utility of going to trial.

PROPOSITION 2: (a) A risk-neutral plaintiff who has brought suit and a risk-neutral defendant will go to trial rather than settle if and only if the plaintiff's estimate of the expected value of the trial judgment exceeds the defendant's estimate of it by at least the sum of their respective expected legal costs. Hence, under the American system, there will be a trial when

$$p \int_a^b w dF(w) - q \int_a^b w dG(w) \geq x + y; \quad (11)$$

under the British system, there will be a trial when

$$p \int_a^b w dF(w) - q \int_a^b w dG(w) \geq (1 - p)(x + y) + q(x + y); \quad (12)$$

under the system favoring the plaintiff, there will be a trial when

$$p \int_a^b w dF(w) - q \int_a^b w dG(w) \geq (1 - p)x + (qx + y); \quad (13)$$

and under the system favoring the defendant, there will be a trial when⁶²

$$p \int_a^b w dF(w) - q \int_a^b w dG(w) \geq [x + (1 - p)y] + qy. \quad (14)$$

The comparison of rates of litigation (conditional on suit having been brought) depends on a variety of factors. (i) Suppose first that the plaintiff's chance of prevailing exceeds the defendant's estimate of the plaintiff's chance; that is, sup-

⁶² Moreover, it will be clear that (conditional on suit having been brought) under the American system if (11) does not hold, the settlement s will be in the interval $(p \int_a^b w dF(w) - x, q \int_a^b w dG(w) + y)$; under the British system, if (12) does not hold, s will be in $(p \int_a^b w dF(w) - (1 - p)(x + y), q \int_a^b w dG(w) + q(x + y))$; under the system favoring the plaintiff, if (13) does not hold, s will be in $(p \int_a^b w dF(w) - (1 - p)x, q \int_a^b w dG(w) + qx + y)$; and under the system favoring the defendant, if (14) does not hold, s will be in $(p \int_a^b w dF(w) - [x + (1 - p)y], q \int_a^b w dG(w) - qy)$.

pose $p > q$. Then the likelihood of trial will be greatest under the British system and least under the American system, and the likelihood of trial under the system favoring the plaintiff will be greater or less than that under the system favoring the defendant according as the plaintiff's legal costs are greater or less than the defendant's. On the other hand, (ii) suppose that the plaintiff's chance of prevailing is less than the defendant's estimate; that is, $p < q$.⁶³ Then the situation is reversed: The likelihood of trial will be greatest under the American system and least under the British system, and the likelihood of trial under the system favoring the plaintiff is greater or less than that under the system favoring the defendant according as the plaintiff's legal costs are less or greater than the defendant's. However, (iii) if the plaintiff's chance of prevailing equals the defendant's estimate, $p = q$, then the likelihood of trial is not affected by the system determining payment of legal costs. (These results are illustrated in fig. 3.)

b) If one or both of the parties are risk averse, then they will settle strictly more often than if both were risk neutral.

PROOF: To prove part *a*, let l_π designate the plaintiff's expected legal costs and l_δ the defendant's. Thus under the American system, $l_\pi = x$ and $l_\delta = y$; under the British system, $l_\pi = (1 - p)(x + y)$ and $l_\delta = q(x + y)$; under the system favoring the plaintiff, $l_\pi = (1 - p)x$ and $l_\delta = qx + y$; and under the system favoring the defendant, $l_\pi = x + (1 - p)y$ and $l_\delta = qy$.

Under our assumptions, there will be a settlement if there exists a settlement amount s such that both the plaintiff and the defendant would be made better off by settling for s than if they had gone to trial. Thus, there will be a settlement if there is an s such that

$$s > p \int_a^b w dF(w) - l_\pi \quad (15)$$

and

$$s < q \int_a^b w dG(w) + l_\delta. \quad (16)$$

Subtracting (16) from (15) and rearranging terms gives

$$p \int_a^b w dF(w) - q \int_a^b w dG(w) < l_\pi + l_\delta. \quad (17)$$

Conversely, if (17) holds, there exist s satisfying (15) and (16). For example if

$$s = p \int_a^b w dF(w) - l_\pi + 1/2 \left\{ l_\pi + l_\delta - \left[p \int_a^b w dF(w) - q \int_a^b w dG(w) \right] \right\}, \quad (18)$$

⁶³ This case and the next, when $p = q$, are probably much less important than the first. The reason is that when $p < q$, the absolute number of trials is likely to be very low under all the systems. In fact, suppose (as might often be true) that when the plaintiff's opinion about the likelihood of prevailing is pessimistic relative to the defendant's, so is the plaintiff's opinion about the trial judgment. Then there will be no litigation whatever: if when $p < q$, then $\int_a^b w dF(w) < \int_a^b w dG(w)$, it follows that (11)–(14) cannot be satisfied.

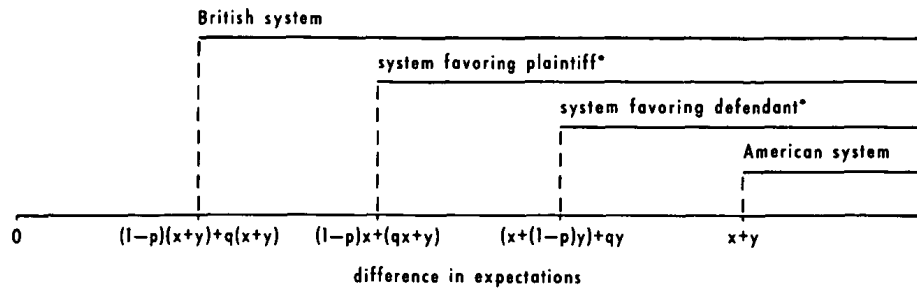
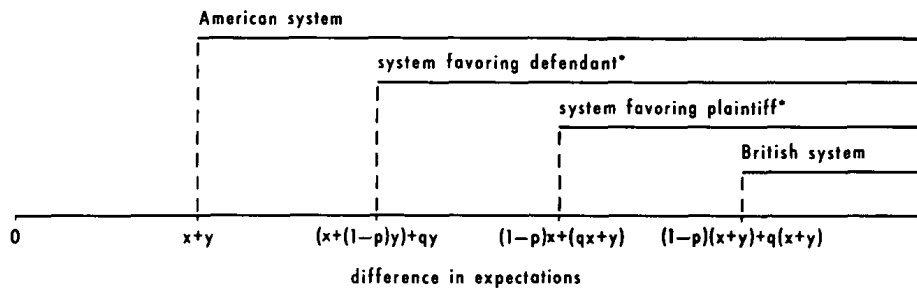
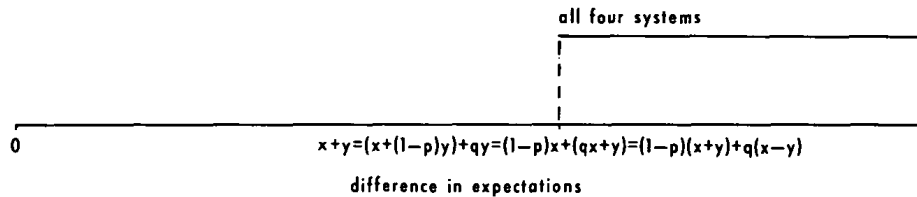
(a) case when $p > q$ (b) case when $p < q$ (c) case when $p = q$ 

FIG. 3

REGION OF LITIGATION

(as a function of the difference between the plaintiff's estimate of the expected judgment and the defendant's) (asterisk indicates that the relative positions of those two systems are as shown if $x > y$; if $x < y$, they are reversed; if $x = y$, they are coincident)

then (15) holds; and so does (16), since from (18) and (17),

$$s < p \int_a^b w dF(w) - l_\pi + \left\{ l_\pi + l_\delta - \left[p \int_a^b w dF(w) - q \int_a^b w dG(w) \right] \right\} \quad (19)$$

and the right-hand side reduces to

$$q \int_a^b w dG(w) + l_\delta.$$

Thus, there will be settlement if and only if (17) holds, and hence trial if and only if the opposite inequality holds. From this and the definitions of l_π and l_δ (and what they were observed to equal under the different rules), (11)–(14) immediately follow. And from (11)–(14), items i–iii and figure 3 also follow. For example, suppose $p > q$. Then

$$(x + y) > (1 - p + q)(x + y) = (1 - p)(x + y) + q(x + y). \quad (20)$$

Thus, (11) and (12) imply that a trial is more likely under the British system. The other relationships are similarly established.

Part *b* will be shown only for the American system. (The arguments under the other systems are similar.) It will first be shown that if a risk-neutral plaintiff and risk-neutral defendant would settle, then a risk-neutral plaintiff and a risk-averse defendant would also settle. Then it will be shown that the converse may not be true. (The cases of a risk-averse plaintiff and a risk-neutral defendant or of a risk-averse plaintiff and a risk-averse defendant are similar.) Now if a risk-neutral plaintiff and a risk-neutral defendant settle for s , then, in particular, rewriting (16) for the American system, we obtain

$$s < q \int_a^b (w + y) dG(w) + (1 - q)(y). \quad (21)$$

Multiplying by -1 , and adding to both sides v , the initial wealth of the defendant, we get

$$v - s > q \int_a^b (v - w - y) dG(w) + (1 - q)(v - y). \quad (22)$$

If the defendant is risk averse, then (22) and Jensen's inequality imply

$$\begin{aligned} V(v - s) &> V\left(q \int_a^b (v - w - y) dG(w) + (1 - q)(v - y)\right) \\ &> q \int_a^b V(v - w - y) dG(w) + (1 - q)(v - y). \end{aligned} \quad (23)$$

Thus the risk-averse defendant and the risk-neutral plaintiff would also settle. On the other hand, consider a p , F , q , G , x , and y such that

$$p \int_a^b w dF(w) - q \int_a^b w dG(w) - (x + y) > 0, \quad (24)$$

and define positive probabilities $p(\epsilon)$ for small nonnegative ϵ by

$$p(\epsilon) \int_a^b w dF(w) - q \int_a^b w dG(w) - (x + y) = \epsilon. \quad (25)$$

Note that this means that for $\epsilon > 0$, a risk-neutral plaintiff would bring suit (for [25] implies $p(\epsilon) \int_a^b w dF(w) > x$), but would not settle with a risk-neutral defendant (by [11]). Now assume that the defendant is risk averse and consider a potential settlement offer from him of $s(\epsilon)$, where

$$s(\epsilon) = q \int_a^b w dG(w) + y + 2\epsilon. \quad (26)$$

By Jensen's inequality,

$$\begin{aligned} & q \int_a^b V(v - w - y) dG(w) + (1 - q)V(v - y) \\ & < V\left(v - \int_a^b w dG(w) - y\right) = V(v - s(0)), \end{aligned} \quad (27)$$

so that the defendant would strictly prefer to settle for $s(0)$ than to go to trial. Since the inequality is strict and $s(\epsilon)$ is continuous in ϵ , the defendant would also strictly prefer to settle for $s(\epsilon)$ for $\epsilon > 0$ and sufficiently small. Also, the plaintiff who thinks he will prevail with probability $p(\epsilon)$ would prefer to settle for $s(\epsilon)$, assuming $\epsilon > 0$, than to go to trial: from (25) and (26), we have

$$s(\epsilon) = p(\epsilon) \int_a^b w dF(w) - x - \epsilon + 2\epsilon > p(\epsilon) \int_a^b w dF(w) - x \quad (28)$$

for $\epsilon > 0$. Thus, if ϵ is sufficiently small, there would be a settlement between a risk-neutral plaintiff and the risk-averse defendant even though there would be a trial between a risk-neutral plaintiff and a risk-neutral defendant. Q.E.D.

