Sanctioning Frivolous Suits: An Economic Analysis

A. Mitchell Polinsky

Daniel L. Rubinfeld

Berkeley Law

Follow this and additional works at: https://scholarship.law.berkeley.edu/facpubs

Part of the Economics Commons, and the Law Commons

Recommended Citation
ARTICLES

Sanctioning Frivolous Suits: An Economic Analysis

A. MITCHELL POLINSKY AND DANIEL L. RUBINFELD*

An often-voiced concern in the United States is that there are too many frivolous lawsuits.¹ In this article we analyze the desirability of imposing sanctions on plaintiffs found to have filed frivolous suits. We also describe the optimal design of such sanctions, including the choice of their magnitude and whether to award them to defendants.

Our interest in this topic was stimulated by the controversy surrounding Rule 11 of the Federal Rules of Civil Procedure, which provides for the imposition of sanctions on individuals who present in court a “pleading, written motion, or other paper” that is deemed to be frivolous.² Although

---

¹ For example, in a recently reported survey of jurors in cases in which businesses or corporations were defendants, 83% of the jurors indicated that they “agree/strongly agree” with the statement, “[t]here are far too many frivolous lawsuits today.” Only 5% said that they “disagree/strongly disagree” with the statement. Valerie P. Hans & William S. Lofquist, Jurors’ Judgments of Business Liability in Tort Cases: Implications for the Litigation Explosion Debate, 26 LAW & SOC’Y REV. 85, 95 (1992).

² The rule reads in part as follows:

(b) Representation to Court. By presenting to the court . . . a pleading, written motion, or other paper, an attorney or unrepresented party is certifying that to the best of the person’s knowledge, information, and belief, formed after an inquiry reasonable under the circumstances,—

(1) it is not being presented for any improper purpose, such as to harass or to cause unnecessary delay or needless increase in the cost of litigation;

(2) the claims, defenses, and other legal contentions therein are warranted by existing law or by a nonfrivolous argument for the extension, modification, or reversal of existing law or the establishment of new law; . . .

c) Sanctions. If . . . the court determines that subdivision (b) has been violated, the court may . . . impose an appropriate sanction upon the attorneys, law firms, or parties that have violated subdivision (b) or are responsible for the violation.

(1) How Initiated.

397
the system of sanctions that we analyze does not attempt to capture all of the legal and institutional details of Rule 11, it parallels that rule in several important respects. For example, under Rule 11, financial penalties are the norm (usually equal to the nonfrivolous party’s litigation costs); the penalty typically is imposed as a result of a motion filed by the opposing party; and the penalty usually is given to that party. Similarly, in our analysis, the sanction consists of a financial penalty; the penalty is imposed as the result of a legal action initiated by the nonfrivolous party; and at

(A) By Motion. . . . If warranted, the court may award to the party prevailing on the motion the reasonable expenses and attorney’s fees incurred in presenting or opposing the motion. . . .

(B) On Court’s Initiative. . . .

(2) Nature of Sanction; Limitations. A sanction imposed for violation of this rule shall be limited to what is sufficient to deter repetition of such conduct or comparable conduct by others similarly situated. . . . [T]he sanction may consist of, or include, . . . an order to pay a penalty into court, or, if imposed on motion, and warranted for effective deterrence, an order directing payment to the movant of some or all of the reasonable attorneys’ fees and other expenses incurred as a direct result of the violation.


3. A study of all district court cases between August 1, 1983 and August 1, 1985 in which Rule 11 sanctions were considered found that in 96% of the cases in which sanctions were imposed the courts awarded “‘reasonable’ costs and attorney fees” to the opposing party and “[o]nly a handful of courts have imposed sanctions other than, or in addition to, fees and expenses.” Melissa L. Nelken, Sanctions Under Amended Federal Rule 11—Some “Chilling” Problems in the Struggle Between Compensation and Punishment, 74 GEO. L.J. 1313, 1333 (1986) (footnote omitted). A survey of nearly 4500 attorneys practicing in district courts in three federal judicial circuits found that “the overwhelming majority (95%) of sanctions were . . . monetary.” Lawrence C. Marshall et al., The Use and Impact of Rule 11, 86 NW. U. L. Rev. 943, 956-57 (1992) (footnote omitted); see also Georgene M. Vairo, Rule 11: A Critical Analysis, 118 F.R.D. 189, 227 (1988) (“[T]he numerous cases decided under amended Rule 11 thus far suggest that a monetary sanction providing attorney fees is the preferred sanction.”); Elizabeth C. Wiggins et al., The Federal Judicial Center’s Study of Rule 11, 1991 FJC DIRECTIONS 18 (“Rule 11 has operated predominantly as a cost-shifting device.”). But see Herbert Kritzer et al., Rule 11: Moving Beyond the Cosmic Anecdote, 75 JUDICATURE 269, 270 (1992) (“[T]he dollar amount of these [Rule 11] sanctions calls into question conclusions by some that . . . cost shifting has become the norm.”).

4. Although the court has the power to initiate a Rule 11 sanction, it is much more common for the sanction to be imposed as a result of a motion by one of the parties. See Vairo, supra note 3, at 220 (“[M]ost courts seem to prefer to have a party make a Rule 11 motion.”).

5. In a study in the Third Circuit, researchers found that of the 21 cases in which monetary sanctions were imposed (out of a total of 27 cases in which sanctions of any kind were imposed), 18 were awarded to the opposing party. STEPHEN B. BURBANK, RULE 11 IN TRANSITION 36-37 (1989). Nelken’s 1986 study of Rule 11 sanctions found an even greater likelihood of the award going to the opposing party. See Nelken, supra note 3, at 1333 (finding that in 96% of the cases in which sanctions were imposed, the courts awarded reasonable costs and attorney fees to the opposing party); see also Wiggins et al., supra note 3, at 18 (“Rule 11 sanctions have typically taken the form of monetary fees payable to an opposing party.”).
least some of the penalty is awarded to that party. Additionally, Rule 11 motions most commonly are brought in response to the filing of frivolous suits (or claims); our focus, too, is on the filing of frivolous suits. Because the system of sanctions that we analyze resembles the operation of Rule 11, we will use Rule 11 as a springboard for our discussion.

Rule 11 sanctions have been the subject of an extensive debate by legal practitioners, judges, and academic commentators, especially since 1983, when the availability of such sanctions was substantially expanded. Some

6. See Marshall et al., supra note 3, at 953 (finding that "the most common reason for actual sanctions, as well as for motions and show cause orders, is the filing of allegedly frivolous suits or claims").

7. We recognize that the application of Rule 11 is not the only way to control frivolous litigation:

Rule 11 is one of several devices in the procedural system for controlling abusive conduct by lawyers. As the Supreme Court has recently told us, there is inherent power to deal with gross out-of-court abusive conduct. Chambers v. Nasco, Inc., 59 L.W. (June 6, 1991). Rule 37 deals with discovery abuse, 28 U.S.C. § 1927 with unreasonable and vexatious litigation, and Rule 16 provides speedy means to rid cases of frivolous positions. Rule 11 is confined to filed writings.

A. Leon Higginbotham, Jr., et al., Bench-Bar Proposal to Revise Civil Procedure Rule 11, 137 F.R.D. 159, 167 (1991); see generally Ronald J. Gilson, The Devolution of the Legal Profession: A Demand Side Perspective, 49 Md. L. Rev. 869 (1990); David B. Wilkins, Who Should Regulate Lawyers?, 105 Harv. L. Rev. 799 (1992) (discussing various methods for controlling lawyer misconduct, including Rule 11 sanctions). Although we do not discuss these alternatives to Rule 11, it will be clear that our analysis applies to any rule that provides for financial penalties on frivolous litigants upon initiation by nonfrivolous litigants.

8. This expansion occurred through amendments to the Federal Rules of Civil Procedure. The rapidly increasing use of Rule 11 sanctions since 1983 is widely acknowledged. See, e.g., Wilkins, supra note 7, at 838 ("In the eight years since rule 11 was amended . . . there has been an explosion of activity.")}; Marshall et al., supra note 3, at 948 ("Unquestionably, these amendments led to a dramatic increase in Rule 11 activity.").


The controversy over Rule 11 also has percolated into the legal and popular press. See, e.g., Randall Samborn, Rule 11 Reforms Are Criticized, Nat'l L.J., May 25, 1992, at 3, 28 (outlining debate over the proposed Rule 11); Stephen Labaton, Solution to Wasteful Lawsuits Becomes a Problem, N.Y. Times, June 14, 1992, at E2 (same).
of the principal controversies surrounding Rule 11 can be summarized by the following questions:

—Should sanctions be allowed, given that they increase the complexity, and therefore the cost, of litigation?  

—Assuming that sanctions are allowed, should the amount of the penalty be based on the goal of compensating the nonfrivolous party for litigation expenses, or on the goal of deterring potential frivolous litigants?

—If the penalty is based on the deterrence goal, should some or all of the penalty be given to the court or to the general treasury rather than to the opposing party?

9. The committee that submitted the Bench-Bar Proposal to Revise Civil Procedure Rule 11 stated that “[S]ome of us would repeal the Rule . . . .” Higginbotham, Jr., et al., supra note 7, at 162; see also A PROJECT OF THE FEDERAL PROCEDURE COMM., AMERICAN BAR ASS’N, SECTION OF LITIGATION, SANCTIONS: RULE 11 AND OTHER POWERS 16 (2d ed. 1988) ("[A] legitimate question is whether the federal courts need Rule 11 at all."); Gregory Joseph, Redrafting Rule 11, NAT’L L.J., Oct. 1, 1990, at 13 (“There is a strong case to be made that Rule 11 should be abolished.”).

There is widespread agreement that the existence of Rule 11 increases the cost of litigation. See, e.g., Higginbotham, Jr., et al., supra note 7, at 167 (“There has been too much costly satellite litigation over the application of the rule.”). Higher litigation costs undoubtedly are a factor leading some commentators to recommend constraining or abolishing Rule 11. See, e.g., Wiggins et al., supra note 3, at 6 (“One of the issues important to the Advisory Committee was whether the financial cost in satellite litigation exceeded the benefits of the rule.”). But see BURBANK, supra note 5, at 83 (“[S]atellite litigation on Rule 11 issues does not seem to us a serious problem for either litigants or district judges in the Third Circuit.”); Marshall et al., supra note 3, at 985 (“[A]ssertions that large blocks of lawyers’ time are being devoted to Rule 11 activities are not substantiated by the data.”).

10. After stating that “[s]erious consideration should be given to the question of the primary purpose of Rule 11,” Georgene Vairo proceeds to contrast the effects of “shifting the focus from the compensatory purpose to the specific deterrent purpose.” Vairo, supra note 3, at 233. Current practice appears to favor compensation. See supra note 3. The Supreme Court, however, has emphasized the deterrence purpose of Rule 11. See Cooter & Gell v. Hartmarx Corp., 496 U.S. 384, 393 (1990) (Rule 11’s “central purpose . . . is to deter baseless filings”). The amendments to Rule 11 proposed by the Advisory Committee on Civil Rules and recently adopted by the Supreme Court also stress deterrence. See FED. R. CIV. P. 11, supra note 2.

The Committee Notes to the Judicial Conference’s amendments are even more explicit. REPORT ON THE PROPOSED AMENDMENTS, supra note 8, at 53 (“[T]he purpose of Rule 11 sanctions is to deter rather than to compensate.”). The Bench-Bar Proposal to Revise Civil Procedure Rule 11 also adopts a deterrence orientation. See Higginbotham, Jr., et al., supra note 7, at 165 (“The sanction shall be limited to what is sufficient to deter comparable conduct by persons similarly situated.”).

11. In the past, Rule 11 sanctions typically have been given to the other party. See supra note 5. Under the amendments proposed by the Advisory Committee on Civil Rules and adopted by the Supreme Court, however, monetary sanctions “ordinarily [would] be paid into court as a penalty” (although they may be paid to the nonfrivolous party if such payments are needed to achieve deterrence). See Order, supra note 8; REPORT ON THE PROPOSED AMENDMENTS, supra note 8, at 53-54. Under the Bench-Bar Proposal to Revise Civil Procedure Rule 11, all of the penalty would be paid to the court. Higginbotham, Jr., et al., supra note 7, at 165.
—Since sanctions might be imposed by mistake, will they discourage some legitimate plaintiffs from suing, given the risk of being found frivolous?  

—If sanctions are not allowed, are there alternative policies that can discourage frivolous litigation?

This article answers these questions using an economic model of litigation. Part I describes the model in an informal and nontechnical way (a mathematical version appears in the Appendix). Parts II through VI use the model to address the preceding questions.

Throughout this article we will refer to the party who brings the original suit as the “plaintiff” even though that party will be a defendant in an action brought against it for having filed a frivolous suit. Similarly, we will refer to the party against whom the original suit is brought as the “defendant” even though that party will be the plaintiff in an action claiming that the original suit was frivolous.

The key points that emerge from our analysis may be summarized as follows:

First, because of the increased litigation costs associated with sanctions, they should not be universally available. We describe the circumstances under which sanctions against frivolous suits are, and are not, beneficial. All else being equal, sanctions tend to be more desirable the lower the litigation costs associated with their use, the higher the litigation costs in the original trial, the higher the ratio of potential frivolous plaintiffs to legitimate plaintiffs, and the higher the probability that a plaintiff (whether frivolous or legitimate) will prevail in the original trial.

12. Melissa L. Nelken observes, for example, that “[t]he threat of sanctions may deter not only frivolous cases, but also potentially meritorious cases from being filed and pursued.” Nelken, supra note 3, at 1340 (footnote omitted). In a subsequent survey in the Northern District of California, Nelken reported:

When asked whether they had not asserted claims or defenses that they considered meritorious due to concern about Rule 11 sanctions, 15 per cent [of the lawyers] said yes; and 14 per cent said that they had not accepted certain cases for the same reason.

Melissa L. Nelken, The Impact of Federal Rule 11 on Lawyers and Judges in the Northern District of California, 74 JUDICATURE 147, 150 (1990); see also Vairo, supra note 3, at 232-33 (suggesting that abuse of Rule 11 “is chilling meritorious litigation and effective advocacy”).

13. Most commentaries about Rule 11 do not focus on alternative policies. But see Wiggins et al., supra note 3, at 31-32 (reporting on judges' opinions regarding the effectiveness of Rule 11 compared with other methods for controlling groundless litigation); see generally Wilkins, supra note 7 (discussing a variety of methods for controlling lawyer misconduct, including Rule 11 sanctions). We will consider two natural alternatives to Rule 11 that have been suggested for their own merits. See infra Part VI.

14. Rule 11 usually is invoked by motion in the original litigation. Although we will treat the imposition of a sanction against a frivolous plaintiff as the result of a separate action, nothing material to our analysis depends on this characterization.
Second, when sanctions are used, their level generally should be set so as to deter potential frivolous litigants rather than to compensate nonfrivolous parties for their litigation expenses. Although there is a compensation-based argument for imposing penalties, we find the deterrence rationale more convincing (for reasons to be explained). We show that to accomplish deterrence the level of the sanction should exceed the cost to the nonfrivolous party of bringing an action for sanctions; the resulting sanction bears only a tangential relationship to the nonfrivolous party's litigation costs.

Third, assuming that sanctions are designed to deter frivolous behavior, they should be “decoupled” so as to impose a greater penalty on the frivolous party than the amount awarded to the nonfrivolous party; the difference would go to the court or to the general treasury. Decoupling sanctions in this way can accomplish the same degree of deterrence of frivolous suits as a “coupled” system, but with a lower level of litigation costs. In short, if the award to nonfrivolous parties is lowered, fewer actions for sanctions will be brought, thereby saving litigation costs. At the same time, if the amount paid by frivolous parties is raised, deterrence of frivolous behavior can be maintained.

Fourth, although the possibility of mistaken imposition of sanctions tends to discourage legitimate plaintiffs from suing, this effect can be offset by raising the award to successful plaintiffs in the original action. We show that if sanctions are used and the award to successful plaintiffs in the original trial is raised so as to keep the expected costs imposed on the defendant constant, the expected payoff to a legitimate plaintiff from filing a suit actually will rise relative to the payoff in the absence of sanctions. In other words, legitimate plaintiffs are better off if sanctions are used, provided that the award in the original trial is adjusted appropriately.

Fifth, if sanctions are not used because of their greater complexity and cost, there are two alternatives that also can deter frivolous suits: a version of the English rule of fee allocation and the decoupling of liability in the initial trial. The English rule, under which the loser pays the winner's legal costs, tends to discourage frivolous suits because, if frivolous plaintiffs have a higher probability of losing than legitimate plaintiffs, making a losing plaintiff pay the winning defendant's legal costs imposes a higher expected cost on frivolous plaintiffs than on legitimate plaintiffs. We explain, however, why it might be necessary to modify the traditional English rule by making the loser pay an amount greater than the winner's legal costs.

Decoupling liability in the initial trial, such that the award to a prevailing plaintiff is less than what the defendant pays, also can discourage frivolous suits for a similar reason: If frivolous plaintiffs have a lower probability of winning than legitimate plaintiffs, the award can be lowered so that frivolous, but not legitimate, plaintiffs will be discouraged from suing (while the same expected costs can be imposed on defendants by...
making losing defendants pay a higher amount than is awarded to plaintiffs).\textsuperscript{15}

\textsuperscript{15} Our article, together with an independently and concurrently written article, Bruce H. Kobayashi & Jeffrey S. Parker, \textit{No Armistice at 11: A Commentary on the Supreme Court’s 1993 Amendment to Rule 11 of the Federal Rules of Civil Procedure}, 3 SUP. CT. ECON. REV. 93 (1993), is the first to use the economic theory of litigation to analyze the incentives of a nonfrivolous party to bring an action for sanctions against a frivolous party. While Kobayashi and Parker’s framework overlaps with ours in several respects, their focus is on a different aspect of frivolous litigation. Specifically, they analyze the effects of giving a plaintiff the option to withdraw without sanction the filing that provoked a Rule 11 motion by the defendant.

Several law and economics scholars have written about frivolous litigation generally. See, e.g., Lucian Ayre Bebchuk, \textit{Suing Solely to Extract a Settlement Offer}, 17 J. LEGAL STUD. 437 (1988); Avery Katz, \textit{The Effect of Frivolous Lawsuits on the Settlement of Litigation}, 10 INT’L REV. L. & ECON. 3 (1990); David Rosenberg & Steven Shavell, \textit{A Model in which Suits are Brought for their Nuisance Value}, 5 INT’L REV. L. & ECON. 3 (1985). Their focus is on explaining why a defendant might make a settlement offer to a frivolous plaintiff even though the plaintiff would drop the case if no offer were forthcoming (and, to a lesser extent, they discuss how different fee allocation rules affect the incentive to bring frivolous suits). They do not consider actions for sanctions initiated by the nonfrivolous party against the frivolous party.

Three related articles in the law and economics literature deserve special mention. Katz analyzes the effects of penalties imposed on frivolous plaintiffs that are paid entirely to the government. Katz, supra, at 19-20. Since the penalties considered by Katz are paid entirely to the government, presumably they would have to be initiated by the government. He suggests, though not at length, that a scheme that paid the penalty to the defendant would operate similarly. Id. at 19 n.18. Bebchuk and Chang demonstrate that making a plaintiff pay the defendant’s legal fees if a plaintiff’s probability of prevailing is sufficiently low can induce plaintiffs to sue if and only if they believe their cases are sufficiently strong. They also interpret Rule 11 as an example of this type of fee-shifting rule, but do not consider the role of the defendant in implementing this scheme. Lucian Ayre Bebchuk & Howard F. Chang, \textit{An Analysis of Fee-Shifting Based on the Margin of Victory: On Frivolous Suits, Meritorious Suits, and the Role of Rule 11}, Harvard Law School, Program in Law and Economics, Discussion Paper No. 135 (Oct. 1993) (on file with The Georgetown Law Journal). Finally, in another article, we derive the optimal award to a winning plaintiff and the optimal penalty on a losing plaintiff when the probability of prevailing varies among plaintiffs and neither the defendant nor the court can identify a particular plaintiff’s chance of success. A. Mitchell Polinsky & Daniel L. Rubinfeld, \textit{Optimal Awards and Penalties When the Probability of Prevailing Varies Among Plaintiffs}, National Bureau of Economic Research, Discussion Paper No. 4507 (Oct. 1993) (on file with The Georgetown Law Journal). In that article we, too, ignore the role of the defendant in effecting the penalty.

Although many articles in legal periodicals and monographs have been written about Rule 11 sanctions, none of them has directly relied on the economic theory of litigation to inform thinking about the subject. See, e.g., supra note 8. But cf. Gilson, supra note 7; Stephen G. Bené, \textit{Note, Why Not Fine Attorneys?: An Economic Approach to Lawyer Disciplinary Sanctions}, 43 STAN. L. REV. 907 (1991) (considering the relevance of the economic theory of fines, but not the economic theory of litigation, to the issue of lawyer discipline, including Rule 11 sanctions). For related discussions, see John C. Coffee, Jr., \textit{Understanding the Plaintiff’s Attorney: The Implications of Economic Theory for Private Enforcement of Law Through Class and Derivative Actions}, 86 COLUM. L. REV. 669 (1986) (endorsing economic incentives under the private attorney general rules); William C. Baskin III, \textit{Note, Using Rule 9(b) to Reduce Nuisance Securities Litigation}, 99 YALE L.J. 1591, 1594-99 (1990) (discussing how the economic motives of plaintiffs’ counsel in securities class actions leads to the filing of nuisance suits).
I. A FRAMEWORK FOR ANALYZING SANCTIONS

This part describes a relatively simple model of a plaintiff's decision whether to bring a suit and a defendant's decision whether to bring an action for sanctions against the plaintiff for having filed a frivolous suit.\(^{16}\)

A. BASIC ASSUMPTIONS

For simplicity, we presume that if a suit is brought by the plaintiff, it results in a trial (and that any action for sanctions brought by the defendant is resolved similarly).\(^ {17}\) Both the plaintiff and the defendant then incur some trial costs. Furthermore, suppose there are a fixed number of each of two types of potential plaintiffs, frivolous ones and legitimate ones.

It is not necessary for our purposes to define what constitutes a frivolous suit (an issue about which there is much controversy), and we do not propose to undertake that task here. All we need to assume is that, however a frivolous suit is defined, a frivolous plaintiff bringing the same claim as a legitimate plaintiff has a lower probability of prevailing at trial than the legitimate plaintiff.\(^ {18}\) We believe that this assumption is reasonable. For example, suppose that the definition of a frivolous suit includes a suit in which the plaintiff knows that some alleged fact is false; the probability of the plaintiff "proving" this fact at trial presumably will be lower than if the alleged fact were true.\(^ {19}\)

The analysis in this article can be applied to different categories of suits, when (among other things) the probabilities of prevailing for frivolous and legitimate plaintiffs differ among the categories. Thus, for example,
our framework does not preclude the possibility that in one category legitimate suits have a low probability of prevailing (but still higher than for comparable frivolous suits),\footnote{20} or the possibility that in another category frivolous suits have a high probability of prevailing (but still lower than for comparable legitimate suits). The specific conclusions of the analysis may, of course, differ from category to category, but the general thrust of the analysis will be the same.

If a plaintiff of either type wins at trial, he receives some award from the defendant. We assume that the defendant will not then request sanctions. If a plaintiff loses at trial, he pays nothing to the defendant unless the defendant seeks sanctions against him and the defendant prevails in the resulting proceeding.\footnote{21} In that case, the plaintiff pays a penalty to the defendant. If the defendant brings an action for sanctions, both parties incur some additional litigation costs.

The sequence of the decisions of the parties and the outcomes is summarized in Figure 1.\footnote{22}

In both the original trial and the sanctions proceeding, each side bears its own litigation costs; that is, the American rule of fee allocation governs. (This assumption, however, does not prevent the award in the initial trial or the penalty in the sanctions proceeding from including the prevailing party's litigation costs; it simply recognizes that the inclusion of litigation costs is not automatic.)

Neither the defendant nor the court is able to distinguish between frivolous and legitimate plaintiffs at the initial trial. For simplicity, though, we assume that if the defendant seeks sanctions, the type of plaintiff will

\footnote{20. An example of a low-probability legitimate suit might be a suit filed in good faith that attempts to reverse a long-standing law.}

\footnote{21. The assumption that an action for sanctions is brought after the original trial is not essential. For our purposes, the important point is that sanctions are sought after the defendant obtains some information (aside from the fact that a suit has been filed) that allows her to assess the likelihood that the plaintiff's suit is frivolous. For simplicity, we incorporate this point by assuming that the defendant decides whether to bring an action for sanctions after she learns whether the plaintiff has won or lost—plaintiffs who lose are, everything else equal, more likely to be frivolous. (In the context of Rule 11, the information that leads a party to file a motion for sanctions often is obtained early in the litigation process.)}

\footnote{22. Our analysis of a possible action for sanctions following the initial trial is, along with Kobayashi & Parker, supra note 15, one of the first to model litigation that occurs in stages. Of related interest analytically, however, are two articles by William Landes. In the first of these, William M. Landes, \textit{Sequential versus Unitary Trials: An Economic Analysis}, 22 J. LEGAL STUD. 99 (1993), he discusses “sequential” trials, in which a trial to determine the plaintiff's damages occurs only if the plaintiff prevails in a prior trial. In his second article, William M. Landes, \textit{Counterclaims: An Economic Analysis} (Sept. 1993) (unpublished manuscript, on file with \textit{The Georgetown Law Journal}), he considers the possibility that the defendant will bring a “countersuit” against the plaintiff. The second article is closer in spirit to our own, although it differs in two important respects. First, it allows for the possibility of a countersuit even in the absence of an initial suit. Second, it is not concerned with frivolous litigation.}
be correctly identified in the resulting proceeding. (We discuss in a later part the implications of mistakes in the sanctions proceeding.)

Finally, we assume that all parties are risk neutral. (This assumption is modified in a subsequent part.) A risk-neutral party decides what to do on the basis of the “expected value” of a prospective outcome—the probability that the outcome will occur multiplied by the gain or loss if the outcome occurs.

B. DEFENDANT’S DECISION TO BRING AN ACTION FOR SANCTIONS

It will be useful to begin with the second stage of the litigation process, when the defendant decides whether to request sanctions if she has prevailed in the original suit. Although the defendant does not know whether the specific plaintiff over whom she has prevailed is frivolous or legitimate, we assume that the defendant knows the general tendencies of potential frivolous and potential legitimate plaintiffs to sue, and therefore, that she knows the probability that the plaintiff over whom she has prevailed is frivolous. (To be precise, the defendant also must be presumed to know the relative numbers of potential frivolous and legitimate plaintiffs and their respective probabilities of prevailing at trial.)

Given our earlier assumption that a sanctions proceeding will correctly identify the plaintiff’s type, the probability that the defendant will obtain a penalty if she seeks sanctions equals the probability that a losing plaintiff is frivolous. The defendant will bring an action for sanctions if the expected value of the penalty—which equals the probability that the plaintiff is frivolous multiplied by the amount of the penalty—exceeds the defendant’s litigation costs associated with the sanctions proceeding. Thus, if the
penalty is set high enough, the defendant can be induced to request sanctions (provided at least some frivolous plaintiffs sue).

C. PLAINTIFFS' DECISIONS TO SUE

A frivolous plaintiff's decision to file a suit depends on whether the defendant will bring an action for sanctions against a losing plaintiff. If the defendant would bring an action, then if the plaintiff sues and loses he will anticipate having to pay a penalty and having to bear the additional litigation costs associated with the sanctions proceeding. Thus, a frivolous plaintiff would sue only if the expected value of the award in the initial trial (the probability that he will win multiplied by the award to a prevailing plaintiff) exceeds the sum of his litigation costs in the initial trial and his expected costs associated with the sanctions proceeding (the probability that the plaintiff will lose the initial trial multiplied by the sum of the penalty and his litigation costs in the sanctions proceeding).

If the defendant would not request sanctions, then a frivolous plaintiff will bring a suit if the expected value of the award in the initial trial exceeds his litigation costs in the initial trial.

Now consider a legitimate plaintiff's decision. That decision, too, depends on whether the defendant will bring an action for sanctions against a losing plaintiff. For a legitimate plaintiff, however, the only cost of an action against him is his litigation cost (no penalty will be imposed given our assumption that the sanctions inquiry correctly identifies the plaintiff's type).

If the defendant would request sanctions, then a legitimate plaintiff will sue initially if the expected value of the award in the initial trial exceeds the sum of his litigation costs in the initial trial and his expected litigation costs associated with the sanctions proceeding (the probability that the plaintiff will lose the initial trial multiplied by his litigation costs in the sanctions proceeding).

If the defendant would not request sanctions, then a legitimate plaintiff will sue initially if the expected value of the award in the initial trial exceeds his litigation costs in the initial trial.

We assume that if actions for sanctions are not brought, then both frivolous and legitimate plaintiffs will sue. If this were not the case, the analysis would be uninteresting: for then either just legitimate plaintiffs would sue or no one would sue, and there would be no need to consider sanctions to control frivolous litigation.

Note that there are two important differences between a frivolous plaintiff and a legitimate plaintiff in terms of the decision to sue: the probability of losing the original trial is higher for a frivolous plaintiff; and only a frivolous plaintiff faces the possibility of a penalty at the sanctions
stage (under our current assumptions). Both of these differences imply that the incentive to sue is lower for frivolous plaintiffs.

D. TYPES OF EQUILIBRIA

The preceding discussion shows that the defendant’s decision whether to bring an action for sanctions depends on the plaintiffs’ decisions whether to sue, and vice versa. An equilibrium occurs in the model when each party is doing the best it can for itself given its beliefs about what the others will do and when these beliefs are correct.23

One possible equilibrium occurs when the defendant does not request sanctions and both types of plaintiff file suit. (Recall that we assumed above that both types of plaintiff will sue if the defendant does not seek sanctions.) This equilibrium could occur, for example, if the penalty is relatively small or if the probability that a losing plaintiff is frivolous is relatively small, so that it is not worthwhile for the defendant to incur the litigation costs associated with a sanctions proceeding. We will refer to this as the “no-sanctions equilibrium.”

A second possible equilibrium can occur when the defendant does request sanctions, all legitimate plaintiffs file suit, and some, but not all, frivolous plaintiffs file suit. The reason some frivolous plaintiffs must file suit in this equilibrium is that, if none did, the defendant would expect that every losing plaintiff was a legitimate plaintiff and therefore would not find it worthwhile to seek sanctions (assuming, as we do for now, that there are no mistakes in the sanctions inquiry). We will refer to this equilibrium as the “sanctions equilibrium.”

In the sanctions equilibrium, frivolous plaintiffs must be indifferent between suing and not suing. For if they definitely preferred not to sue, then the defendant would not request sanctions. Conversely, if they definitely preferred to sue, then all would sue, rather than just some. (There is another possible equilibrium in which the defendant seeks sanctions and all frivolous plaintiffs sue. However, this equilibrium is not of interest from a public policy perspective because it is dominated by the no-sanctions equilibrium.24 As a consequence, we have defined the sanctions equilibrium to be one in which frivolous plaintiffs are indifferent between suing and not suing, and in which only a fraction of them sue.)25

---

23. To the extent that parties’ expectations about what others will do are not correct, the parties (or their lawyers) will learn from their experience and adjust their expectations accordingly.

24. The no-sanctions equilibrium is preferable because the same degree of deterrence of the defendant’s conduct can be achieved without incurring the litigation costs associated with actions for sanctions. For further discussion of this point, see infra Appendix.

25. The general results in this article do not depend on the “knife-edge” character of the sanctions equilibrium, in which frivolous plaintiffs are indifferent between suing and not suing. However, this equilibrium is the simplest one that allows us to develop the main points of this article. See infra notes 68, 70.
E. SOCIAL WELFARE AND PUBLIC POLICY

In general, social welfare includes the gain to the defendant from engaging in the conduct that gives rise to the suits, the harm caused by the defendant's conduct, and the litigation costs borne by all parties.\textsuperscript{26} A discussion of how the level of social welfare in the no-sanctions equilibrium compares to the level of social welfare in the sanctions equilibrium will be deferred to Part II.

For now, just note that either type of equilibrium can be achieved by the appropriate choice of the level of the penalty. If the penalty is set low enough, the defendant will not request sanctions, so that, given our earlier assumption, both types of plaintiff will sue. This results in the no-sanctions equilibrium. If this equilibrium is desirable, it is most natural to assume that it is obtained simply by prohibiting actions for sanctions.

Alternatively, the level of the penalty can be chosen so as to achieve the sanctions equilibrium. That there exists a level of the penalty which makes frivolous plaintiffs indifferent between suing and not suing is plausible since, if the penalty is very low, frivolous plaintiffs will prefer to sue, and if it is very high, they will prefer not to sue.\textsuperscript{27}

F. NUMERICAL EXAMPLE

To make the model introduced in this part more concrete, consider the following numerical example, which will be used throughout the article to illustrate our points.\textsuperscript{28} Suppose there are one hundred potential frivolous plaintiffs and one hundred potential legitimate plaintiffs. In the original trial, frivolous plaintiffs prevail with probability .3, legitimate plaintiffs prevail with probability .8, and each party (including the defendant) incurs litigation costs of $10,000. If sanctions are not allowed, we assume that an award of $100,000 to a prevailing plaintiff maximizes social welfare.

Note that in the absence of sanctions both types of plaintiffs will sue: the expected award at trial for a frivolous plaintiff is $30,000 \((= .3 \times \$100,000)\);

\textsuperscript{26} Although it is common in the literature on the economic analysis of law to include the defendant's gain as a component of social welfare, it is not essential to our analysis or our conclusions to do so. As the reader will see, this is because our emphasis will be on combinations of the award (to a winning plaintiff) and the sanction (on a losing frivolous plaintiff) that do not affect the defendant's expected costs from engaging in the conduct that gives rise to the suits, so that the defendant's behavior will be the same under the different combinations. See infra note 31 and accompanying text.

\textsuperscript{27} One might wonder why the penalty is not raised above the level that makes frivolous plaintiffs indifferent between suing and not suing, so that they prefer not to sue. Recall that this outcome is not an equilibrium in our model since, if only legitimate plaintiffs sued, the defendant would never win an action for sanctions and therefore would not be willing to bring them. (In a more general model—such as one in which legitimate plaintiffs might mistakenly be found to be frivolous—the condition that frivolous plaintiffs be indifferent between suing and not suing would not be necessary. See infra note 70.)

\textsuperscript{28} The assumptions embodied in the example have been chosen to illustrate the points of this article as simply as possible.
the expected award for a legitimate plaintiff is $80,000 ($ = 0.8 \times$100,000); and litigation costs are only $10,000 for each type of plaintiff. This situation illustrates the no-sanctions equilibrium since the defendant does not seek sanctions and both types of plaintiff file suit.

Now suppose that sanctions are available and that each party incurs an additional $5,000 in litigation costs in a sanctions proceeding. Assume further that the award to a prevailing plaintiff in the original trial is raised to $135,250 and that the penalty in the sanctions proceeding is $38,679 (the reason for choosing these numbers will be explained in Part II).

If actions for sanctions are brought, frivolous plaintiffs will be indifferent between suing and not suing: their expected award in the initial trial is $40,575 ($ = 0.3 \times$135,250); this equals the sum of their litigation costs in the initial trial, $10,000, and their expected costs associated with a sanctions proceeding, $30,575 ($ = 0.7 \times (38,679 + 5,000)$). Legitimate plaintiffs, however, will prefer to sue: their expected award in the initial trial is $108,200 ($ = 0.8 \times$135,250), which greatly exceeds the sum of their litigation costs in the initial trial, $10,000, and their expected litigation costs associated with a sanctions proceeding, $1,000 ($ = 0.2 \times 5,000$).

Finally, consider whether the defendant will seek sanctions. Suppose that forty percent of the potential frivolous plaintiffs sue. This implies that the probability that a losing plaintiff is frivolous is 0.583. Thus, the defendant will request sanctions because the expected value of the penalty that she will receive, $22,550 ($ = 0.583 \times 38,679$), greatly exceeds her sanctions-related litigation costs, $5,000.

The preceding situation illustrates the sanctions equilibrium since the defendant brings actions for sanctions, all legitimate plaintiffs file suit, and some, but not all, frivolous plaintiffs file suit.

II. SHOULD SANCTIONS BE ALLOWED, GIVEN THAT THEY INCREASE THE COMPLEXITY, AND THEREFORE THE COST, OF LITIGATION?

Within our framework, the answer to the question of whether sanctions should be permitted depends on whether the level of social welfare is higher under the no-sanctions equilibrium or under the sanctions equilibrium. A comparison of social welfare under these two equilibria is complicated by the fact that if the award in the initial trial is the same in both cases, the expected costs of the defendant—and therefore her incentive to engage in the conduct that gives rise to the suits—generally will be different. To simplify matters, we will assume in this part that the award in the initial trial is adjusted in the sanctions equilibrium so that the

29. Since the expected number of losing frivolous plaintiffs is 28 ($ = 0.7 \times 0.4 \times 100$) and the expected number of losing legitimate plaintiffs is 20 ($ = 0.2 \times 100$), the fraction of losing plaintiffs who are frivolous is 0.583 ($ = 28/(28 + 20)$).
defendant's expected costs are the same as under the no-sanctions equilib-
rium.\textsuperscript{30}

Given this assumption, the conduct of the defendant that gives rise to
the initial suits will be the same in the two equilibria; consequently, the
gain to the defendant from engaging in the conduct, as well as the harm
caused by the defendant, will be the same.\textsuperscript{31} Thus, the social welfare
comparison between the two equilibria depends solely on their respective
litigation costs.\textsuperscript{32}

In the no-sanctions equilibrium, both types of plaintiff sue, but by
definition there are no subsequent actions for sanctions. Litigation costs in
this equilibrium therefore consist solely of the plaintiffs' and the defendant's
litigation costs in the original trials.

In the sanctions equilibrium, all legitimate plaintiffs sue, only some
frivolous plaintiffs sue, and actions for sanctions are brought against losing
plaintiffs. Litigation costs therefore consist of the plaintiffs' and the
defendant's litigation costs in the trials resulting from the plaintiffs who
sue, plus the plaintiffs' and the defendant's litigation costs associated with
the sanctions proceedings brought against the subset of plaintiffs who lose.

This discussion suggests that either type of equilibrium could be prefer-
able. On one hand, prohibiting actions for sanctions obviously saves the
litigation costs that would arise at the sanctions stage. On the other hand,
allowing such actions deters some frivolous plaintiffs from suing in the first
place and therefore saves litigation costs associated with initial trials.
Depending on the circumstances, either equilibrium could lead to lower
litigation costs.

To explain the circumstances under which each equilibrium would be
preferred, we will assume initially that sanctions are not allowed and then
consider whether it would be desirable to allow them.

\textsuperscript{30} That this can be accomplished is demonstrated in the Appendix.

\textsuperscript{31} To understand intuitively why the defendant's behavior will be the same, consider the
example of a firm that is contemplating manufacturing a new item that will cause a certain
amount of pollution as an inevitable byproduct. Without sanctions, suppose that 100 legiti-
mate plaintiffs and 100 frivolous plaintiffs would sue the firm, resulting in expected liability
and defense costs equal to $1 million. With sanctions, suppose that 100 legitimate plaintiffs
and only 40 frivolous plaintiffs would sue the firm, but that the award to prevailing plaintiffs
is raised so that the expected liability and defense costs from these suits again equals $1
million. Because the defendant's costs are the same in the two scenarios, the defendant's
decision whether to manufacture the new item will be the same.

\textsuperscript{32} Although our discussion of whether sanctions are desirable focuses on their effect on
aggregate litigation costs, one could easily incorporate into our analysis a preference for
detering frivolous suits independently of their effect on litigation costs (and independently
of their effect on the defendant's behavior). Although the existence of such a preference
obviously would put more weight on the scale in favor of the sanctions equilibrium—and
thereby affect our discussion in this part—the remainder of the article would not be changed
because, after the present part, we assume that the sanctions equilibrium is preferable to the
no-sanctions equilibrium.
Not surprisingly, sanctions tend to be desirable the lower the parties' litigation costs associated with a sanctions proceeding, for then the deterrence of frivolous suits can be achieved at low cost. Also, sanctions tend to be beneficial the higher the parties' litigation costs in the original trial. This is because the deterrence of frivolous suits then avoids more litigation costs.

Sanctions also tend to enhance social welfare the higher the probability that a plaintiff (of either type) will prevail in the original trial. The explanation of this result is slightly more complicated. The higher the probability that a plaintiff will prevail in the original trial, the lower the probability that there will be an action for sanctions (since a request for sanctions occurs in our model only if a plaintiff loses). Thus, the beneficial effects of sanctions (in deterring some frivolous suits) can be achieved with lower expected litigation costs as the probability that a plaintiff will prevail in the original trial rises.

Sanctions also tend to be desirable the lower the number of legitimate plaintiffs. If there are fewer legitimate plaintiffs, then there will be fewer legitimate plaintiffs who will lose the initial trial and therefore fewer (wasteful) actions for sanctions against legitimate plaintiffs.

Finally, and perhaps most surprisingly, the desirability of sanctions depends in an ambiguous way on the number of potential frivolous plaintiffs. On one hand, the greater the number of potential frivolous plaintiffs, the greater the litigation costs saved by deterring frivolous suits initially. On the other hand, the greater the number of frivolous plaintiffs, the greater the number of actions for sanctions that will result and the greater the litigation costs associated with these actions. The latter effect could dominate if the costs of sanctions proceedings are relatively high, if the percentage of potential frivolous plaintiffs deterred is relatively low, and if the probability that a frivolous plaintiff will lose the original trial is relatively high (resulting in a larger number of sanctions proceedings).

Sanctions, thus, may or may not be socially desirable. A key determinant of whether they should be permitted is the cost of litigation in the original trial relative to the cost in the sanctions proceeding. If litigation costs are relatively high in the original trial, then the litigation-cost savings from deterring some frivolous suits through the use of sanctions will more than compensate for the increased litigation costs resulting from the sanctions proceedings.

It will be instructive to illustrate some of the points in the present part with the numerical example from Part I. First observe from the figures given that the expected costs borne by the defendant in the no-sanctions equilibrium are $13 million: $4 million associated with suits by frivolous plaintiffs \((= 100 \times [(.3 \times $100,000) + $10,000])\), plus $9 million associated with suits by legitimate plaintiffs \((= 100 \times [(.8 \times $100,000) + $10,000])\). In the sanctions equilibrium, if the award is $135,250 and the penalty is
$38,679, the expected costs borne by the defendant also will be $13 million:
$1.08 million associated with suits by frivolous plaintiffs ($= [.4 \times 100] \times ((.3 \times $135,250) - (.7 \times [$38,679 - $5,000]) + $10,000)), plus $11.92 million associated with suits by legitimate plaintiffs ($= 100 \times [(.8 \times $135,250) + (.2 \times $5,000) + $10,000])

If the defendant’s costs are the same in the two equilibria, the social welfare comparison between them depends solely on their respective litigation costs. In the example, litigation costs in the no-sanctions equilibrium are $4 million ($= [100 + 100] \times [$10,000 + $10,000]), while litigation costs in the sanctions equilibrium are $3.28 million. Thus, in this example, sanctions are socially desirable—the savings in litigation costs in the initial trials as a result of deterring sixty percent of the potential frivolous plaintiffs more than offsets the additional litigation costs associated with sanctions.

The example easily could be modified, however, to show that sanctions are not desirable under different assumptions. For instance, if the litigation cost per party in a sanctions proceeding were much higher than $5,000, then the no-sanctions equilibrium would be preferable.

In order to focus on the optimal design of a system of sanctions, we will assume for the remainder of this article that actions for sanctions are socially desirable.

III. SHOULD THE AMOUNT OF THE PENALTY BE BASED ON THE GOAL OF COMPENSATING THE NONFRIVOLOUS PARTY FOR LITIGATION EXPENSES OR ON THE GOAL OF DETERRING POTENTIAL FRIVOLOUS LITIGANTS?

Upon first consideration, it might seem that the goals of compensation and deterrence are not in substantial conflict since making the frivolous party pay for the nonfrivolous party’s litigation costs will both compensate the nonfrivolous party and deter, at least to some extent, the frivolous party. The level of the penalty that achieves proper deterrence, however, is only tangentially related to the nonfrivolous party’s litigation expenses. Therefore, designing a system of sanctions requires making a choice between these two goals.

33. Litigation costs in the sanctions equilibrium are:

\[
((.4 \times 100) \times [$10,000 + $10,000 + (.7 \times [$5,000 + $5,000]]) + [100 \times [$10,000 + $10,000 + (.2 \times [$5,000 + $5,000])] = $3,280,000.
\]

34. As this discussion shows, some of our specific conclusions may differ from one category of suit to another, depending on the category-specific values of the cost of litigation and other variables. To the extent that information about the values of these variables cannot be obtained, it may be necessary to adopt a uniform policy across categories (such as a uniform policy of permitting actions for sanctions or a uniform policy of prohibiting them).
To understand the economic rationale for setting the penalty equal to the nonfrivolous party's litigation costs, it is necessary to expand the framework from Part I to include the possibility that one or both of the parties is risk averse rather than risk neutral (as was assumed initially).\(^3\) Risk-averse persons suffer disutility from the bearing of risk to the extent that they are not insulated from risk by insurance or other means. Since social welfare depends on the utility of individuals, their bearing of risk lowers social welfare.

Risk aversion provides a possible economic justification for setting the penalty equal to the nonfrivolous party's litigation costs. Suppose that the frivolous party is a plaintiff and the nonfrivolous party is the defendant, as we have been assuming, and that the defendant is risk averse. It is desirable, therefore, that the defendant not bear any unnecessary risk. However, the defendant faces two risks with respect to potential frivolous plaintiffs: that she will be sued and lose notwithstanding the fact that the plaintiff is frivolous; or that she will be sued and win but still bear her litigation costs. The latter risk can be eliminated by making a losing frivolous plaintiff pay a penalty equal to the defendant's litigation costs.\(^3\)

Note that this reasoning implies that the defendant's litigation costs in both the original trial and the action for sanctions should be included in the penalty. Otherwise, the defendant still would bear some risk even if she prevails in the original suit and in the sanctions proceeding.\(^3\)

To see this in the numerical example, note that if a frivolous plaintiff sues the defendant and loses, the defendant still suffers a loss of $10,000—her litigation costs in the original trial. A penalty equal to just the defendant's cost of bringing an action for sanctions, $5,000, would not reduce this loss at all (it would simply compensate the defendant for the cost of bringing the action). However, a penalty equal to the sum of the defendant's litigation costs in the original trial and the sanctions proceeding, $15,000, would eliminate it entirely. (For a reason explained below, the fact that the penalty should include both sets of litigation costs contrib-

---

35. Risk aversion means, in essence, that a person cares not only about the expected value of a risky situation, but also about the absolute magnitude of the risk. For example, a risk-averse person would not be indifferent between losing $1,000 with certainty and a 10% chance of losing $10,000, even though the expected loss is $1,000 in the latter case. She would prefer the $1,000 loss with certainty or, equivalently, to buy an insurance policy for $1,000 that would compensate her fully in the event that she lost $10,000.

36. For reasons explained in the Appendix, the ability of sanctions to reduce the bearing of risk by the defendant may be more limited than is suggested here.

37. Rule 11 sanctions presumably include all of the defendant's litigation costs when the plaintiff's original complaint is the basis for the Rule 11 motion. Rule 11 states that "the sanction may consist of, or include, . . . an order directing payment to the movant of some or all of the reasonable attorneys' fees and other expenses incurred as a direct result of the violation." Fed. R. Civ. P. 11 (emphasis added).
utes to the conflict between the compensatory rationale for penalties and the deterrence rationale.)

Although there is nothing inappropriate per se about the compensatory rationale for penalties, another observation suggests that it may not be very important. If the risk aversion of defendants were of great concern, one would expect a fee allocation rule to have been adopted that would always require a losing plaintiff to compensate the defendant for her litigation costs—that is, regardless of whether the defendant requests sanctions. In fact, such a rule is not the norm in the United States. It would be peculiar, therefore, to rely heavily on risk-bearing considerations to justify the use of sanctions or to justify setting the level of the penalty equal to the nonfrivolous party's litigation costs.

For this reason, we give greater weight to the deterrence rationale for imposing sanctions. We will now show that a deterrence-based penalty must exceed the cost to the nonfrivolous party of bringing an action for sanctions and that such a penalty is related only tangentially to the nonfrivolous party's litigation costs.

To deter frivolous suits in the model described in Part I, the defendant must have an incentive to seek sanctions if she wins the original trial. (Otherwise, given our assumptions, frivolous plaintiffs would sue, knowing that they would not face penalties.) Recall that the defendant will request sanctions if the expected value of the penalty that she will receive—the probability that the defendant will prevail in the sanctions proceeding multiplied by the amount of the penalty—exceeds the defendant's litigation costs associated with the proceeding.

If all losing plaintiffs were frivolous plaintiffs, the defendant would be certain to win the action for sanctions (assuming, as we do for now, that there are no mistakes). Then the penalty would only have to equal the defendant's litigation cost in the action (plus a little more) to induce her to bring such an action.

---

38. See, e.g., Steven Shavell, Suit, Settlement, and Trial: A Theoretical Analysis Under Alternative Methods for the Allocation of Legal Costs, 11 J. LEGAL STUD. 55-56 n.2 (1982) (noting that the system in which both sides bear their own litigation costs "is the prevailing method for allocating legal costs in the United States, although exceptions to its use are made in certain jurisdictions for certain categories of cases"). The only example Shavell provides of a rule favoring defendants is in medical malpractice cases in Florida. Id.

39. This conclusion is premised on the assumption that the compensatory rationale for sanctions is based on the desirability of reducing risk. However, a compensatory rationale for sanctions could be derived from other considerations. For example, if one believes on fairness grounds that nonfrivolous parties have a right to be free of the costs imposed by frivolous parties, then setting penalties equal to the nonfrivolous party's litigation costs would be desirable even if such a policy increased the risk borne by the nonfrivolous party. To the extent that fairness or other considerations not based on risk aversion provide a compensatory rationale for sanctions, the conclusion in the text might have to be modified.
Some losing plaintiffs, however, will be legitimate plaintiffs; therefore, the probability that the defendant will prevail in the action for sanctions is less than one. To make up for the chance that the defendant will lose, the penalty must exceed the defendant's cost of bringing the action. For example, if half of all losing plaintiffs are frivolous, then the penalty has to be at least twice the defendant's sanctions-related litigation costs; if one-quarter of the losing plaintiffs are frivolous, the penalty has to be at least four times the defendant's sanctions-related litigation costs, and so forth.  

This discussion provides two reasons why the level of the penalty that deters frivolous litigation bears little relationship to the level that compensates the nonfrivolous party for her litigation costs. First, a deterrence-based penalty does not depend on the nonfrivolous party's litigation costs in the original trial. Those costs are "sunk costs" at the time the nonfrivolous party is deciding whether to request sanctions. A compensation-based penalty, however, would include those costs. Second, a deterrence-based penalty generally must exceed the nonfrivolous party's cost of seeking sanctions. A compensation-based penalty, in contrast, would include (without adjustment) the nonfrivolous party's cost of seeking sanctions.

The numerical example can be used to illustrate the differences between a compensation-based and a deterrence-based penalty. As noted above, a compensation-based penalty would be $15,000, the sum of the defendant's litigation costs in the original trial and the action for sanctions. A deterrence-based penalty would have to be a multiple of the $5,000 cost to the defendant of bringing an action. Since the probability that a losing plaintiff is frivolous is .583 (see the discussion in Part I.F.), this is also the probability that the defendant will prevail in a sanctions proceeding. Hence, the defendant's sanctions-related litigation costs have to be multiplied by at least $1.72 ( = 1/.583) to make it worthwhile for the defendant to bring an action, which means that the penalty must be at least $8,600 ( = 1.72 \times 5,000).  

For the reasons discussed in this part, we assume for the remainder of this article that penalties are designed to achieve proper deterrence of

---

40. Under Rule 11, the payment to the party bringing a motion for sanctions normally equals the litigation costs incurred by that party. See supra note 2. Under the amendments to Rule 11 proposed by the Advisory Committee on Civil Rights and recently adopted by the Supreme Court, the party bringing a Rule 11 motion would receive no more than her litigation costs. See Report on the Proposed Amendments, supra note 8, at 47-48, 54. Thus, the need to multiply litigation costs to make up for the chance of not prevailing on the Rule 11 motion apparently has been overlooked. (Because only the litigation costs associated with the Rule 11 motion need to be multiplied, awarding a party her entire litigation costs might provide a sufficient inducement to bring the motion, but this would be fortuitous.)

41. In fact, the deterrence-based penalty in the numerical example is $38,679. It needs to be this large because any lower penalty would lead all potential frivolous plaintiffs to sue.
frivolous litigation, rather than to compensate nonfrivolous parties for their litigation costs.

IV. SHOULD SOME OR ALL OF THE PENALTY BE GIVEN TO THE COURT OR TO THE GENERAL TREASURY RATHER THAN TO THE OPPOSING PARTY?

Now suppose that the sanction imposed on a frivolous plaintiff can differ from what the defendant receives. Although the plaintiff’s payment could be greater than or less than the defendant’s award, we will concentrate on the former case (for reasons that will become clear below). We assume, moreover, that the excess of what the plaintiff pays over what the defendant receives will go to the court or to the general treasury.

Everything else equal, if the defendant does not receive the full amount of the penalty paid by a frivolous plaintiff, the defendant’s incentive to seek sanctions will be lessened. The payment to the defendant cannot be lowered too much, however, or the defendant will not be willing to bring such actions. As discussed previously, the payment cannot fall below an appropriate multiple of the defendant’s sanctions-related litigation costs.

To understand why it might be desirable to give some of a frivolous plaintiff’s penalty to the court or to the general treasury rather than to the defendant, it will be useful to add to the framework of Part I the possibility that potential defendants differ in terms of their costs of bringing actions for sanctions. Such cost variations might occur, for example, because defendants value their time differently or because they do not “shop around” for legal counsel.

We will assume that the court awards the same amount to every defendant who prevails in an action for sanctions. Given this common payment, it is clear that defendants for whom the cost of bringing an action is relatively high will not bring such actions and that defendants for whom the cost is relatively low will bring them. Moreover, if the payment is lowered, fewer defendants will be willing to request sanctions. In terms of social welfare, a reduction in the number of sanctions proceedings is desirable, everything else equal, because the cost of litigation is lowered.

If fewer actions for sanctions are brought, however, frivolous plaintiffs might not be adequately deterred from suing. This potential problem can be offset by raising the penalty that a frivolous plaintiff has to pay after losing an action. For example, suppose that a potential frivolous plaintiff believes that there is only a fifty percent chance that if he sues and loses,
the defendant will seek sanctions against him. If the penalty paid by a frivolous plaintiff is doubled, the expected penalty will not have changed. The penalty actually needs to be raised by more than this to offset the plaintiff's expected savings in litigation costs as a result of a lower chance of an action being brought against him. An appropriate increase in the penalty will leave frivolous plaintiffs' behavior unaffected despite the reduced probability of sanctions.

The discussion in the preceding several paragraphs implies that the sanction should be "decoupled" by paying a defendant who prevails in an action for sanctions less than the amount that has to be paid by the frivolous plaintiff who loses the action. Paying defendants less reduces the number of actions; making frivolous plaintiffs pay more maintains the deterrent effect of such actions. Relative to the outcome when the amount received by the defendant is equated to the amount paid by a frivolous plaintiff, this decoupled system lowers litigation costs without sacrificing the deterrent effect of sanctions.

The preceding discussion also suggests why it is not desirable to decouple the sanction by awarding the defendant more than a frivolous plaintiff pays: doing so would encourage more requests for sanctions, thereby increasing litigation costs.

When sanctions are decoupled optimally, there will be a surplus equal to the difference between what a frivolous plaintiff pays and what the defendant receives. Provided that this surplus is not returned to the parties, it is immaterial to whom it is given. As noted, we assume that the surplus goes to the court or to the general treasury.

To see in the numerical example how decoupling the penalty can improve social welfare, suppose that the cost of bringing an action for sanctions varies among defendants, with half of them having costs of $2,500 and half having costs of $7,500. Assuming that the probability that a

43. This belief would be reasonable, for instance, if there are an equal number of high-litigation-cost and low-litigation-cost defendants of whom only the latter will seek sanctions, and the plaintiff does not know which type he has sued.

44. See Katz, supra note 15, at 19-20 (showing that a sufficiently large penalty paid by frivolous plaintiffs to the government could deter frivolous suits). Under this form of decoupling, the government rather than the opposing party must take the initiative in enforcing the sanction.

45. There are some circumstances, however, in which the defendant should be awarded more than what a frivolous plaintiff pays. For example, if the wealth of frivolous plaintiffs is quite limited, they will not be able to pay very much. Hence, to deter them, the probability of an action for sanctions being brought must be relatively high, which in turn might require awarding the defendant an amount greater than what frivolous plaintiffs pay. Our discussion in the text implicitly assumes that these circumstances do not apply. See generally A. Mitchell Polinsky & Yeon-Koo Che, Decoupling Liability: Optimal Incentives for Care and Litigation, 22 RAND J. ECON. 562 (1991).

46. Under both the amendments to Rule 11 adopted by the Supreme Court and the Bench-Bar Proposal to Revise Civil Procedure Rule 11, the surplus (in the latter case, the entire penalty) would be paid to the court. See supra note 11.
defendant will prevail in an action is still .583, the amount of the penalty given to defendants must be just over $4,288 (= [1/.583] x $2,500) to induce low-litigation-cost defendants to seek sanctions (their expected award then will just exceed their $2,500 cost of bringing the action). Since high-litigation-cost defendants clearly will not request sanctions if the award is this low, only one-half of all potential actions will be brought. To continue to appropriately deter frivolous plaintiffs, the penalty must be more than doubled (to offset the plaintiff’s expected savings in litigation costs) from $38,679 to $82,358. Given this penalty, the behavior of potential frivolous plaintiffs will not be affected. But decoupling the penalty in this way is desirable because the litigation costs associated with sanctions fall from $480,000 to $180,000, a savings of $300,000. In addition, a surplus of $1.87 million (= [(.4 x 100 x .7) + (100 x .2)] x .5 x ($82,358 - $4,288)) will accrue to the court or to the general treasury.

V. SINCE SANCTIONS MIGHT BE IMPOSED BY MISTAKE, WILL THEY DISCOURAGE SOME LEGITIMATE PLAINTIFFS FROM SUING?

Before discussing whether legitimate plaintiffs will be discouraged from suing because of the risk that sanctions will be imposed on them by mistake, it will be useful to consider how such plaintiffs are affected by actions for sanctions even when there are no mistakes.

The framework in Part I presumed, for simplicity, that the sanctions proceeding correctly identified the plaintiff’s type. Yet, even within that framework, legitimate plaintiffs are adversely affected by the prospect of sanctions. Although they will not be found liable after a sanctions proceeding, they will incur litigation costs as a result of it. Everything else equal, the prospect of these costs reduces their incentive to bring a suit initially.

47. If actions for sanctions are brought with certainty and the penalty is $38,679, a frivolous plaintiff’s costs as a result of an action, including his litigation costs, are $43,679 (= $38,679 + $5,000). If actions are brought with a probability of .5 and the penalty is $82,358, a frivolous plaintiff’s expected costs as a result of an action also are $43,679 (= .5 x ($82,358 + $5,000)).

48. When the penalty is not decoupled and both types of defendant seek sanctions, litigation costs are

\[
\left[(.4 \times 100 \times .7) + (100 \times .2)\right] \times .5 \times (\$2,500 + \$5,000) + \left[(.4 \times 100 \times .7) + (100 \times .2)\right] \times .5 \times (\$7,500 + \$5,000) = \$480,000.
\]

The first term in braces represents litigation costs associated with actions for sanctions by low-litigation-cost defendants, and the second term represents the corresponding litigation costs associated with high-litigation-cost defendants. Litigation costs when the penalty is decoupled and only low-litigation-cost defendants request sanctions are

\[
[(.4 \times 100 \times .7) + (100 \times .2)] \times .5 \times (\$2,500 + \$5,000) = \$180,000.
\]

49. Within the model, the probability that legitimate plaintiffs bear these costs equals the probability that they lose the initial trial.
Now suppose, more realistically, that mistakes can occur in the sanctions process, and that legitimate plaintiffs sometimes have to pay sanctions.\(^5\) This possibility reinforces the conclusion that the existence of sanctions, even without mistakes, reduces the incentive of legitimate plaintiffs to bring suits.

However, if the cost of defending against an action for sanctions and the risk of mistaken imposition of a penalty would cause legitimate plaintiffs to forgo suing, their incentive to sue can be restored by raising the award to prevailing plaintiffs in the original trial. In other words, no matter how adverse the effects of sanctions, legitimate plaintiffs still will sue if the award in the original trial is high enough.

It might seem that raising the initial award also will induce more frivolous plaintiffs to bring suits, thereby subverting the purpose of a system of sanctions. It is true that raising the award will increase the incentive of both legitimate and frivolous plaintiffs to sue. However, because legitimate plaintiffs have a higher probability of prevailing in the original trial than do frivolous plaintiffs, raising the award will benefit them more. As a result—and because raising the penalty adversely affects frivolous plaintiffs more than legitimate plaintiffs—it is possible to raise the award in the initial trial and the penalty so that, despite the possibility that legitimate plaintiffs will be mistakenly held liable in sanctions proceedings, the incentive of legitimate plaintiffs to sue can be maintained without increasing the incentive of frivolous plaintiffs to sue. Everything else equal, the greater the likelihood that legitimate plaintiffs will be mistakenly found liable in sanctions proceedings, the higher must be the optimal award to prevailing plaintiffs in the original trial and the optimal penalty.

The numerical example can be used to illustrate the effects of mistakes on the optimal award and penalty. Recall that, when the possibility of mistakes was ignored, the optimal award was $135,250 and the optimal penalty was $38,679. Now suppose that 25% of the legitimate plaintiffs who lose the initial trial are incorrectly determined to be frivolous plaintiffs in sanctions proceedings. Raising the award in the original trial by $2,484—to $137,734—and raising the penalty by $1,064—to $39,743—will restore the incentive of potential legitimate plaintiffs to sue while still appropriately deterring potential frivolous plaintiffs.\(^1\)

50. The chance that frivolous plaintiffs may escape liability because of mistakes in the sanctions process will be discussed briefly below.

51. It was shown in Part I.F. that if the award is $135,250 and the penalty is $38,679 and if there are no mistakes, legitimate plaintiffs will sue because their expected award in the initial trial ($108,200) exceeds the sum of their litigation costs in the initial trial ($10,000) and their expected litigation costs associated with the sanctions proceeding ($1,000). The resulting value of a suit to a legitimate plaintiff is $97,200 ($108,200 − $10,000 − $1,000). If 25% of the legitimate plaintiffs who lose the initial trial are incorrectly determined to be frivolous plaintiffs in sanctions proceedings, their incentive to sue—as measured by the value of a suit—falls by $1,934 (0.2 × 0.25 × $38,679). (Although in this example legitimate
Even though the adverse effects of mistakes on legitimate plaintiffs' incentives to sue can be offset, accuracy in the sanctions process still is a desirable goal if the parties are risk averse.\(^5\) As noted above, to offset the effects of mistakes requires raising the award in the initial trial as well as the penalty in the sanctions proceeding. Both of these changes increase the range of possible outcomes of a suit. They therefore impose greater risk on the parties, which lowers social welfare if the parties are risk averse.\(^5\) Thus, even though it is possible to maintain legitimate plaintiffs' incentives to sue, accuracy in the sanctions process is desirable because the award and the penalty will have to be increased by lesser amounts, thereby reducing the imposition of risk.\(^5\)

This part has focused on the mistake of imposing a sanction on a legitimate plaintiff. Another type of mistake that can occur in an action for sanctions is not imposing a penalty on a frivolous plaintiff. Obviously, this mistake will increase a frivolous plaintiff's incentive to sue. For reasons analogous to those discussed above, the effects of this mistake can be offset by increasing the penalty (which will adversely affect frivolous plaintiffs more than legitimate plaintiffs) as well as the award in the original trial.

VI. IF SANCTIONS ARE NOT ALLOWED, ARE THERE ALTERNATIVE POLICIES THAT CAN DISCOURAGE FRIVOLOUS LITIGATION?

We will discuss two alternatives to sanctions that can discourage frivolous suits: a modified version of the English rule for allocating legal costs and the “decoupling” of liability in the initial trial.
A. MODIFYING THE ENGLISH RULE

Under the English rule, the loser of the initial suit pays the winner’s legal costs. Compared to the American rule, the English rule increases the award to a winning plaintiff—by the amount of the plaintiff’s litigation costs—and imposes a penalty on a losing plaintiff—equal to the defendant’s litigation costs. Because legitimate plaintiffs have a higher probability of winning than frivolous plaintiffs, raising the award benefits them more. And because frivolous plaintiffs have a higher probability of losing than legitimate plaintiffs, raising the penalty hurts them more. Thus, adopting the English rule would lower the value of suits for frivolous plaintiffs relative to legitimate plaintiffs.55

However, the conventional version of the English rule may not discourage frivolous plaintiffs enough to keep them from suing. In other words, there is no reason to believe that imposing a penalty on a losing plaintiff equal to the defendant’s litigation costs will deter a plaintiff who otherwise would have brought a suit. For example, if the defendant’s litigation costs—and therefore the penalty—are low enough, such a plaintiff will not be discouraged from suing. Thus, while the English rule disfavors frivolous plaintiffs relative to legitimate plaintiffs, it might not disfavor them enough to deter their suits.56

This potential problem with the English rule can be overcome by dropping the requirement that the penalty on a losing plaintiff equal the defendant’s litigation costs. If the penalty is made large enough, frivolous plaintiffs can be discouraged from suing. However, since legitimate plaintiffs will face the higher penalty too—even with a lower probability—it might be necessary to raise the award to a winning plaintiff in the original trial in order to provide enough of an inducement to legitimate plaintiffs to continue bringing suits. (This point is analogous to the possible need to raise the award in the initial trial in order to offset the risk of mistakenly imposing sanctions on legitimate plaintiffs.)

Elsewhere, we have shown that it is feasible to choose a penalty to impose on losing plaintiffs and an award to give to winning plaintiffs that, together, can forestall frivolous suits without discouraging legitimate suits (and that also can appropriately discourage the defendant’s conduct).57

55. However, if the probability of prevailing is sufficiently high for frivolous plaintiffs (but still below that of legitimate plaintiffs), switching to the English rule could increase frivolous plaintiffs’ incentive to sue. Conversely, if the probability of prevailing is sufficiently low for legitimate plaintiffs (but still above that of frivolous plaintiffs), switching to the English rule could lower their incentive to sue.

56. For an even more skeptical view of the ability of the English rule to deter frivolous suits, see Katz, supra note 15, at 17-19 (“[T]he English rule provides no remedy to the problem of strike suits.”).

57. See Polinsky & Rubinfeld, supra note 15. To be more precise, what we show is that when the probability of prevailing varies among plaintiffs, it is possible to choose a penalty and an award that can deter suits by plaintiffs whose probability of success is below some
For the reasons suggested in the previous paragraph, however, the required penalty might exceed the defendant's legal costs; and for similar reasons, the required increase in the award might differ from the plaintiff's litigation costs. We will, accordingly, refer to this system as a "modified" English rule and assume that the award and penalty are chosen so as to deter frivolous suits but not legitimate suits.

In the numerical example, adopting the conventional English rule would be equivalent to raising the award to a winning plaintiff by $10,000 (the plaintiff's litigation costs in the original trial) and imposing a penalty on a losing plaintiff of $10,000 (the defendant's litigation costs in the original trial). Starting from the benchmark of the no-sanctions equilibrium (an award of $100,000 and no penalty on a losing plaintiff), this results in an award of $110,000 and a penalty of $10,000. This award and penalty combination will not deter potential frivolous plaintiffs from suing since their expected award net of their expected penalty, $26,000 (\(= .3 \times 110,000\) - \(.7 \times 10,000\)), far exceeds their litigation costs of $10,000. Consider, however, a modified English rule consisting of an award of $164,000 and a penalty of $56,000.\(^{58}\) With this award and penalty, frivolous plaintiffs will not sue because their expected award net of their expected penalty now is $10,000 (\(= .3 \times 164,000\) - \(.7 \times 56,000\)). Moreover, this award and penalty will cause the defendant's expected costs to be the same as that under the sanctions approach when the award is $135,250 and the penalty is $38,679.\(^{59}\)

B. DECOUPLING LIABILITY IN THE ORIGINAL TRIAL

Another alternative to sanctions is to decouple liability in the original trial.\(^{60}\) We already have discussed, in Part IV, the desirability of decoupling the penalty in the sanctions proceeding. Similar reasoning can show that, by decoupling the award in the original trial, frivolous plaintiffs can be discouraged from suing without deterring suits by legitimate plaintiffs.

To see this within the framework of Part I, suppose again that the American rule governs, so that each party bears his or her own litigation costs and that there are no penalties on losing plaintiffs.\(^{61}\) Given the cost

\[\text{threshold without deterring suits by plaintiffs whose probability is above the threshold (and that also can appropriately discourage the defendant's conduct). A highly abbreviated version of this analysis is contained in the Appendix in the special case in which the plaintiffs' probability of prevailing assumes only two values.}\]

\(^{58}\) See infra note 76 for an explanation of how these numbers are derived.

\(^{59}\) Under the modified English rule, the defendant's expected costs are

\[100 \times [(.8 \times 164,000) - (.2 \times 56,000) + 10,000] = 13,000,000,\]

the same as under the sanctions approach (see supra Part II).

\(^{60}\) See generally Polinsky & Che, supra note 45.

\(^{61}\) If penalties were allowed in this discussion, it would not be necessary to consider
of bringing a suit and the respective probabilities of prevailing for frivolous and legitimate plaintiffs, there clearly exists some award to a winning plaintiff that is low enough to deter suits by frivolous plaintiffs but not so low as to deter suits by legitimate plaintiffs. (Such an award exists because frivolous plaintiffs have a lower probability of winning.) We will assume hereafter that the award is set at this level so that only legitimate plaintiffs sue. 62

However, because the award might have to be relatively low—to discourage potential frivolous plaintiffs from suing—the expected costs imposed on the defendant might be too low if all she has to pay if she loses the initial suit is the award to the prevailing plaintiff. In other words, the defendant might not be adequately discouraged from engaging in the conduct that gave rise to the suits. This problem, if it exists, can be easily remedied by making the defendant pay an amount greater than the award to the plaintiff, with the excess given to the court or to the general treasury. (Similarly, if the expected costs imposed on the defendant would be too high if she has to pay the award to the plaintiff, the amount that the defendant has to pay can be lowered.) Thus, if liability in the original trial is decoupled, frivolous plaintiffs can be discouraged from suing, legitimate plaintiffs can be encouraged to sue, and expected costs can be imposed on the defendant so that her conduct will be appropriate.

In the numerical example, this result can be achieved by awarding $33,333 to a winning plaintiff in the original trial, but making the losing defendant pay $150,000. Frivolous plaintiffs then will not sue because their expected award is $10,000 (= .3 x $33,333), which equals their cost of litigation. Also, this system of decoupled liability will cause the defendant's expected costs to be the same as that under the sanctions approach when the award is $135,250 and the penalty is $38,679. 63

The discussion in this part has shown that either a properly modified English rule or an appropriately decoupled system of liability in the initial trial can achieve the same outcome as the sanctions approach. Moreover, both of these alternatives have an advantage over actions for sanctions in that they do not require a costly inquiry to determine whether the plaintiff was frivolous. For these reasons, they deserve serious consideration. 64

62. There is a range of values of the award that will achieve the desired goal.
63. Under this system of decoupled liability the defendant's expected costs are

\[ 100 \times [(0.8 \times $150,000) + $10,000] = $13,000,000, \]

the same as under the sanctions approach (see supra Part II).
64. One qualification to the discussion in this part should be mentioned. We have assumed throughout this article that all legitimate suits have the same probability of prevailing and that all frivolous suits have the same (but lower) probability of prevailing. If,
SANCTIONING FRIVOLOUS SUITS

CONCLUSION

In this article we have used an economic model of litigation to discuss the appropriateness of imposing sanctions on plaintiffs found to have filed frivolous suits. We also have described certain desirable properties of a system of sanctions. Our key conclusions are that:

—Sanctions should be used to promote the goal of deterring frivolous litigation, rather than to promote the goal of compensating nonfrivolous parties for their litigation costs.

—The payment to the nonfrivolous party should be a multiple of that party's cost of bringing an action against the frivolous party (when the multiple depends on the nonfrivolous party's probability of prevailing in the action).

—The penalty paid by the frivolous party generally should be higher, with the excess paid to the court or the general treasury. (In other words, sanctions should be decoupled.)

—The award in the initial trial often should be raised to offset the adverse effects on legitimate plaintiffs of mistakes in the sanctioning process.

The system of Rule 11 sanctions has, until recently, differed significantly from the recommended system in every dimension: the Rule 11 award to the nonfrivolous party usually has been limited to that party's litigation costs; the Rule 11 penalty paid by the frivolous party almost always has equalled the amount paid to the nonfrivolous party; and the award in the original trial never (to our knowledge) has been changed to offset the effects of mistakes in the Rule 11 process.65 Only with the recent amendment of Rule 11 has decoupling become the norm.66

Even if the Rule 11 system were fully reformed in the ways that we have suggested, the use of Rule 11 sanctions still would require costly inquiries

more realistically, there are variations in the probability of prevailing within each category of suits, the argument for using the sanctions approach improves. This is because the sanctions approach attempts to determine whether a particular suit is frivolous or not; hence, it will be more likely than the alternatives discussed in this part to deter frivolous suits that have a relatively high probability of prevailing and to encourage legitimate suits that have a relatively low probability of prevailing. To the extent that it is desirable to deter frivolous suits and encourage legitimate suits independently of the consequences for the injurer's gain, the victims' damages, and the parties' litigation costs, the sanctions approach is more attractive than the alternatives discussed in this part.

65. See supra notes 3, 5.

66. Both the amendments to Rule 11 proposed by the Advisory Committee on Civil Rules and adopted by the Supreme Court, and those proposed by the Bench-Bar Proposal to Revise Civil Procedure Rule 11, make decoupling the norm and also emphasize the deterrence purposes of Rule 11. See supra notes 10-11. Neither of these proposals, however, explicitly recognizes the need to enhance the award to the nonfrivolous party to make up for the chance that that party will not prevail on the Rule 11 motion. See supra note 40 and accompanying text. Nor does either explicitly recognize the desirability of raising the award in the initial trial to offset the effects of mistakes in the Rule 11 process. See supra Part V.
to determine whether a litigant presented in court a frivolous "pleading, written motion, or other paper." For the reasons discussed in this article, the additional litigation costs associated with Rule 11 inquiries might more than offset the beneficial effects of Rule 11 sanctions in terms of deterring frivolous litigation. Moreover, as we have shown, there may be less costly ways to deter frivolous litigation.

APPENDIX: A FORMAL MODEL OF SANCTIONS

Because the outline and basic assumptions of the model were presented in Part I, the first part of the Appendix will be relatively brief. The following notation will be used:

\[ N_f = \text{number of potential frivolous plaintiffs}; \]
\[ N_e = \text{number of potential legitimate plaintiffs}; \]
\[ q = \text{probability that a frivolous plaintiff will prevail in the initial trial}; \]
\[ p = \text{probability that a legitimate plaintiff will prevail in the initial trial (} p > q); \]
\[ c_p = \text{a plaintiff's litigation cost in the initial trial}; \]
\[ c_{p'} = \text{a plaintiff's litigation cost in an action for sanctions}; \]
\[ c_d = \text{the defendant's litigation cost in the initial trial (per suit)}; \]
\[ c_{d'} = \text{the defendant's litigation cost in an action for sanctions (per action)}; \]
\[ x = \text{award paid to a plaintiff from the defendant if the plaintiff prevails in the initial trial (} x \geq 0); \]
\[ y = \text{penalty paid by a plaintiff to the defendant if, after an action for sanctions, the plaintiff is found to be frivolous (} y \geq 0); \]
\[ \alpha = \text{fraction of frivolous plaintiffs who sue; and} \]
\[ \phi(\alpha) = \text{probability that a losing plaintiff is frivolous (given } \alpha). \]

DEFENDANT'S DECISION TO BRING AN ACTION FOR SANCTIONS

If the defendant wins the initial trial, she will bring an action if

\[ \phi(\alpha)y > c_{d'}, \]  

\[ \phi(\alpha) = \frac{\alpha N_f(1 - q)}{[\alpha N_f(1 - q) + N_e(1 - p)]}. \]

67. Throughout the numerical example, \( N_f = 100, N_e = 100, q = .3, p = .8, c_p = $10,000, c_{p'} = $5,000, c_d = $10,000, \) and \( c_{d'} = $5,000. \) In the no-sanctions equilibrium, \( x = $100,000, y = $0, \alpha = 1, \) and \( \phi(1) = .778. \) In the sanctions equilibrium, \( x = $135,250, y = $38,679, \alpha = .4, \) and \( \phi(.4) = .583. \) Because most of the details of the numerical example already have been provided in the text or in footnotes accompanying the text, we will not further elaborate on the numerical example in the Appendix.
Let \( \hat{\phi} \) be the value of \( \phi \) below which the defendant will not seek sanctions and above which she will:

\[
\hat{\phi} = \frac{c_d}{y}.
\]  

(3)

**PLAINTIFFS’ DECISIONS TO SUE**

Assuming actions for sanctions will be brought, a frivolous plaintiff will file a suit if

\[
qx - (1 - q)y > c_p + (1 - q)c_p'.
\]  

(4)

Let \( \hat{y} \) be the value of \( y \) below which a frivolous plaintiff will sue and above which he will not sue:

\[
\hat{y} = \frac{[qx - c_p - (1 - q)c_p']}{(1 - q)};
\]  

(5)

if \( y = \hat{y} \), frivolous plaintiffs are indifferent between suing and not suing. If sanctions are prohibited or are permitted but not brought, it is assumed that \( x \) is such that a frivolous plaintiff will sue—that is,

\[
qx > c_p.
\]  

(6)

Similarly, a legitimate plaintiff will file a suit when actions for sanctions are brought if

\[
px > c_p + (1 - p)c_p'.
\]  

(7)

It is assumed that \( x \) is such that (7) holds. Obviously, then, a legitimate plaintiff also will sue if sanctions are prohibited or are permitted but not brought.

**TYPES OF EQUILIBRIA**

First note that an equilibrium cannot exist if \( y > \hat{y} \): if \( y \) were that high, frivolous plaintiffs would not sue; the defendant then would not request sanctions; but if the defendant does not request sanctions, frivolous plaintiffs would sue. We will assume hereafter, therefore, that \( y \leq \hat{y} \).68

---

68. The argument in this paragraph presumes that the defendant either brings an action for sanctions with certainty or does not bring one with certainty. If a mixed strategy were permitted, in which the defendant brings an action with some probability, an equilibrium apparently would exist in which the penalty, \( y \), is as high as possible, the fraction of frivolous plaintiffs who sue, \( \alpha \), is relatively small, and the probability of the defendant seeking sanctions is correspondingly small. We do not consider this possibility both because it is more complicated and because it is not necessary to develop the main points of this article.
In the “no-sanctions equilibrium,” actions for sanctions are not brought and all frivolous plaintiffs sue. This equilibrium will occur if \( \phi(1) < \hat{\phi} \)—that is, if it is not worthwhile for the defendant to bring actions even when her probability of prevailing is relatively high (because all frivolous plaintiffs sue). This equilibrium will result if \( y \) is set low enough, since for any \( y \) less than \( c_d' \), \( \hat{\phi} > 1 \).

In the “sanctions equilibrium,” sanctions are requested and some frivolous plaintiffs sue. We assume that \( y = \hat{y} \) in this equilibrium for the following reason. If \( y \) were less than \( \hat{y} \), then all frivolous plaintiffs would sue and actions for sanctions would be brought against them. In this case, it would be socially preferable to set \( y \) low enough to switch to the no-sanctions equilibrium since the costs of sanctions proceedings would be avoided.\(^{69}\)

A sanctions equilibrium will occur if \( \phi(\alpha) > \hat{\phi} \)—that is, if the defendant’s chance of prevailing in an action for sanctions is sufficiently high. Given \( y = \hat{y} \), each frivolous plaintiff is indifferent between suing and not suing. In general, there will be a range of \( \alpha \) (the proportion of frivolous plaintiffs who sue) between some positive fraction and unity such that \( \phi(\alpha) \) exceeds \( \hat{\phi} \); an equilibrium corresponding to each \( \alpha \) in this range is possible. We assume, however, that \( \alpha < 1 \) since, if \( \alpha = 1 \), it would again be socially preferable to set \( y \) low enough to switch to the no-sanctions equilibrium (for the reason discussed in the previous paragraph).\(^{70}\)

**PART II ANALYSIS**

We will first show that it is possible to choose \( x \) and \( y \) in the sanctions equilibrium so as to maintain the level of the defendant’s costs that occurs in the no-sanctions equilibrium. Suppose initially that \( x \) is the same in both equilibria and let

\[
D_1 = \text{defendant's costs in the no-sanctions equilibrium};
\]

\[
D_2 = \text{defendant's costs in the sanctions equilibrium}.
\]

---

69. To complete this argument, it must be shown that the level of the defendant’s expected costs in the no-sanctions equilibrium can duplicate the level of the defendant’s expected costs that results if \( y \) is less than \( \hat{y} \), all frivolous plaintiffs sue, and actions for sanctions are brought. Note that the defendant’s expected costs in the latter case must be lower than in a no-sanctions equilibrium in which the award to a prevailing plaintiff is the same, because the defendant has the option of not seeking sanctions and therefore must be better off if she chooses to request them. Consequently, there generally exists a no-sanctions equilibrium with a lower award that will achieve the same level of the defendant’s expected costs.

70. Note that, if legitimate plaintiffs sometimes are mistakenly determined to be frivolous, a third type of equilibrium could occur in our model: the defendant seeks sanctions and no frivolous plaintiffs file suit.
Observe that:

\[ D_1 = N_f(qx + c_d) + N_e(px + c_d), \]  

and

\[ D_2 = \alpha N_f[qx + c_d - (1 - q)(\hat{y} - c_{d'})] + N_e[px + c_d + (1 - p)c_{d'}]. \]  

In the sanctions equilibrium, \( \phi(\alpha) > \hat{\phi} \), or equivalently (using (2) and (3)),

\[ \alpha N_f(1 - q)(\hat{y} - c_{d'}) > N_e(1 - p)c_{d'}. \]  

Substituting the right-hand-side of (10) for \( \alpha N_f(1 - q)(\hat{y} - c_{d'}) \) in (9) yields a higher level of the defendant's costs,

\[ \hat{D}_2 = \alpha N_f[qx + c_d] + N_e[px + c_d]. \]  

Thus, \( D_2 < \hat{D}_2 < D_1 \), where the second inequality follows from the assumption that \( \alpha < 1 \) in the sanctions equilibrium. In other words, if the award in the sanctions equilibrium is the same as in the no-sanctions equilibrium, the defendant's costs will be lower in the sanctions equilibrium. However, since \( D_2 \) is continuously increasing in \( x \) without bound, there exists a higher \( x \) in the sanctions equilibrium that can impose the same level of the defendant's costs as in the no-sanctions equilibrium. Given the resulting \( x, y \) is determined by (5).

Now assuming that the level of the defendant's costs is the same in the two equilibria, the social welfare comparison of the two equilibria becomes a comparison of their respective litigation costs. Let

\[ L_1 = \text{litigation costs in the no-sanctions equilibrium}; \]
\[ L_2 = \text{litigation costs in the sanctions equilibrium}. \]

Note that

\[ L_1 = (N_f + N_e)(c_p + c_d), \]  

and

\[ L_2 = (\alpha N_f + N_e)(c_p + c_d) + [\alpha N_f(1 - q) + N_e(1 - p)](c_{p'} + c_{d'}). \]  

The sanctions equilibrium will be preferable to the no-sanctions equilibrium if and only if

\[ L_1 - L_2 = [(1 - \alpha)N_f](c_p + c_d) \]
\[ - [\alpha N_f(1 - q) + N_e(1 - p)](c_{p'} + c_{d'}) > 0. \]
The first term, which is positive, is the savings in litigation costs in the original trials due to the deterrence of some frivolous suits by actions for sanctions. The second term, which is negative, is the increase in litigation costs associated with the actions.

The results in Part II regarding when sanctions are desirable can be demonstrated by taking the derivative of (14) with respect to each parameter. The sign of the derivative is unambiguous for the following parameters (the sign is indicated in parentheses): \( c_p (+), c_d (+), c_p' (-), c_d' (-), p (+), q (+), \) and \( N_f (-) \). A positive sign indicates that increases in the parameter favor sanctions, everything else equal, while a negative sign indicates that decreases in the parameter favor sanctions.

The only parameter whose derivative has an ambiguous sign is \( N_f \):

\[
\frac{\partial (L_1 - L_2)}{\partial N_f} = (1 - \alpha)(c_p + c_d) - \alpha(1 - q)(c_p' + c_d').
\]

(15)

The first term reflects the additional litigation costs saved by deterring frivolous suits initially as the number of potential frivolous plaintiffs increases. The second term represents the increase in litigation costs associated with the greater number of actions for sanctions as \( N_f \) increases. Clearly, either term could dominate.

PART III ANALYSIS

As noted in the text, if risk aversion were to provide a rationale for penalties, then the situation that would favor penalties most strongly would be when the defendant is risk averse and the plaintiff is risk neutral.

It will be useful to think of the risk faced by the defendant in the litigation process as stemming from two sources: the uncertainty whether, given a suit, she will win or lose; and the uncertainty whether she will be sued. We will now show that setting the penalty equal to the defendant’s litigation costs will, relative to not having sanctions at all, actually increase the first type of risk, but tend to decrease the second type.

First assume that the defendant has been sued by a frivolous plaintiff. In the absence of sanctions, the defendant would bear some risk as a result of such a suit since her costs would be \( c_d \) if she wins the suit and \( x + c_d \) if she loses the suit. The difference between these two outcomes, \( x \), provides a measure of the risk she bears. Now suppose that penalties are available and that they are set so as to compensate the defendant for her litigation costs (both the cost of the original trial and the cost of the sanctions proceeding). The defendant’s costs then would be zero if she wins the suit and brings an action for sanctions and \( x + c_d \) if she loses the suit. The difference has increased to \( x + c_d \). In other words, given a suit, awarding the defendant a penalty equal to her litigation costs creates greater uncertainty.

The second source of litigation risk for the defendant is whether she will
be sued at all. Within our model, there is no risk of this type since the fraction of suits brought is fixed (in equilibrium). More generally, however, a defendant could face substantial uncertainty with respect to whether she will be sued. If a suit is not brought, she bears no cost. If a suit is brought by a frivolous plaintiff and sanctions are not available, her expected payment is \( qx + c_d \) (since she pays \( c_d \) if she wins and \( x + c_d \) if she loses). If sanctions are available, however, her expected payment falls to \( qx + qc_d \) (she pays nothing if she wins and brings an action for sanctions, and she pays \( x + c_d \) if she loses). Thus, on average sanctions reduce the difference between the outcome of not being sued and being sued, and tend to reduce risk for this reason. But since, as noted above, they increase risk with respect to suits that are brought, their overall effect on the bearing of risk by the defendant is ambiguous. Thus, risk aversion does not provide a strong rationale for sanctions.

We will now explain why the level of a penalty that deters frivolous litigation bears little relationship to the level of a penalty that compensates the nonfrivolous party for her litigation costs. In our model, in which the purpose of penalties is deterrence, the penalty must equal \( \hat{y} \) defined by (5) in order to be consistent with the sanctions equilibrium. Because this penalty makes frivolous plaintiffs indifferent between suing and not suing, it should not be surprising that it does not depend at all on the defendant’s litigation costs. (It depends solely on the plaintiff’s probability of prevailing and the award in the original trial, and the plaintiff’s litigation costs in the original trial and in the sanctions proceeding.) A penalty that compensates the defendant for her litigation costs—that is, a penalty equal to \( c_d + c_d' \)—bears no relationship to \( \hat{y} \) and could be less than or greater than \( \hat{y} \).

In the text we emphasized that a deterrence-oriented penalty must be at least some multiple of the defendant’s cost of bringing an action for sanctions, in order to induce the defendant to bring such an action. This condition, too, must be satisfied in the sanctions equilibrium. It corresponds to (1) above, which can be rewritten as

\[
y > \frac{c_d}{\phi(\alpha)}.
\]

In a sanctions equilibrium, therefore, \( y \) must equal \( \hat{y} \) and (16) must be satisfied, which implies that \( \hat{y} > \frac{c_d}{\phi(\alpha)} \). In other words, the penalty must exceed the defendant’s cost of bringing an action for sanctions multiplied by the reciprocal of her chance of prevailing in that action. Viewing a deterrence-oriented penalty from this perspective suggests a closer relationship to a compensation-oriented penalty since the defendant’s cost of

71. A more complete analysis also would take into account the risks associated with suits by legitimate plaintiffs. Taking such risks into account, however, would not affect our general points regarding the risk-allocation effects of sanctions.
seeking sanctions enters as a factor in both. Even from this perspective, however, the relationship is only tangential (for the reasons discussed in the text).

PART IV ANALYSIS

We will now show why it generally is desirable to give some of the penalty paid by a frivolous party to the court or the treasury rather than to the nonfrivolous party. Let $y_p$ be the penalty paid by a frivolous plaintiff and let $y_d$ be the amount paid to the defendant. Thus, the proposition is that $y_d$ is less than $y_p$ when both are chosen optimally.

As discussed in the text, to demonstrate this result we assume that defendants differ in terms of their cost of bringing actions for sanctions. For simplicity, suppose that there are just two types of defendant—a fraction $\gamma$ with low sanctions-related litigation costs, $c_{d1}$, and a fraction $(1 - \gamma)$ with high litigation costs, $c_{d2}$.

Imagine starting from the usual sanctions equilibrium in which the penalty is not decoupled and defendants always seek sanctions against losing plaintiffs. Using the present notation, this corresponds to assuming that $y_d = y_p = \hat{y}$ and $c_{d1} < c_{b1} < \phi(\alpha)\hat{y}$.

Now suppose that the penalty is decoupled and that the payment to the defendant ($y_d$) is lowered so that $c_{d1} < \phi(\alpha)y_d < c_{b1}$. Consequently, only defendants with low litigation costs will request sanctions. Since the probability that a losing plaintiff will have an action for sanctions brought against him now is $\gamma$, a frivolous plaintiff will be indifferent between suing and not suing if

$$q_x - (1 - q)\gamma y_p = c_p + (1 - q)\gamma c_p'.$$

It is straightforward to show, using (6), that the $y_p$ that solves (17) is decreasing in $\gamma$; thus, if $\gamma$ is less than 1, $y_p$ exceeds $\hat{y}$. Consequently, in the new sanctions equilibrium with decoupled liability, $y_d < \hat{y} < y_p$.

An obvious advantage of decoupling the penalty in this way is that litigation costs are reduced since fewer actions for sanctions are brought. Moreover, the level of the defendants' costs that occurred in the original sanctions equilibrium (and therefore the defendants' incentive to engage in the conduct giving rise to the suits) can be maintained in the decoupled sanctions equilibrium: Both types of defendant are worse off initially (their expected costs rise) because their payment has been lowered. But this effect can be offset by lowering the award they have to pay to successful plaintiffs in the original trial, $x$. 72

---

72. This statement is correct in terms of the aggregate costs borne by the defendants. A more complete analysis would have to take into account that the two types of defendant are made worse off by different degrees because they behave in different ways. (One requests sanctions and the other does not.)
PART V ANALYSIS

Now suppose that mistakes can occur in the sanctions proceeding. In particular, assume that the court incorrectly imposes penalties on losing legitimate plaintiffs with probability θ.

This possibility obviously lowers legitimate plaintiffs' incentive to sue. Now a legitimate plaintiff will bring a suit if

\[ px - (1 - p)θy > c_p + (1 - p)c_p'. \]  \hspace{1cm} (18)

The important point to note is that if θ is high enough to reverse (18), then raising the award in the initial trial, x, can restore the incentive for legitimate plaintiffs to sue.

If x is increased, however, the penalty (y) also must be raised so that frivolous plaintiffs remain indifferent between suing and not suing. We will now show that it is possible to pick a higher x and a higher y such that the incentive of legitimate plaintiffs to sue is maintained at the level that would occur in the absence of mistakes and such that frivolous plaintiffs remain indifferent between suing and not suing. Let \( x^0 \) and \( y^0 \) be the new award and penalty respectively when there are mistakes, and let x and y be the original values when there are not mistakes. Then the following two conditions must hold:

\[ px^0 - (1 - p)θy^0 - c_p - (1 - p)c_p' = px - c_p - (1 - p)c_p'; \]  \hspace{1cm} (19)

\[ y^0 = (qx^0 - c_p - (1 - q)c_p')/(1 - q). \]  \hspace{1cm} (20)

If (19) is satisfied, the value of suing for legitimate plaintiffs is preserved, and if (20) is satisfied, frivolous plaintiffs remain indifferent between suing and not suing. It is clear from (19) and (20) that there exist an \( x^0 \) and a \( y^0 \) that satisfy both conditions.\(^{74}\)

---

\(^{73}\) In order to focus on the effects of mistakes in the sanctions process, we return to our initial assumption that sanctions are not decoupled, so that y again represents both what a plaintiff who is found to be frivolous pays and what the defendant receives.

\(^{74}\) The resulting \( x^0 \) and \( y^0 \) also impose the same expected costs on the defendant and therefore leave her behavior unchanged. To see this, recall from (9) that in the absence of mistakes the defendant's expected costs are

\[ αN_f[qx + c_d - (1 - q)(\tilde{y} - c_d')] + N_d[px + c_d + (1 - p)c_d']. \]  \hspace{1cm} (i)

In the presence of mistakes (and with \( x^0 \) and \( y^0 \)), the defendant's expected costs are

\[ αN_f[qx^0 + c_d - (1 - q)(y^0 - c_d')] + N_d[px^0 - (1 - p)θy^0 + c_d + (1 - p)c_d']. \]  \hspace{1cm} (ii)

The first term in brackets in (i) equals the first term in brackets in (ii) if
To see that the new $x^0$ and $y^0$ exceed the original values, (19) and (20) can be differentiated totally and solved to obtain

$$dx^0/d\theta = [(1 - q)(1 - p)y^0]/[p(1 - q) - (1 - p)q\theta]; \quad (21)$$

$$dy^0/d\theta = [q(1 - p)y^0]/[p(1 - q) - (1 - p)q\theta]. \quad (22)$$

Since $p > q$, $1 - q > 1 - p$, and $0 \leq \theta \leq 1$, the denominator in (21) and (22) is positive. Therefore, both $dx^0/d\theta$ and $dy^0/d\theta$ also are positive. Note from (19) and (20) that if $\theta = 0$, then $x^0 = x$ and $y^0 = y$. It follows that, if the probability of mistakenly imposing a penalty on a legitimate plaintiff is positive, then $x^0 > x$ and $y^0 > y$.\(^{75}\)

**PART VI ANALYSIS**

The first alternative to sanctions discussed in the text is a modified English rule in which a losing plaintiff in the original trial pays a penalty to the defendant (and no inquiry is made about the plaintiff's type). Let $z$ be this penalty.

Under the modified English rule, frivolous plaintiffs will be indifferent between suing and not suing if

$$qx - (1 - q)z = c_p. \quad (23)$$

Since $p > q$, legitimate plaintiffs will strictly prefer to sue when (23) holds. Thus, it is possible to choose $x$ and $z$ such that the left-hand side of (23) is slightly less than the right-hand side—thereby deterring all frivolous plaintiffs—while still providing an incentive for legitimate plaintiffs to sue.

It also is possible under the modified English rule to choose $x$ and $z$ to impose any level of expected costs on the defendant. Under this rule, the defendant's expected costs are

$$Ne[px - (1 - p)z + c_d]. \quad (24)$$

Solving for $z$ as a function of $x$ from (23) and substituting the result in (24) yields an expression equal to

$$Ne[(p - q)x/(1 - q)] \quad (25)$$

From (5), $qx - (1 - q)y = c_p + (1 - q)c_{p'}$. From (20), $qx^0 - (1 - q)y^0 = c_p + (1 - q)c_{p'}$. Hence, (iii) is satisfied. The second term in brackets in (i) equals the second term in brackets in (ii) if $px = px^0 - (1 - p)\theta y^0$, which follows immediately from (19). Thus, (i) and (ii) are equal.

75. It also can be shown using a similar analysis that the optimal award in the initial trial and the optimal penalty would increase if there were a positive probability of not imposing a penalty on a frivolous plaintiff.
plus a constant term, which grows without bound as \( x \) increases. Thus, the level of expected costs imposed on the defendant under a sanctions equilibrium can be duplicated under the modified English rule.\(^7\)

The second alternative to sanctions is to decouple liability in the original trial. Let \( x_p \) be the award given to a successful plaintiff and let \( x_d \) be the amount paid by a losing defendant. Assuming the American rule for allocating legal costs governs, frivolous plaintiffs can be made indifferent between suing and not suing if \( x_p \) is chosen to satisfy \( qx_p = c_p \). Since \( p > q \), legitimate plaintiffs will strictly prefer to sue. Thus, by lowering \( x_p \) slightly, all frivolous plaintiffs will be deterred, but legitimate plaintiffs still will sue. If \( x_p \) is chosen in this way, the expected costs borne by the defendant are \( N_d[px_d + c_d] \). By choosing \( x_d \) appropriately, any level of expected costs imposed on the defendant can be achieved.

---

76. This statement might need to be qualified if the level of expected costs imposed on the defendant under the sanctions equilibrium is relatively low, because it is possible that the lowest level of expected costs under the modified English rule (which occurs when \( z = 0 \) and \( x = c_p/q \)) will exceed this level.

The values of the award and the penalty used in the numerical example in Part VI to illustrate the modified English rule can be calculated from the preceding formulas. First, set (25) plus the referred-to constant term equal to $13,000,000 and solve for \( x \) to get $164,000. Then, using this result, solve for \( z \) from (23) to get $56,000.