

4-1-1988

# Statistical Error and Legal Error—Type One and Type Two Errors and the Law

R. S. Radford

---

## Recommended Citation

R. S. Radford, *Statistical Error and Legal Error—Type One and Type Two Errors and the Law*, 21 Loy. L.A. L. Rev. 843 (1988).  
Available at: <https://digitalcommons.lmu.edu/llr/vol21/iss3/2>

This Article is brought to you for free and open access by the Law Reviews at Digital Commons @ Loyola Marymount University and Loyola Law School. It has been accepted for inclusion in Loyola of Los Angeles Law Review by an authorized administrator of Digital Commons@Loyola Marymount University and Loyola Law School. For more information, please contact [digitalcommons@lmu.edu](mailto:digitalcommons@lmu.edu).

# STATISTICAL ERROR AND LEGAL ERROR

## Type One and Type Two Errors and the Law

*R.S. Radford\**

### I. INTRODUCTION

In *Philadelphia Newspapers Inc. v. Hepps*,<sup>1</sup> the Supreme Court wrestled with the fundamental inadequacy of the legal system as a means of determining truth. Justice O'Connor expressed the Court's frustration in the context of libel suits:

There will always be instances when the fact finding process will be unable to resolve conclusively whether . . . speech is true or false; it is in those cases that the burden of proof is dispositive. Under a rule forcing the plaintiff to bear the burden of showing falsity, there will be some cases in which plaintiffs cannot meet their burden despite the fact that the speech is in fact false. . . . Similarly, under an alternative rule placing the burden of showing truth on defendants, there would be some cases in which defendants could not bear their burden despite the fact that the speech is in fact true. . . . Under either rule, then, the outcome of the suit will sometimes be at variance with the outcome that we would desire if all speech were either demonstrably true or demonstrably false.<sup>2</sup>

Justice O'Connor's concern with the consequences of uncertainty and error was by no means unique. These issues have a long history in Western jurisprudence, reaching into all branches of the law. The *Hepps* passage is noteworthy mainly for its clarity in highlighting two important points. First, the decision implies a model in which all defamatory statements are either true or false in some objective sense. The accuracy of judicial fact-finding can therefore be verified, at least in theory, by reference to objective reality. This is a completely arbitrary assumption, but one without which the fact-finding process would be meaningless.

Second, Justice O'Connor observes that legal decisions may fail to coincide with objective truth, and that this divergence may take either of

---

\* B.S.B.A., Rockhurst College; M.A., University of Southern California; J.D., University of Southern California. I am grateful for helpful comments by Matt Spitzer, Richard Craswell, Karen Lash and Duane Okamoto.

1. 475 U.S. 767 (1986).

2. *Id.* at 776.

two forms. Through this observation she echoes commentators such as Schauer,<sup>3</sup> Redish<sup>4</sup> and Epstein.<sup>5</sup> Furthermore, the *Hepps* opinion implicitly recognizes that, for a given quantity and quality of evidence, a trade-off exists between the two types of error. Reducing the number of erroneous convictions (or judgments for plaintiffs) is normally achieved at the cost of more erroneous acquittals (or judgments for defendants).<sup>6</sup> The *Hepps* dilemma, seen in this light, seems remarkably similar to the general statistical concept of Type I and Type II errors.<sup>7</sup> However, this identification has been strongly resisted by jurists and scholars who claim that legal decision-making cannot (or should not) be analogized to the quantitative methods of social science.<sup>8</sup> This Article will review the main themes of this dispute, and it will show that a legal model based on statis-

3. Schauer has written:

[E]rrors produced can be errors of under-inclusion or errors of over-inclusion. Because we do not have certain knowledge of the true state of affairs, we cannot be absolutely certain that a verdict of innocent is not a decision to free one who is in fact guilty (an error of under-inclusion), or that a verdict of guilty is not a decision to punish one who is in fact innocent (an error of over-inclusion).

F. SCHAUER, *FREE SPEECH: A PHILOSOPHICAL ENQUIRY* 137 (1982).

4. Redish has observed:

If the asserted governmental interest is of overwhelming importance . . . and there is serious doubt concerning the viability of less restrictive means of achieving that end, the risk of an overestimation of the viability of alternative measures might be significant, and might cause the Court to think twice before it invalidates legislation on overbreadth grounds. If, on the other hand, the government interest, while not insubstantial, is such that no overriding threat to physical safety or national security would result from failure to promote it, the Court might be more willing to take the risk of overestimating the viability of less drastic means.

M. REDISH, *FREEDOM OF EXPRESSION: A CRITICAL ANALYSIS* 235 (1984).

What Redish does not make explicit is that the Court must weigh these risks under uncertainty as to *both* the seriousness of the threat to state interests and the viability of less restrictive means.

5. "The constant challenge . . . is to balance two kinds of error. Errors of overinclusion occur when the regulation sweeps wider than necessary to control the identified evil; those of underinclusion occur when the regulation does not reach all instances of the evil in question."

R. EPSTEIN, *TAKINGS: PRIVATE PROPERTY AND THE POWER OF EMINENT DOMAIN* 127-28 (1985).

6. Kaye, *The Paradox of the Gatecrasher and Other Stories*, 1979 ARIZ. ST. L.J. 101, 103 n.13. The same two kinds of error, and the relationship between them, have been recognized in the application of administrative regulations. See Craswell, *Interpreting Deceptive Advertising*, 65 B.U.L. REV. 657, 678, 710-13 (1985).

7. For an explanation of Type I and Type II errors, see *infra* text accompanying note 57.

Justice O'Connor's account of the problem is a special case because both types of error may occur regardless of which party bears the burden of proof. Defendants may be judged to have satisfied their burden even when the speech was in fact false; when plaintiffs bear the burden they may prevail even though the speech was in fact true. In these cases it is equally true that the outcome will "be at variance with the outcome we would desire if all speech were either demonstrably true or demonstrably false." *Hepps*, 475 U.S. at 776.

8. See *infra* text accompanying notes 74-122.

tical error can be especially useful in evaluating the impact of alternative standards and burdens of proof.

## II. UNCERTAINTY AND ERROR

### A. *Uncertainty and the Law*

Regardless of whether reality is objectively determinate,<sup>9</sup> human beings lack the means of knowing reality with certainty.<sup>10</sup> The available data are always imperfect, as is our ability to interpret the data.<sup>11</sup> What we call knowledge of facts is more accurately described as an awareness of a distribution of probabilities.

"Probabilities are as real as masses," said H. Margenau. Yet the truth is rather the reverse: masses are as real as probabilities. Indeed, anything we can now say about masses depends on what we can say about probabilities. . . . The verification of propositions about probabilities is . . . the only fundamental issue. Everything else depends upon this verification.<sup>12</sup>

Saying that *X* has been proven true amounts to a claim that *X* has been associated with a very high probability. Conversely, saying that *X* is very highly probable can be the equivalent of claiming that *X* is true. Estimates of the likelihood of *X* can serve to express either our certainty of truth or the extent of our ignorance.<sup>13</sup>

The problem of uncertainty is especially acute in the context of trials. An essential function of a trial is the resolution of uncertainty.<sup>14</sup> Yet uncertainty always implies error. While courts are required to resolve disputes of fact and law, the evidence is always doubtful; even a defendant's confession or admission of fault cannot raise the probability of guilt or liability to 100%.<sup>15</sup> Worse, the facts in dispute usually involve past

---

9. The assumption of objective determinacy, which underlay Justice O'Connor's opinion in *Hepps*, will be employed throughout this Article.

10. See, e.g., R. NOZICK, *PHILOSOPHICAL EXPLANATIONS* (1981).

11. See Heiner, *Imperfect Decisions and the Law: On the Evolution of Legal Precedent and Rules*, 15 J. LEGAL STUD. 227 (1986).

12. N. GEORGESCU-ROEGEN, *THE ENTROPY LAW AND THE ECONOMIC PROCESS* 54 (4th printing 1981) (footnote omitted).

13. Ball, *The Moment of Truth: Probability Theory and Standards of Proof*, 14 VAND. L. REV. 807, 809 (1961).

14. Zeisel, *Statistics as Legal Evidence*, 15 INT'L ENCYCLOPEDIA SOC. SCI. 246, 246 (1968). Of course, uncertainty can be "resolved" in at least two different ways. One is to actually determine the underlying truth of the matter; the other is to settle the issue by declaratory fiat. Which of these methods is more appropriate to judicial factfinding is disputed. See *infra* text accompanying notes 83-92.

15. See Jaffee, *Of Probativity and Probability: Statistics, Scientific Evidence, and the Calculus of Chance At Trial*, 46 U. PITT. L. REV. 925, 935 (1985).

events. Even if crystal-clear evidence was once available, its utility may have faded or disappeared before it could be presented in court. Thus, from a purely practical viewpoint, certainty of legally relevant facts may never be attainable without the aid of time travel.<sup>16</sup>

Uncertainty is also felt in the legal process at other levels. This Article, like most of the existing literature on the subject, will deal mainly with "rule uncertainty"—uncertainty in the application of legal rules to individuals. Rule uncertainty is a fairly obvious phenomenon, having directly observable effects in the form of erroneous verdicts and other outcomes. However, legal rules are not imposed for their own sake. In most cases they are intended to have some incentive effect in influencing social behavior.<sup>17</sup> The incentives created by a given rule may or may not promote the behavior the rule seeks to encourage—a situation that might be described as "incentive uncertainty." Incentive uncertainty reflects the risk of specification error (a simple mismatch between the formulation of a rule and the desired behavioral norm) and the risk of error in predicting incentive effects. Behavioral incentives are shaped not by rules per se, but by rules as they are applied—including any uncertainty in their application. Thus, it may sometimes happen that rules entailing unnecessary risks of error in application create optimal incentives for complying with desired behavioral norms.<sup>18</sup> Do courts use the dynamics of uncertainty in this way, accepting rule uncertainty and errors at trial in order to maximize compliance with underlying norms? No evidence exists that courts possess the sophistication needed to pursue a conscious policy of this sort. Nonetheless, any model of legal uncertainty must recognize this dimension of the issue.

### B. *Legal Accommodation of Uncertainty and Error*

Although it is seldom examined and leaves some unmoved,<sup>19</sup> awareness of the risk of error pervades the theory and practice of law.<sup>20</sup> Its presence can be felt in the requirements of due process and the prac-

---

16. See Kaye, *Paradoxes, Gedanken Experiments, and the Burden of Proof: A Response to Dr. Cohen's Reply*, 1981 ARIZ. ST. L.J. 635, 642.

17. See, e.g., Leubsdorf, *Remedies For Uncertainty*, 61 B.U.L. REV. 132 (1981).

18. See, e.g., Craswell & Calfee, *Deterrence and Uncertain Legal Standards*, 2 J. L. ECON. & ORGANIZATION 2 (1986).

19. "[H]e who falls by a mistaken sentence, may be considered as falling for his country." Paley, *Moral and Political Philosophy*, in 3 THE WORKS OF WILLIAM PALEY, D.D. 315, 315 (1838), quoted in Birmingham, *Remarks on "Probability" in Law: Mostly, a Casenote and a Book Review*, 12 GA. L. REV. 535, 537 (1978).

20. See, e.g., Radin, *Risk-Of-Error Rules and Non-Ideal Justification*, 28 NOMOS 33 (J. Pennock & J. Chapman eds., 1986).

tice of deferring to precedent. A conscious risk of error is implicit in the structure of criminal penalties, as in the practice of punishing attempted crimes less severely than completed crimes.<sup>21</sup> The potential cost of error is a factor in determining the amount of resources devoted to litigation,<sup>22</sup> and the risk of error has been cited as a reason for not awarding punitive damages in tort cases.<sup>23</sup> Posner has described the purpose of legal procedures as the minimization of direct costs plus the costs of error.<sup>24</sup> This is consistent with the Supreme Court's practice in procedural due process cases of balancing the marginal costs of stricter procedural protections against the marginal benefits of reduced error.<sup>25</sup> Finally, the law's accommodation of uncertainty and error is nowhere more evident than in the assignment of the burden and choice of a standard of proof.

### 1. The standard of proof

Courts do not attempt to establish the certainty of disputed facts, but only their probability under the circumstances presented by the evidence.<sup>26</sup> The law recognizes that some degree of uncertainty will persist to the conclusion of trial. It therefore specifies the residual doubt compatible with a finding of guilt or liability, expressed as the appropriate standard of proof. These standards are simply guidelines as to the unresolved risk of error that courts are willing to accept when reaching a decision. The amount of residual uncertainty tolerated by a standard tends to vary inversely with the potential loss risked by defendants in cases coming under it.<sup>27</sup> Regardless of the magnitude of the potential loss, however, even the strictest standard is understood to mean something less than proof to an absolute certainty.<sup>28</sup>

In most civil litigation, proof is established by a preponderance of the evidence.<sup>29</sup> On its face, the preponderance standard simply means

---

21. See, e.g., R. POSNER, *ECONOMIC ANALYSIS OF LAW* § 7.3 (3d ed. 1986).

22. See, e.g., *id.* at § 7.2.

23. See, e.g., *id.* at § 6.10.

24. Posner, *An Economic Approach to Legal Procedure and Judicial Administration*, 2 J. LEGAL STUD. 399, 399-401 (1973).

25. See, e.g., *Matthews v. Eldridge*, 424 U.S. 319 (1976).

26. See Broun & Kelly, *Playing the Percentages and the Law of Evidence*, 1970 U. ILL. L. FORUM 23, 26-27.

27. In at least one field of law, the Court has rejected the reasonable doubt standard specifically because the state might never be able to win a case under it. See *Addington v. Texas*, 441 U.S. 418, 429 (1979).

28. Bass, Gesser & Mount, *Scientific Statistical Methodology and the Doctrine of "Reasonable Doubt" in Criminal Law; (With Special Reference to the Breath Analysis for Blood Alcohol) Empirical Fact or Legal Fiction?*, 5 DALHOUSIE L.J. 350 (1979).

29. See, e.g., R. LEMPERT & S. SALTZBURG, *A MODERN APPROACH TO EVIDENCE* 797 (2d ed. 1982).

that something has been shown to be more probable than not.<sup>30</sup> More formally, the standard can be said to require a showing of probabilities greater than 0.5,<sup>31</sup> or "odds of at least 51 to 49 that such-and-such has taken place or will do so."<sup>32</sup>

However, a fact proven for some judicial purposes is not necessarily proven for others. For some quasi-criminal issues, proof must be shown by clear and convincing evidence,<sup>33</sup> while criminal convictions require proof beyond a reasonable doubt.<sup>34</sup> These differing standards can be interpreted as variations in the degree of probability needed to establish issues bearing on guilt or liability.<sup>35</sup> This is the sense in which Zeisel refers to standards of proof as "canon[s] of probability."<sup>36</sup> A requirement of proof to an absolute certainty would be tantamount to saying that there must be a 100% probability that a premise is true.<sup>37</sup> Below this extreme case, transforming standards of proof into probability estimates requires some arbitrariness. For instance, if the preponderance standard can be expressed as a probability greater than 0.5, then proof beyond a reasonable doubt might be defined as a probability greater than 0.9 or 0.95.<sup>38</sup> In any event, no reason exists to suppose the standards cannot be quantified.<sup>39</sup> Indeed, research has shown that judges can in fact assign probability values to standards of proof.<sup>40</sup> In a 1969 survey, state and federal judges associated the phrase "beyond a reasonable

---

30. See Kaye, Book Review, 89 YALE L.J. 601, 603 (1980). But see Brook, *Inevitable Errors: The Preponderance of the Evidence Standard in Civil Litigation*, 18 TULSA L.J. 79, 88-96 (1982) (requiring "actual belief" of truth of contested facts). Cohen takes yet another approach, showing that two probabilistic variables are involved in reaching verdicts: (1) the threshold probability of guilt (generally defined as >0.50; i.e., more probable than not) and (2) the confidence level used in constructing the interval estimate about this figure. N. Cohen, *Confidence in Probability: Burdens of Persuasion in a World of Imperfect Knowledge*, 60 N.Y.U. L. REV. 385 (1985). In Cohen's model, the key variable for distinguishing between standards of proof is the confidence level. I use the more traditional approach, varying the threshold probability.

31. See Brook, *supra* note 30, at 85. But see Jaffee, *supra* note 15, at 936-38 (preponderance standard interpreted as weight of evidence, not probability).

32. Lord Simon of Glaisdale, quoted in R. EGGLESTON, *EVIDENCE, PROOF, AND PROBABILITY* 132 (2d ed. 1983).

33. See, e.g., R. LEMPERT & S. SALTZBURG, *supra* note 29.

34. See *In re Winship*, 397 U.S. 358 (1970).

35. See Broun & Kelly, *supra* note 26, at 26-27.

36. Zeisel, *supra* note 14, at 246.

37. Bass, Gesser & Mount, *supra* note 28, at 351.

38. Kaye, *supra* note 6, at 103.

39. See Williams, *The Mathematics of Proof—I*, 1979 CRIM. L. REV. 297, 303.

40. Simon, *Judges' Translations of Burdens of Proof into Statements of Probability*, 1969 THE TRIAL LAW. GUIDE 103. This study suffers from serious methodological defects, stemming in part from a low response rate. See *id.* at 104-06. For comparisons of quantitative

doubt" with a mean probability of 89%;<sup>41</sup> "by a preponderance of the evidence" was associated with a mean probability of 61%.<sup>42</sup>

What is the point of using tougher standards of proof in criminal and quasi-criminal cases? More stringent proof is meant to require greater certainty—i.e., a higher degree of probability—in cases involving more costly errors. Despite references to balancing their risks, the legal system does not always regard the two kinds of error as equally weighted. This is especially clear in criminal trials, where Anglo-American procedures still reflect Blackstone's dictum that it is better for ten guilty defendants to go free than for one of the innocent to be punished.<sup>43</sup> In economic terms, convicting the innocent involves greater negative externalities—social costs beyond those borne by the parties—than acquitting the guilty.<sup>44</sup> The various standards try to balance the relative costs of the two types of error, *given* that the costs of errors against defendants are accorded greater weight as they rise.<sup>45</sup>

Changes in standards of proof implicitly trade off one sort of error for the other. In *Addington v. Texas*,<sup>46</sup> the Supreme Court came close to suggesting that errors in favor of plaintiffs occur under the preponderance standard simply because factfinders do not take the cost of error seriously enough. "Increasing the burden of proof is one way to impress the factfinder with the importance of the decision and thereby perhaps to reduce the chances" of erroneous decisions.<sup>47</sup> Yet in the same opinion, Chief Justice Burger described the reasonable doubt standard as an at-

---

interpretations by judges, jurors and students, see Simon & Mahan, *Quantifying Burdens of Proof: A View from the Bench, the Jury, and the Classroom*, 5 L. & SOC'Y REV. 319 (1971).

In a similar study at the University of Manchester, however, subjects were reported to have grossly varied in their interpretation of the reasonable doubt standard. One-third fixed it below .70, while another one-third put it above .90. R. EGGLESTON, *supra* note 32, at 118-19, citing J. COHEN & I. CHRISTENSEN, INFORMATION AND CHOICE 62 (1970).

41. Simon, *supra* note 40, at 108.

42. *Id.* at 112.

43. See Zeisel, *supra* note 14, at 246-47. See also Nagel & Neef, *Deductive Modeling to Determine an Optimum Jury Size and Fraction Required to Convict*, 1975 WASH. U.L.Q. 933, 945-46. Both Fortescue and Frederick the Great apparently preferred a ratio of twenty-to-one in favor of acquitting the guilty. Fletcher, *Two Kinds of Legal Rules: A Comparative Study of Burden-of-Persuasion Practices in Criminal Cases*, 77 YALE L.J. 880, 882 (1968). Fletcher notes the seeming anomaly that the common law often shifts the burden of proof to defendants pleading a justification or excuse in murder cases. *Id.* at 882-83. This implies a preference to condemn wrongly a defendant who killed in self defense, rather than to mitigate the charges against a knowing murderer.

44. See R. POSNER, *supra* note 21, at § 21.3.

45. See Brook, *supra* note 30, at 85; Thompson, *Liability and Individualized Evidence*, 49 LAW & CONTEMP. PROBS. 199, 215 (1986).

46. 441 U.S. 418 (1979).

47. *Id.* at 427.



tempt to minimize the risk of erroneous convictions "even at the risk that some who are guilty might go free."<sup>48</sup> This description obscures the key point that some who are guilty will go free (and some who are innocent will be convicted) under *any* standard applied in American courts. Applying the reasonable doubt test means that *more* of the guilty will be allowed to go free, as a cost of reducing the number of innocent who would be convicted under a less rigorous standard.

Whether innocent defendants actually benefit from more stringent standards of proof depends on whether there is an offsetting adjustment in the severity of penalties. A given level of deterrence may be achieved by relatively light penalties at the preponderance standard, or relatively heavy penalties under the reasonable doubt standard. It is not obvious how to gauge the social benefit of convicting fewer of the innocent, but sentencing them to longer terms.<sup>49</sup>

## 2. The burden of proof

In *Hepps*, Justice O'Connor observed how the placement of the burden of proof affects the relative risk of the two types of error.<sup>50</sup> As she implied, this burden can be shifted from one party to the other to correct what the courts perceive as an imbalance between the two risks. The burden of proof is composed of the burden of production (the requirement of advancing evidence to support one's case) and the burden of persuasion (the requirement that one's case must meet some threshold of plausibility).<sup>51</sup> This Article emphasizes the burden of persuasion, but the analysis will apply generally to the burden of production as well.

The significance of the burden of persuasion is that, when the evidence on a disputed point is evenly balanced (allowing for the relevant standard of proof), the party bearing the burden loses.<sup>52</sup> Some of these "tie" verdicts will be mistaken, and all of the mistakes will go against the same side. By shifting the burden of persuasion, the courts shift this block of errors from defendants to plaintiffs or vice versa. This can be an effective way of announcing a judicial reevaluation of the relative costs of error.<sup>53</sup>

---

48. *Id.* at 428 (citation omitted).

49. See G. TULLOCK, TRIALS ON TRIAL: THE PURE THEORY OF LEGAL PROCEDURE 82-86 (1980).

50. See *supra* text accompanying note 2.

51. See Kaye, *Statistical Significance and the Burden of Persuasion*, 46 LAW & CONTEMP. PROBS. 13, 15 n.11 (1983).

52. See R. EGGLESTON, *supra* note 32, at 105.

53. See citations in Kaye, *supra* note 51, at 16-17.

### C. Statistical Error and Legal Error

In some sense, every trial can be viewed as a hypothesis test.<sup>54</sup> The disputed issues are generally empirical propositions, which may be shown to be more or less probable under the circumstances of the case, but can seldom be proven as absolutely true.<sup>55</sup> In criminal cases the presumption of innocence is analogous to the null hypothesis that the defendant is not guilty. In civil trials the null hypothesis is that the party who does not bear the burden of proof should prevail. Establishing guilt or liability requires evidence sufficiently probative to meet the standard of proof. That is, the data tested must be conclusive enough to reject the null hypothesis at a specified confidence level.<sup>56</sup>

Statisticians are familiar with two types of error that may arise in testing hypotheses. Rejecting a null hypothesis that is in fact true is known as Type I error; accepting (failing to reject) a hypothesis that is in fact false is Type II error.<sup>57</sup> In general, an inverse relationship exists between the relative incidence of the two kinds of error.<sup>58</sup> Type I errors can be reduced merely by increasing the level of confidence needed to reject the null hypothesis; however, this will simultaneously increase the risk of Type II error.<sup>59</sup> In a model of trials as hypothesis tests, Type I errors would be mistaken convictions or findings of liability; Type II errors would be mistaken verdicts in favor of defendants.<sup>60</sup> For example, imagine a new evidentiary rule excluding identifications based on police

---

54. For a review of hypothesis testing and the use of null hypotheses, see R. WEHMHOFER, *STATISTICS IN LITIGATION: PRACTICAL APPLICATIONS FOR LAWYERS* § 3.25 *et. seq.* (1985). The model of trials as hypothesis tests is clearest for binary verdict sets: guilty or not guilty, liable or not liable. For decisions along a continuum, such as determining the amount of damages or the length of discretionary sentences, the simple analogy to statistical tests would not apply.

55. See Ball, *supra* note 13.

56. "At the ninety-five percent confidence level, five percent of all defendants who would win if their true probability of liability were known will incorrectly lose." Cohen, *supra* note 30, at 411.

57. See, e.g., T. WONNACOTT & R. WONNACOTT, *INTRODUCTORY STATISTICS FOR BUSINESS AND ECONOMICS* 272-73 (3d ed. 1984); P. EWART, J. FORD & C. LIN, *APPLIED MANAGERIAL STATISTICS* 188-93 (1982).

58. Although the two risks of error are inversely related, they are not simple complements. See, e.g., Cohen, *supra* note 30. The combined risks of Type I and Type II error would equal unity only on the assumption that all cases are decided incorrectly. If decisions were handed down randomly there should be a 50 percent likelihood that neither error will occur. See *infra* note 60 and accompanying text.

59. R. WEHMHOFER, *supra* note 54, at § 3.27.

60. The probability of Type I error is traditionally designated as  $\alpha$ ; the risk of Type II error is  $\beta$ . The relationship between the two kinds of error can be illustrated with a contingency table:

line-ups. Assuming that both guilty and innocent defendants are sometimes convicted on the marginal weight of such evidence, exclusion would reduce the number of mistaken convictions (lowering Type I error), but simultaneously increase the number of mistaken acquittals (raising Type II error).

The Anglo-American legal system has traditionally reserved its greatest concern for avoiding Type I errors.<sup>61</sup> Presumably, it would be ideal to reduce the risk of such errors to zero.<sup>62</sup> However, since this is statistically impossible so long as defendants are tried and found guilty, what should be the practical goal? At least five alternative approaches to legal error have been advanced:<sup>63</sup>

1. Ignore the problem.<sup>64</sup>
2. Minimize total errors.<sup>65</sup>
3. Minimize the cost of errors.<sup>66</sup>
4. Equalize the incidence of errors.<sup>67</sup>
5. Minimize the number of large errors.<sup>68</sup>

Each of these policies would require a different matrix of standards and

STATE OF THE WORLD	DECISION	
	For Plaintiff	For Defendant
	Defendant Is Guilty [Liable]	Type I Error (Probability: $\alpha$ )
Defendant Is Innocent [Not Liable]	Correct Decision (Probability: $1 - \beta$ )	Type II Error (Probability: $\beta$ )

61. See, e.g., Epstein, *Judicial Review: Reckoning on Two Kinds of Error*, 4 CATO J. 711, 717 (1985).

62. See, e.g., T. WONNACOTT & R. WONNACOTT, *supra* note 57; P. EWART, J. FORD & C. LIN, *supra* note 57.

63. The five general approaches listed here are ways of directly coping (or refusing to cope) with uncertainty and error. Another option is to treat legal error as an indirect or instrumental element of the system. See *supra* text accompanying notes 17-18.

64. See, e.g., Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329 (1971).

65. See, e.g., Friedman, *Trial by Jury: Criteria for Convictions, Jury Size and Type I and Type II Errors*, 26 AM. STATISTICIAN 21 (1972).

66. See, e.g., Kaye, *The Laws of Probability and the Law of the Land*, 47 U. CHI. L. REV. 34 (1979).

67. See, e.g., G. TULLOCK, *supra* note 49.

68. See Orloff & Stedinger, *A Framework for Evaluating the Preponderance of the Evidence Standard*, 131 U. PA. L. REV. 1159 (1983).

procedures. A full understanding of the issues requires a closer look at the nature of legal and statistical error.

A trial is an exercise in inferring reality from the appearance of reality. The appearance perceived by the court depends on the quality and quantity of evidence, and the skill with which it is presented. To illustrate, consider the problem addressed in *Hepps*, of determining whether a defamatory statement was true or false. All such statements can be ranked along a scale of perceived or apparent truth, as determined in court.

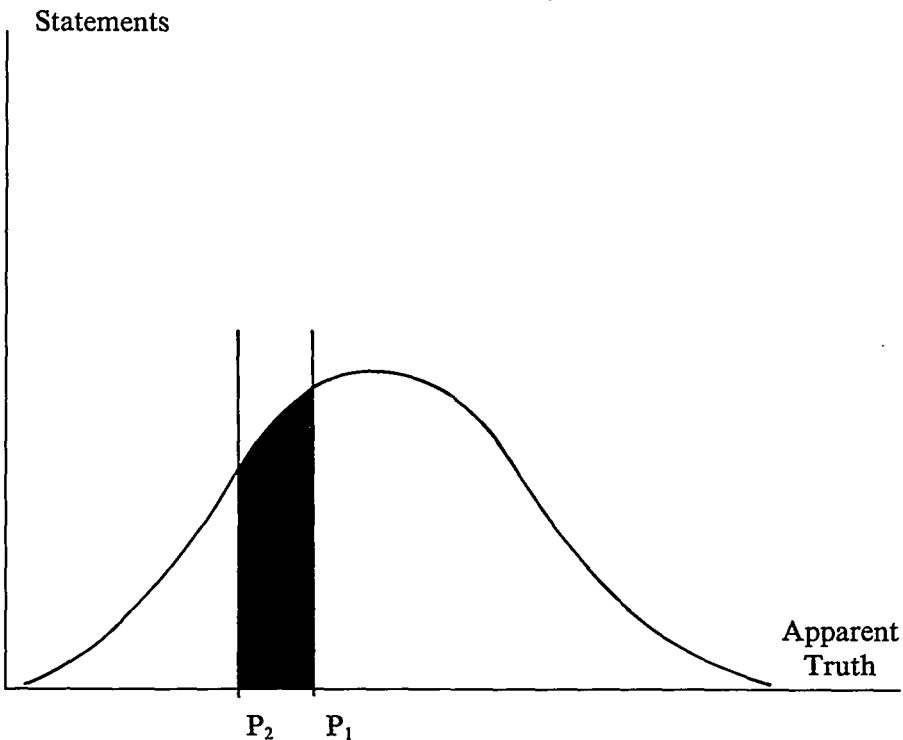


FIGURE 1: Apparent Truth of Defamatory Statements

In Figure 1, the defamations occurring in any time period are assumed to be normally distributed along the Apparent Truth dimension.<sup>69</sup> (The exact shape of the distribution depends on the efficiency of rules of evidence and procedure, and related constraints.) All potentially litigable cases are included in the distribution, but no potential litigant knows

69. The location of the curve will depend in part on the prevailing level of defamatory statements in society. This is treated as exogenous here, but in fact an endogenous link could be specified between the outcomes of defamation cases and the position of the curve.

precisely where his or her case is located. This is because specification of the crucial variable, the appearance of truth, ultimately depends on the perceptions of a court. *A priori*, each potential case is associated with its own probability distribution centered on some point in Figure 1.<sup>70</sup>

If truth were the only defense, the risk of liability for publishing a given defamation would depend on two factors: (1) the standard of proof and (2) the burden of persuasion. In Figure 1, if the standard of proof is set at *P1*, all statements to the left of this point will be judged false, and their publishers will be held liable. If the standard is set at *P2*, many more statements (those falling within the shaded area) can be published with impunity.

The distribution in Figure 1 can be conceptually broken down into component distributions of statements that are "actually" true or false, as in Figure 2. These distributions (labeled *T* and *F*, respectively) will overlap so long as the most credible false statement seems more plausible than the least credible true statement. The extent of the overlap depends largely on the effectiveness of the factfinding process.

Figure 2 highlights the problem of Type I and Type II error. Setting the standard of proof at *P1* rather than at *P2* minimizes the number of false statements that are judged to be true, thereby reducing Type II error (mistaken acquittals). However, it simultaneously increases the number of true statements that are wrongly judged to be false, thereby increasing Type I error (mistaken findings of liability). Setting the standard at *P2* simply reverses the trade-off.

---

70. These distributions depend partly on the amount of resources devoted to litigation. See, e.g., R. POSNER, *supra* note 21, at § 21.8. Therefore, litigants have at least some marginal control over the appearance of their claim, depending on the resources they are able and willing to commit. The problem of uncertain placement is especially acute when a potential defendant does not know prior to litigation whether a published story is true. Cf. Sheer & Zardkoohi, *An Analysis of the Economic Efficiency of the Law of Defamation*, 80 NW. U.L. REV. 364, 412 (1985) (strict liability rule suppresses publication of true statements that are not known to be true at time of publication decision).

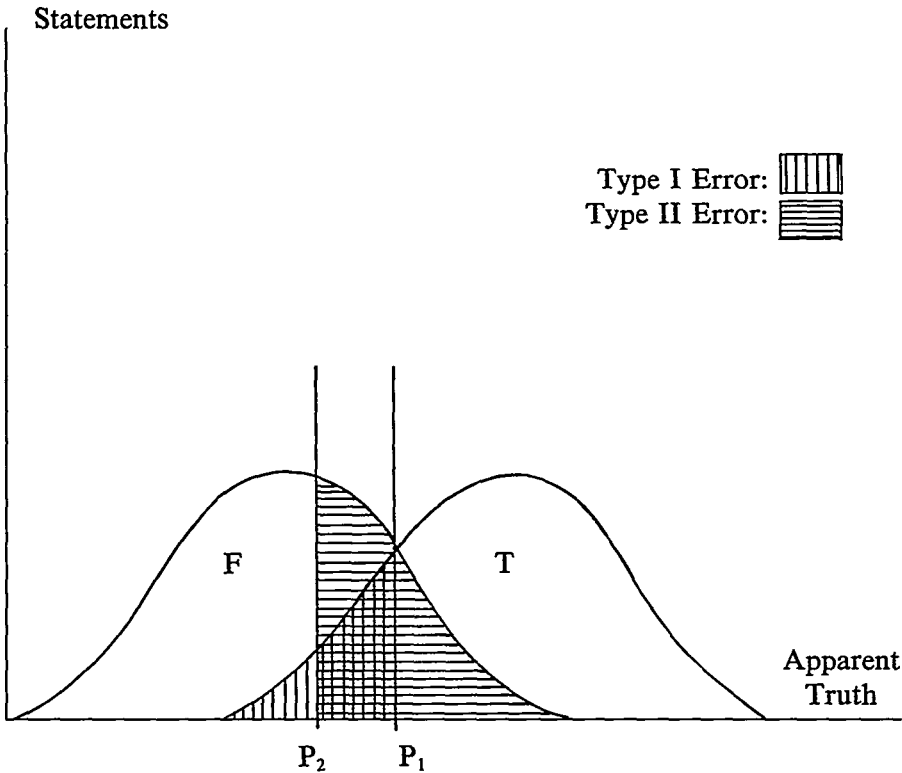


FIGURE 2: Component Distributions by Actual Truth or Falsity

Shifting the burden of proof has a similar effect. If the apparent truth of statement *S* corresponds to the point where the standard of proof is set, liability will depend on which party bears the burden. Statement *S* and the set of all equally credible statements will be judged true if the plaintiff bears the burden of proof, but false if the burden lies with the defendant. This set of statements will include some that are actually true, and some that are false. Shifting the burden of proof between the parties therefore affects the marginal incidence of Type I and Type II error in a way equivalent to changing the standard of proof.<sup>71</sup>

*D. Objections to the Statistical Model*

Nearly a decade ago, in *Ballew v. Georgia*,<sup>72</sup> Justice Blackmun rec-

71. This is, of course, the situation described by Justice O'Connor in *Hepps*. See *supra* text accompanying note 2.

72. 435 U.S. 223 (1978).

ognized the operation of Type I and Type II errors in jury verdicts.<sup>73</sup> This opinion stands as the only one where the Supreme Court has acknowledged the equivalence of legal and statistical error. Courts and commentators generally have been reluctant to view judicial decision-making as a statistical process. Some critics argue that such models are fundamentally inappropriate, regardless of the apparent closeness of the statistical analogy. Others claim that, while the legal system may accurately be described in probabilistic terms, the familiar Pascalian system of mathematical probabilities is unsuited to the task. This Subsection reviews the leading variations of these arguments and suggests that none of them seriously compromises the theoretical value of a statistical model of judicial error.

### 1. The comfortable illusion of certainty

The traditional reluctance to identify legal truth with probabilities may be a holdover from the age of trial by combat or ordeal—procedures by which the truth of competing claims was settled absolutely by divine revelation.<sup>74</sup> Indeed, the tie to divinity may still exist in the minds of some. “When people believed in divine intervention they also believed that true innocence would prevail. In some respects the jury trial still provides a test of innocence in a way that a statistically geared trial cannot. . . . To be believed by a jury seems tantamount to receiving a special moral dispensation.”<sup>75</sup> Whether or not this is so, it is a hallmark of the legal process that once a fact is established at trial as sufficiently probable, it is thereafter treated as if it were absolutely true.<sup>76</sup> For practical purposes, these determinations amount to assertions of objective reality: facts either exist or they do not—statements are either true or false. To think of truth as a matter of degree, as though facts could be partly true and partly false, is alien to this perspective.<sup>77</sup>

A fine psychological line exists between accepting judgments as definitive and actually believing in their certainty. When analysts demand belief rather than merely acceptance, they confuse what the system is capable of generating (acceptable verdicts) with what it is not (verdicts that are certain). Nevertheless, the simple recognition of inevitable error sometimes seems unbearable:

---

73. *Id.* at 234. Only Justice Stevens joined in Justice Blackmun’s opinion of the Court, dealing with optimal jury size. The relevant analysis drew heavily on Nagel & Neef, *supra* note 43.

74. Ball, *supra* note 13, at 810.

75. Zuckerman, *Law, Fact, or Justice?*, 66 B.U.L. REV. 487, 506 (1986).

76. Ball, *supra* note 13, at 808.

77. *Id.* at 808-09.

[I]f we were to set reasonable doubt at ninety-nine percent, are we willing to say that a jury need only be ninety-nine percent sure of its decision, when to do so would be to admit that we would tolerate the conviction of one innocent man out of every hundred convicted?<sup>78</sup>

What is remarkable about this passage is that it describes how the law actually functions every day; what troubles the authors is the possibility that we might *admit* it. Denial of this everyday reality has been vigorously championed by Lawrence Tribe, who suggests that there may be something "intrinsically immoral" about returning a verdict against a defendant while knowing that one might be wrong.<sup>79</sup> Tribe goes on to maintain that the immorality lies not so much in our actual awareness of doubt as in being forced to "proclaim" this awareness.<sup>80</sup>

It is not always clear whether Tribe's position is wholly normative, or if he claims that the law really does not allow convictions when the court is aware of a measurable risk of error.<sup>81</sup> The same ambiguity sometimes arises in similar objections posed by others.<sup>82</sup> What *is* clear is that Tribe's animosity to probabilistic models is based on his view of trials as social ritual. As he ultimately puts it, "[T]he *process*, and not the *result in any particular case*, is all important."<sup>83</sup> Unfortunately, Tribe cuts away his own normative ground with this comment. After all, the worst he can say of his opponents is that their search for precision may lead to erroneous convictions by jurors bedazzled by mathematics.<sup>84</sup> Yet Tribe himself ultimately admits his willingness to accept mistaken verdicts—as the price, not of precision, but of ritual.

Charles Nesson has staked out a position close to Tribe's, based on the view that trials are primarily vehicles for settling disputes, not determining truth.<sup>85</sup> Nesson fears that verdicts will lose their effectiveness if they are understood to be based on degrees of belief in evidence.<sup>86</sup> Nes-

---

78. Broun & Kelly, *supra* note 26, at 31.

79. Tribe, *supra* note 64, at 1372.

80. *Id.* at 1373.

81. *See id.* at 1374; Tribe, *A Further Critique of Mathematical Proof*, 84 HARV. L. REV. 1810, 1818 (1971).

82. *See, e.g.*, Kaplan, *Decision Theory and the Factfinding Process*, 20 STAN. L. REV. 1065, 1073 (1968).

83. Tribe, *supra* note 64, at 1381 (emphasis in original).

84. *Id.*

85. Nesson, *Agent Orange Meets the Blue Bus: Factfinding at the Frontiers of Knowledge*, 66 B.U.L. REV. 521 (1986); Nesson, *The Evidence Or The Event? On Judicial Proof and the Acceptability of Verdicts*, 98 HARV. L. REV. 1357 (1985); Nesson, *Reasonable Doubt and Permissive Inferences: The Value of Complexity*, 92 HARV. L. REV. 1187 (1979).

86. *See* Nesson, *The Evidence or the Event?*, *supra* note 85.



son's argument combines two strands that are related only by assertion. Clearly, even if the main function of trials is to resolve disputes, probabilistic models might still have value. If such tools can increase the accuracy of outcomes without impeding the court's ability to reach decisions (an empirical question), the first element of Nesson's objection loses its force. The remainder of his position collapses into a variant of Tribe's argument: (1) the function of trials is to settle disputes in a socially acceptable way; (2) socially acceptable settlements must include ritualistic claims of infallibility; (3) therefore probabilistic models are never appropriate, even if they enhance accuracy at no cost to facilitating settlements. The main flaw in this argument lies in its empirically doubtful second premise—that popular acceptance demands infallibility. Although this claim is also a cornerstone of Tribe's argument,<sup>87</sup> neither author has shown that it is true—nor why, if true, it should be accepted as inviolate by those who seek to improve the system.<sup>88</sup> In particular, the premise that acceptance rests on infallibility seems no more plausible than the alternative that “[t]he legitimacy of our judicial system . . . depends on the popular assumption that the system produces decisions based on accurate factfinding.”<sup>89</sup>

Nesson's latest work<sup>90</sup> so blurs the contours of his argument that it seems in danger of becoming a semantic game. “At some point,” he acknowledges, “high probability alone is sufficient to produce an acceptable verdict.”<sup>91</sup> But this admission is tempered by his use of “probable” to mean an outcome that “best accomplishes a just and acceptable resolution of the dispute.”<sup>92</sup>

It is hard to avoid the impression that arguments from ritual and acceptability reduce to little more than comfortable self-deception. Probabilistic models hold the risk of error before our eyes; rejecting the model offers a pretense for claiming that errors will not occur.<sup>93</sup> The best response to this line of reasoning has already been given:

It may soothe the conscience of everyone if courtroom controversies were decided and pronouncements made only on the basis of what is held to be “actual belief” in their wisdom. This

---

87. Tribe, *supra* note 64, at 1372-75.

88. See Brook, *The Use of Statistical Evidence of Identification in Civil Litigation: Well-Worn Hypotheticals, Real Cases, and Controversy*, 29 ST. LOUIS U.L.J. 293, 335-38 (1985).

89. Gold, *Jury Wobble: Judicial Tolerance of Jury Inferential Error*, 59 S. CAL. L. REV. 391, 405 (1986).

90. Nesson, *Agent Orange*, *supra* note 85.

91. *Id.* at 522 n.3.

92. *Id.* at 521.

93. See Williams, *supra* note 39, at 304.

practice certainly appears more appropriate than that of relying on the best guess in even the most important situations. But . . . the conscientious juror is enmeshed in a difficult position from which he has no easy escape. Uncertainty makes educated guessers of everyone.<sup>94</sup>

Understandably, jurors (and judges) would prefer to believe in the certainty of their decisions because it is unpleasant to consider the consequences of error. However, these unpleasant consequences—from committing a guiltless person to death to pushing a business into bankruptcy through a successful tort fraud—are an essential fact of the legal system. Minimizing the cost of errors will be impossible unless we are willing to admit that they exist.

## 2. Confused jurors or confused analysis?

The inability of jurors to understand complex mathematical standards is sometimes cited as an objection to probabilistic legal models.<sup>95</sup> However, one may wonder whether the average juror is as awed and mystified by statistics as these scholars contend.<sup>96</sup> Eggleston cites a curious case in which an Australian jury specifically asked a judge to define the level of probability corresponding to the reasonable doubt standard; the judge refused to comply.<sup>97</sup> Whatever lesson may be drawn from this, Eggleston is simply wrong to infer that “it is impossible to specify any particular mathematical level of probability which must be achieved before a verdict of guilty can be returned in a criminal case.”<sup>98</sup>

More important, this argument fails to distinguish between the use of statistical evidence and statistical models of the decision-making process.<sup>99</sup> One can accept the probabilistic nature of proof for purposes of presenting certain kinds of evidence while rejecting the view that legal procedures can (or should) be interpreted in probabilistic terms—or vice versa.<sup>100</sup> Even if all purely statistical evidence were barred from the

---

94. Brook, *supra* note 30, at 92.

95. See Broun & Kelly, *supra* note 26, at 31; Tribe, *supra* note 64.

96. For example, the California Supreme Court may be accused of hyperbole in describing mathematics as “a veritable sorcerer in our computerized society,” that stands ready to “cast a spell over” triers of fact. *People v. Collins*, 68 Cal. 2d 319, 320, 438 P.2d 33, 33, 66 Cal. Rptr. 497, 497 (1968).

97. R. EGGLESTON, *supra* note 32, at 114.

98. *Id.*

99. See Schmalbeck, *The Trouble With Statistical Evidence*, 49 LAW & CONTEMP. PROBS. 221, 231 (1986); Stripinis, *Probability Theory and Circumstantial Evidence: Implications from a Mathematical Analysis*, 21 JURIMETRICS J. 59, 75 (1981).

100. See, e.g., Tribe, *supra* note 64, at 1330-31; Leubsdorf, *supra* note 17, at 148; Broun & Kelly, *supra* note 26, at 27.

courtroom, to avoid confusion or for any other reason, a probabilistic decision model might still capture the essence of the judicial process.<sup>101</sup>

### 3. Questions of interpretation

Probability theory is not a monolithic system of thought. Several distinct modes of probability analysis exist, and to some extent they are incompatible and even mutually antagonistic.<sup>102</sup> The most familiar interpretation of probability is based on relative frequencies, i.e., on repeated observations of identical or very similar events. In this sense, a valid statistical claim that a ten percent probability exists of outcome *X* means that *X* has been observed to occur in ten percent of all cases of the type under review.<sup>103</sup> "If there have been a thousand days with conditions similar to today's and on one hundred of those days it has rained, the probability of rain [today] is the fraction 100/1000, or ten percent."<sup>104</sup> In response, Tribe demands to know what it means to be "four-fifths certain" in a given case.<sup>105</sup> Although this criticism warps the meaning of statistical confidence,<sup>106</sup> it makes the point that probabilities, to be applied at law, must be converted to specific propositions concerning particular cases.<sup>107</sup>

Indeed, one alternative to the frequency-based interpretation is to regard probabilities as expressing the evaluator's subjective confidence that a fact is true. From this perspective, courts can be viewed as forming subjective estimates of the probabilities associated with various possible sets of facts, and basing their decisions on these estimates.<sup>108</sup> In answer to Tribe's "four-fifths certain" jibe,<sup>109</sup> Kaye would liken subjec-

101. This Article does not take up the question of whether any meaningful distinction exists between "probabilistic" and "non-probabilistic" evidence. See, e.g., Kaye, *The Admissibility of "Probability Evidence" in Criminal Trials—Part II*, 27 JURIMETRICS J. 160, 172 (1987); Allen, *A Reconceptualization of Civil Trials*, 66 B.U.L. REV. 401, 414-15 (1986); Leubsdorf, *supra* note 17, at 148.

102. See N. GEORGESCUE-ROEGEN, *supra* note 12, at 52-59.

103. Ball, *supra* note 13, at 810-11.

104. Broun & Kelly, *supra* note 26, at 23-24.

105. Tribe, *supra* note 64, at 1347.

106. "[A] finding that is statistically significant at the .05 level . . . is not an indication that the conclusion is 95 percent correct. Instead, the conclusion is either correct or incorrect. However, if this significance level is attained, the odds are 95 percent that the conclusion is correct." Channels, *The Methods of Social Science and Their Use in Legal Proceedings*, 16 CONN. L. REV. 853, 870 (1984). See also the discussion of confidence levels in Schmalbeck, *supra* note 99, at 225-28.

107. Aickin, *Issues and Methods in Employment Discrimination Statistics*, 26 JURIMETRICS J. 347, 350 (1986).

108. Kaye, *supra* note 16, at 641.

109. See *supra* text accompanying note 105.

tive probabilities to betting odds—estimates based on expected frequencies, tempered by the weight of evidence.<sup>110</sup> “Where subjective probabilities are defined in terms of the odds a person would be willing to accept in betting on the outcome of an event, they can be shown to obey all the rules of probability theory, at least where the judgments satisfy certain conditions of consistency.”<sup>111</sup> This response will not satisfy everyone. Subjective probability theory has been described as “a dangerously inappropriate paradigm for the courts.”<sup>112</sup> Yet as a descriptive model, this interpretation seems to capture much of what courts actually do when deciding the likelihood of facts argued before them.

Brilmayer and Kornhauser raise the objection that the use of statistical methods may endanger the treatment of litigants as unique individuals.<sup>113</sup> Indisputably, criminal convictions and civil liability must be based on the courts’ belief that particular defendants violated particular standards. It can never be enough to say that, based on a statistical analysis, the defendant is the sort of person who is likely to commit this sort of offense. Still, even the most individualized judgment can be redefined as a statement about probabilities.<sup>114</sup> The danger of abuse is apparent,<sup>115</sup> but this is not in itself an indictment of statistical thinking. The crucial issue is whether abuse can be avoided, keeping the system properly focused on the issue of individual responsibility.

L. Jonathon Cohen has advanced perhaps the most persistent argument that standards of proof cannot even in principle be translated into mathematical probabilities.<sup>116</sup> Instead, Cohen has proposed his own theory of “inductive” (i.e., intuitive) probability as more consistent with the meaning of legal standards and the evaluative methods actually used in

110. Kaye, for example, has shown that subjective probability estimates obey all the classical mathematical theories and postulates. Kaye, *supra* note 66, at 42; Kaye, *supra* note 30, at 610; see also L. SAVAGE, *THE FOUNDATIONS OF STATISTICS* (1954); Shafer, *The Construction of Probability Arguments*, 66 B.U.L. REV. 799 (1986).

111. Kaye, *supra* note 66, at 43; see also Kaye, *supra* note 30, at 609; Finkelstein & Fairley, *A Bayesian Approach to Identification Evidence*, 83 HARV. L. REV. 489, 504 (1970).

112. Cohen, *Subjective Probability and the Paradox of the Gatecrasher*, 1981 ARIZ. ST. L.J. 627, 632.

113. Brilmayer & Kornhauser, *Review: Quantitative Methods and Legal Decisions*, 46 U. CHI. L. REV. 116, 149-52 (1978).

114. This is most obvious in cases that do not require bipolar judgments—that is, when the court must choose a value from a continuous liability function. See Thorne, *Mathematics, Fuzzy Negligence, and the Logic of Res Ipsa Loquitur*, 22 JURIMETRICS J. 92, 109-11 (1981).

115. The classic contemporary horror story of statistical abuse at trial is *People v. Collins*, 68 Cal. 2d 319, 438 P.2d 33, 66 Cal. Rptr. 497 (1968). The role of spurious probability estimates in the *Dreyfus* case (1899) should also be remembered. See generally Landsman, *Book Review*, 85 MICH. L. REV. 1095 (1987).

116. L. COHEN, *THE PROBABLE AND THE PROVABLE* 17-39 (1977).

trials.<sup>117</sup> It is undoubtedly true that courts must apply intuitive estimates of relevant probabilities; in most cases they have no alternative. The issue is whether the courts should gauge their intuitive estimates against equally subjective standards, or against some mathematical benchmark.<sup>118</sup> Uncertainty is not the same thing as imprecision, nor does the existence of uncertainty excuse imprecise thinking. A few scholars, for example, have explored the usefulness of "fuzzy" mathematics for modeling the judicial process.<sup>119</sup> This technique accommodates the uncertainty of everyday experience and language within a precise system of mathematical logic. One characteristic of fuzziness is that values are usually expressed as ranges instead of points. Thus, the percentage ranges judges assign to concepts like "beyond a reasonable doubt"<sup>120</sup> could be readily integrated into a fuzzy probability model.

Finally, it has been said that courts differ from social scientists in that courts may not withhold judgment on disputed issues.<sup>121</sup> A scientist is never required to declare the null hypothesis true if a correlation tests out at less than the .95 confidence level; "[c]ourts, by contrast, typically must decide between two competing alternatives."<sup>122</sup> But how viable is this distinction? In most civil and virtually all criminal trials, one party is *ex ante* in possession of some contested good. If the plaintiff or the state meets its burden of proof, the court will transfer a claim to the defendant's property or person. If the burden is not met, the court will refuse to impose the transfer. The latter outcome, however, does not necessarily imply an endorsement of the defendant's position. It can be

---

117. *See id.* at 40-98.

118. "Surely, a logical methodology in support of quantitative probability measures of significance is preferable to the subjective probability estimates of layman jurors." Smith & Charrow, *Upper and Lower Bounds for Probability of Guilt Based on Circumstantial Evidence*, 70 J. AM. STATISTICAL ASSOC. 555, 559 (1975).

Although Professor Cohen has promoted his "Baconian" theory for nearly a decade (for his latest offering see Cohen, *The Role of Evidential Weight in Criminal Proof*, 66 B.U.L. REV. 635 (1986)), it has had relatively little impact on American legal scholarship. For recent thoughtful reevaluations of Cohen's views, see Tillers, *Mapping Inferential Domains*, 66 B.U.L. REV. 883 (1986); Allen, *supra* note 101, at 415-25. More critical appraisals have been leveled by Fienberg, *Misunderstanding, Beyond a Reasonable Doubt*, 66 B.U.L. REV. 651 (1986) and Kaye, *Do We Need a Calculus of Weight to Understand Proof Beyond a Reasonable Doubt?*, 66 B.U.L. REV. 657 (1986). David Schum has suggested that Cohen's approach to probabilities may be superior at certain stages of the judicial process, while Pascalian, Bayesian and other probability systems are better suited for other stages. Schum, *Probability and the Processes of Discovery, Proof, and Choice*, 66 B.U.L. REV. 825 (1986).

119. *E.g.*, Thorne, *supra* note 114.

120. *See* Simon, *supra* note 40.

121. *E.g.*, Lempert, *Statistics in the Courtroom: Building on Rubinfeld*, 85 COLUM. L. REV. 1098 (1985).

122. *Id.* at 1100.

no more than an acknowledgment that the case was not well enough presented to warrant disturbing the status quo. Judgments for defendants are thus closely analogous to a researcher's failure to reject a null hypothesis.

#### 4. Systematic inferential error

Modern cognitive research has shown that systematic biases are often present in subjective probability estimates by laymen and experts alike.<sup>123</sup> Properly developed, these findings might raise grave doubts as to the capacity of judges and juries to do the job the system assigns to them.<sup>124</sup> One area of special concern, from the standpoint of probabilistic models, is evidence that subjects tend to ignore or underutilize base rate data in estimating the likelihood of specific events.<sup>125</sup> Moreover, experimental studies indicate that subjective probability estimates tend to be "conservative"—significantly underestimating the true likelihood of events that are very highly probable—and are only reluctantly revised toward better estimates.<sup>126</sup>

The adversary system ultimately relies on the diligence of counsel to bring out information and minimize error. However, advocates have powerful incentives to *induce* inferential errors in their favor, and gener-

---

123. See Tversky & Kahneman, *Judgment Under Uncertainty: Heuristics and Biases*, 185 SCIENCE 1124 (1974); see also, Kahneman & Tversky, *Subjective Probability: A Judgment of Representativeness*, 3 COGNITIVE PSYCHOLOGY 430 (1972); Tversky & Kahneman, *Belief in the Law of Small Numbers*, 76 PSYCHOLOGICAL BULL. 105 (1971). The first major crossover work applying inferential error theory to the law was R. NISBETT & L. ROSS, *HUMAN INFERENCE: STRATEGIES AND SHORTCOMINGS OF SOCIAL JUDGMENT* (1980). (Some of the underlying research had previously been discussed in Shoben, Book Review, 59 B.U.L. REV. 582 (1979)). The Nisbett & Ross book was introduced to the legal literature by Spitzer, Book Review, 9 HOFSTRA L. REV. 1621 (1981), and Loftus & Beach, *Human Inference and Judgment: Is the Glass Half Empty or Half Full?*, 34 STAN. L. REV. 939 (1982). Victor Gold has waged an unjustifiably lonely campaign to apply this research to the interpretation of evidentiary rules. Gold, *supra* note 89; Gold, *Federal Rule of Evidence 403: Observations on the Nature of Unfairly Prejudicial Evidence*, 58 WASH. L. REV. 497 (1983).

124. "If people are, in general, prone to commit inferential error by relying on judgmental heuristics, there is no reason to believe that such erring will suddenly cease once people become jurors." Spitzer, *supra* note 123, at 1627. Procedural and other reforms to mitigate the effects of cognitive bias, although arguably premature, have been proposed by, for example, Spitzer, *supra* note 123, at 1634-35; Loftus & Beach, *supra* note 123, at 949-50; and Saks & Kidd, *Human Information Processing and Adjudication: Trial by Heuristics*, 15 LAW & SOC'Y REV. 123 (1981).

125. See R. NISBETT & L. ROSS, *supra* note 123, at 141-50; Tversky & Kahneman, *Evidential Impact of Base Rates*, in *JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES* 19 (D. Kahneman, P. Slovic & A. Tversky, eds., 1982); Saks & Kidd, *supra* note 124.

126. See, e.g., Fairley, *Probabilistic Analysis of Identification Evidence*, 2 J. LEGAL STUD. 493, 494-95 (1973); Edwards, *Conservatism In Human Information Processing*, in *FORMAL REPRESENTATIONS OF HUMAN JUDGMENT* 17 (B. Kleinmuntz ed., 1968).

ally lack the training to counter such errors generated by their opponents.<sup>127</sup> Do attorneys routinely (if unconsciously) exploit jurors' propensity to err, and if so, do existing rules and procedures implicitly recognize and offset the practice? These are important empirical questions that have not been adequately tested.

If inferential bias is as universal as the literature suggests, it may actually be less damaging to the statistical paradigm than to alternative decision models. The essence of the problem is that factfinders seem to systematically err in interpreting evidence. If true, this phenomenon can in principle be taken into account by a model that incorporates the known risk of judicial error. In contrast, models based on the myth of certainty are unable to accommodate the effects of inferential error.

### III. THE ERROR FRONTIER AND ATTAINABLE ACCURACY

#### A. *Error in the Law: A Reinterpretation*

Any factfinding system will generate two kinds of outcomes: some that are correct and some that are in error. "Correct" judgments can be defined as those that would have come out the same if all relevant facts were known with certainty. In contrast, "errors" are outcomes that differ from those that would have been reached under conditions of certainty. Erroneous outcomes can be divided into Type I and Type II errors, depending on whether they go against the defendant or the plaintiff, respectively. These simple relationships can be expressed algebraically as:

$$K = C + E \quad (1)$$

and

$$E = \gamma D + \delta P. \quad (2)$$

*K*: the total number of cases decided in some time period *t*;

*C*: the number of correct decisions;

*E*: the number of erroneous decisions;

*D*: the number of cases in which judgment would be for the defendant under conditions of certainty;

*P*: the number of cases in which judgment would be for the plaintiff under conditions of certainty;

$\gamma$ : the probability of error when the defendant is in fact innocent;

$\delta$ : the probability of error when the defendant is in fact guilty or liable.

---

127. Gold, *Jury Wobble*, *supra* note 89, at 399.

The mix of Type I and Type II errors comprising  $E$  depends on the proportion of litigated cases in which the defendant is actually innocent and on the relative values of the coefficients  $\Upsilon$ <sup>128</sup> and  $\delta$ .<sup>129</sup> In the extreme case where no innocent defendant is tried, all errors would be Type II errors, and  $E = \delta P$ . At the other extreme, if all defendants were actually innocent, there would be only Type I errors, and  $E = \Upsilon D$ . Holding  $K$  constant, these two extremes can be graphed as the vertical and horizontal axes of a decision space (see Figure 3a). The various possible combinations of cases,

$$K[i] = D[i] + P[i] \quad (3)$$

can be represented by points along line  $\underline{K}$  (see Figure 3b).

128.  $\Upsilon$  is a function of  $\alpha$ , *see supra* note 60 and accompanying text, but is not identical to  $\alpha$ . Since  $\alpha$  measures the proportion of all innocent potential defendants who would be mistakenly held liable,  $\Upsilon$  would be identical to  $\alpha$  only if all such cases (or a representative sample) were litigated to a decision. This is unlikely for several reasons. In some cases, the potential plaintiff's litigation costs would exceed the expected value of a favorable decision. Moreover, many disputes are settled without litigation, or at least before they are litigated to a decision. Changes in the cost of litigation or in the efficiency of settlement may affect  $\Upsilon$ , but not  $\alpha$ . *See G. TULLOCK, supra* note 49, at 28.

Notice also that the placement of the burden of proof would cause some marginal divergence of  $\Upsilon$  from the "pure" expectation given by  $\alpha$ , even if all cases were litigated. *See supra* text accompanying notes 50-53.

129. The relationship of  $\delta$  to  $\beta$  is analogous to that of  $\Upsilon$  to  $\alpha$ . *See supra* note 128.



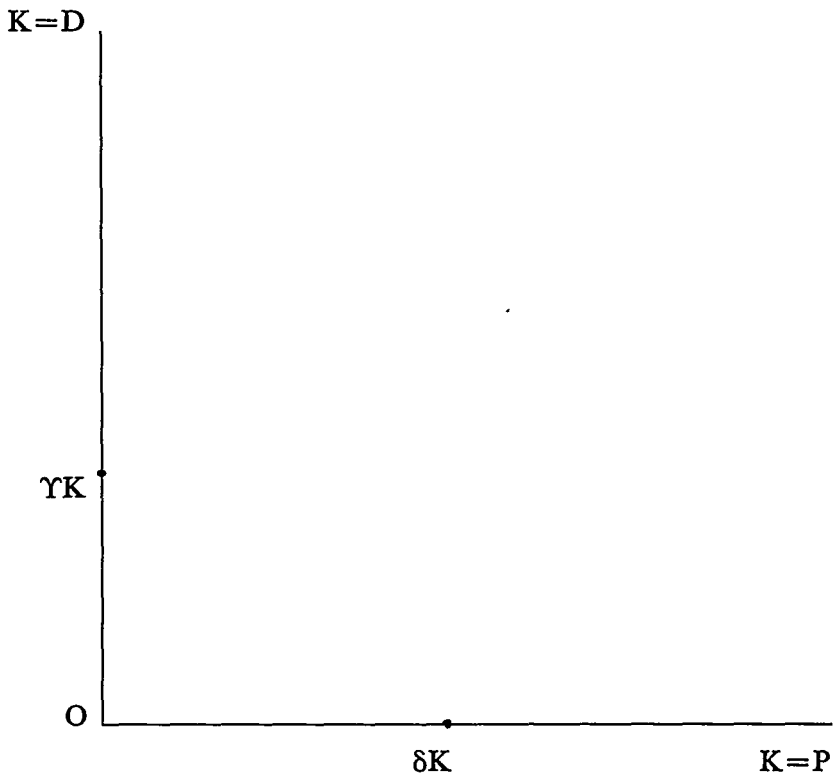


FIGURE 3a

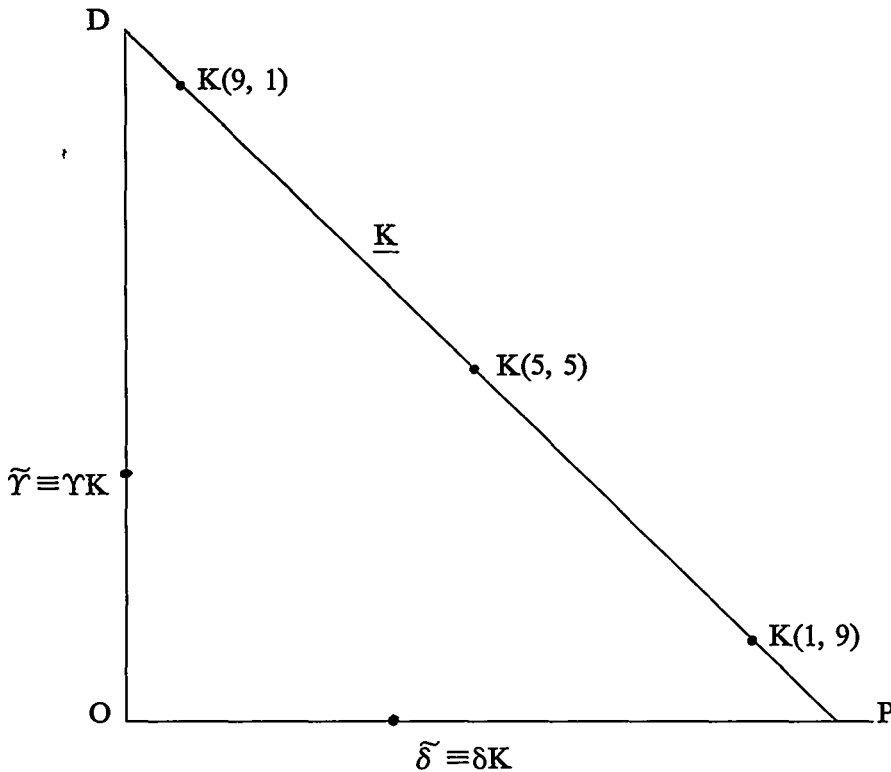


FIGURE 3b

Each of these combinations would entail a given degree of error. For each possible combination  $K[i]$ , the corresponding expected error ( $E[i] = \gamma D[i] + \delta P[i]$ ) can be located on a ray from the origin (see Figure 3c). The set of all possible points  $E[i]$  graphically defines an “error frontier,” shown as the heavy line  $\underline{E}$  in Figure 3d.<sup>130</sup> The shaded area below  $\underline{E}$  represents the minimum overall risk of error in the system.<sup>131</sup> Reducing this risk (or increasing the probability of correct outcomes) would require an accuracy that is not attainable under prevailing institutional,

130. Somewhat similar graphic devices are used by Heiner, *supra* note 11, at 227, and Tribe, *supra* note 64, at 1386-87. Heiner’s reliability curves, based on signal detection theory, show the probability of error as a function of the frequency of decisions using given information. Tribe’s preference contours and “empirical” possibility curves give the proportion of mistaken convictions (i.e.,  $\gamma$ ) as a function of the proportion of correct convictions (i.e.,  $1 - \delta$ ). Tribe varies procedural rules and other constraints along a given curve, while these are held constant in the construction of the error frontier.

131. Total Risk:

$$R = \sum_{i=0}^K (\gamma D[i] + \delta P[K-i]).$$

informational and behavioral constraints.<sup>132</sup> Therefore, the frontier can also be considered the boundary of the feasible set of correct outcomes (the area above  $\underline{E}$ ).

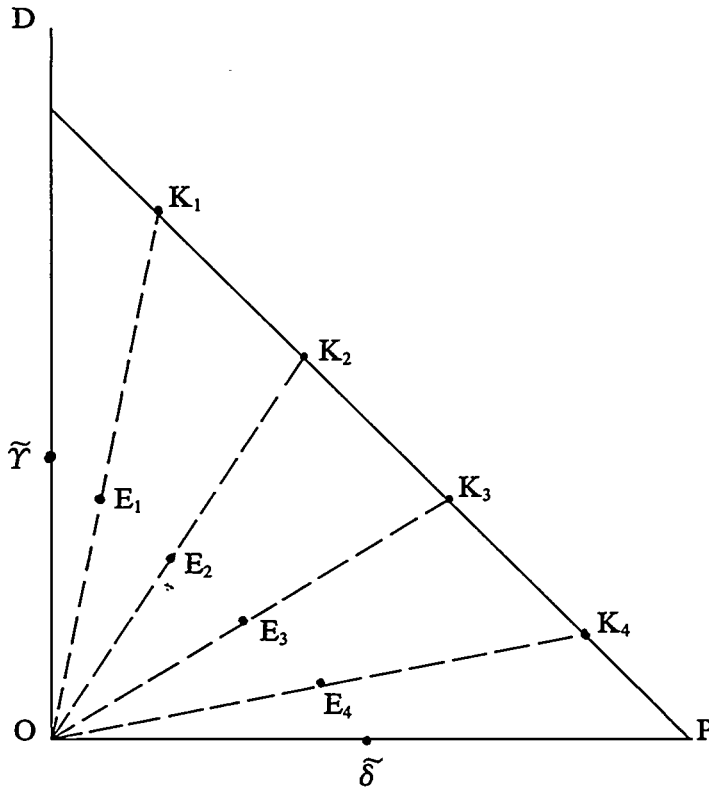


FIGURE 3c

132. Reduced information costs, for example, should improve the overall accuracy of decisions, lowering both  $\gamma$  and  $\delta$  and shifting the error frontier back toward the origin. This result could follow from increasing the quantity and quality of evidence produced at trial, or by clarifying the standards of liability. Epstein makes the latter point in calling for simultaneous reduction of both types of error in judicial review of economic regulation. Epstein, *supra* note 61. Cf. Ehrlich & Posner, *An Economic Analysis of Legal Rulemaking*, 3 J. LEGAL STUD. 257, 264 (1974); Friedman, *supra* note 65, at 23. But cf. Priest, *The Common Law Process and the Selection of Efficient Rules*, 6 J. LEGAL STUD. 65, 68-69 (1977) (arguing that the efficiency of legal rules is not necessarily associated with their clarity).

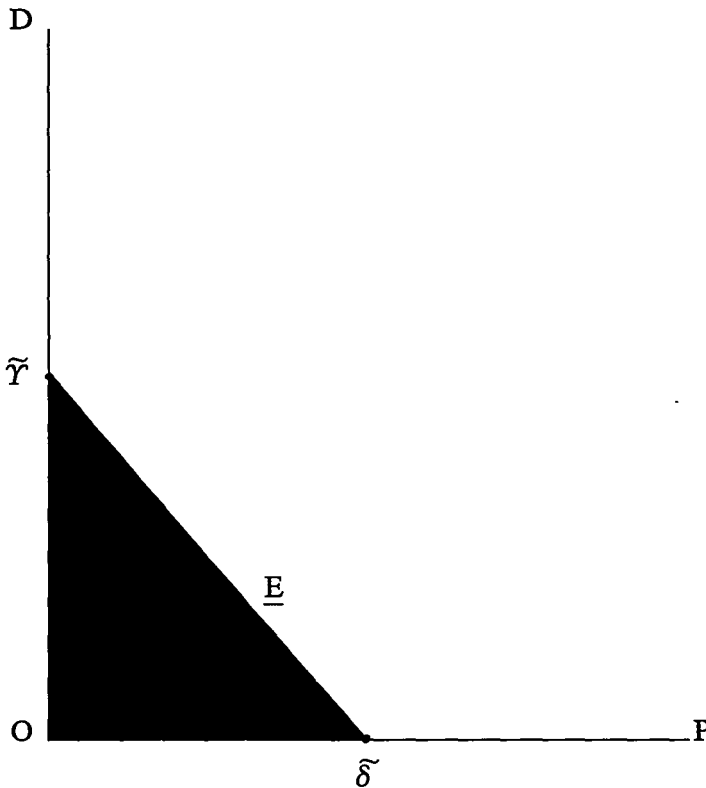


FIGURE 3d: The Error Frontier

The error frontier is drawn here as a straight line, but it may in fact be convex, concave or irregular. A linear  $\underline{E}$  implies that  $\tilde{Y}$  and  $\tilde{\delta}$  do not vary with the relative proportion of cases that are litigated. The reasonableness of this assumption depends on the parties' attitudes toward risk and the efficiency of the settlement process. If cases came to trial randomly, the expected risk of error would be the same for all  $D[i]$  and  $P[i]$ . But in reality, cases are selected for litigation only after passing through a filter of pre-trial negotiations. A vigorous body of literature on the theory of legal settlement exists,<sup>133</sup> but little of it directly addresses the

133. See, e.g., Ordoover & Rubinstein, *A Sequential Concession Game with Asymmetric Information*, 51 Q. J. ECON. 879 (1986); Rosenberg & Shavell, *A Model In Which Suits Are Brought For Their Nuisance Value*, 5 INT'L REV. L. & ECON. 3 (1985); Bebchuck, *Litigation and Settlement Under Imperfect Competition*, 15 RAND J. ECON. 404 (1984); P'ng, *Strategic Behavior in Suit, Settlement, and Trial*, 14 BELL J. ECON. 539 (1983); Priest, *Regulating the Content and Volume of Litigation: An Economic Analysis*, 1 SUP. CT. ECON. REV. 163 (1982); Shavell, *Suit, Settlement, and Trial: A Theoretical Analysis Under Alternative Methods for the Allocation of Legal Costs*, 11 J. LEGAL STUD. 55 (1982); Cooter & Marks with Mnookin, *Bargaining in the*

relationship between the decision to litigate and the risk of judicial error.<sup>134</sup> This question will not be pursued further here, since the analysis in this Article will hold for all symmetrical frontiers.

### B. *Burdens and Standards of Proof*

The placement of the burden of proof always has some effect on the relative risks of Type I and Type II error in the legal system.<sup>135</sup> Shifting the burden between the parties will trade off one type of risk for the other, but will have no direct effect on the total risk of error.<sup>136</sup>

---

*Shadow of the Law: A Testable Model of Strategic Behavior*, 11 J. LEGAL STUD. 225, 243-44 (1982); Ordoover, *On the Consequences of Costly Litigation In the Model of Single Activity Accidents: Some New Results*, 10 J. LEGAL STUD. 269 (1981).

134. The strongest case for convexity follows from Priest & Klein, *The Selection of Disputes for Litigation*, 13 J. LEGAL STUD. 1, 14-15 (1984) (arguing that litigated cases will not be a representative sample of disputes). In their model, the composition of cases shifts to maintain an approximately equal distribution of outcomes. *Id.* at 4-5. Cases closest to the decision standard (thus involving the greatest likelihood of error) are most likely to be litigated. *Id.* at 14-15. Equal probability of outcomes is disputed by Wittman, *Is the Selection of Cases for Trial Biased?*, 14 J. LEGAL STUD. 185 (1985), but his analysis also arguably leads to a convex error frontier. In contrast, Posner argues that the highest-stakes cases are the most likely to be litigated, that parties spend the most on such cases and thereby place the most data in evidence, and consequently that these cases are most likely to be decided correctly. R. POSNER, *supra* note 21, at §§ 21.5, 21.8. This would imply a concave error frontier. The same result is reached from the opposite direction by D'Amato, *Legal Uncertainty*, 71 CAL. L. REV. 1, 16-17 (1983) (holding that the most uncertain claims are likeliest to be settled because of the high cost of litigating them). Finally, Tullock contends that cases with both very high and very low risks of error will be settled, and only moderately risky claims will be litigated. G. TULLOCK, *supra* note 49, at 24-29. Such a model would be consistent with a linear frontier.

Most of the settlement literature focuses on selection of cases for litigation without clearly distinguishing the risk of error. Priest's analysis suggests the most difficult cases are the most likely to be litigated to a decision, since they would involve the greatest difference between plaintiffs' and defendants' expected outcomes. Priest, *supra* note 133. But Rizzo shows this does not follow, given rational expectations and risk neutrality, or when the parties have access to different information. Rizzo, *Can There Be a Principle of Explanation In Common Law Decisions? A Comment On Priest*, 9 J. LEGAL STUD. 423 (1980). Others have maintained that the selection of cases is biased in favor of decisions for plaintiffs, since attorneys working on contingent fees will reject or settle cases where the chances of success are small. See Orloff & Stedinger, *supra* note 68. It has also been argued that the distribution of trial outcomes is biased in favor of the party with the greater economic stake in the decision or in the outcome of future cases of the same type. See Goodman, *An Economic Theory of the Evolution of Common Law*, 7 J. LEGAL STUD. 393 (1978); Rubin, *Why Is the Common Law Efficient?*, 6 J. LEGAL STUD. 51, 55-56 (1977). But it is problematic whether such outcome biases would have a systematic effect on the shape of the error frontier.

135. See *supra* text accompanying notes 50-53.

136. Realigning the relative risks of Type I and Type II error may, however, induce changes in the volume and composition of cases coming to trial in subsequent time periods. See, e.g., Salop & White, *Economic Analysis of Private Antitrust Litigation*, 74 GEO. L. J. 1001, 1055 (1986); Craswell & Calfee, *supra* note 18, at 279; see also R. POSNER, *supra* note 21, § 21.5; G. TULLOCK, *supra* note 49, at 28, 76; Priest, *Reexamining the Selection Hypothesis: Learning*

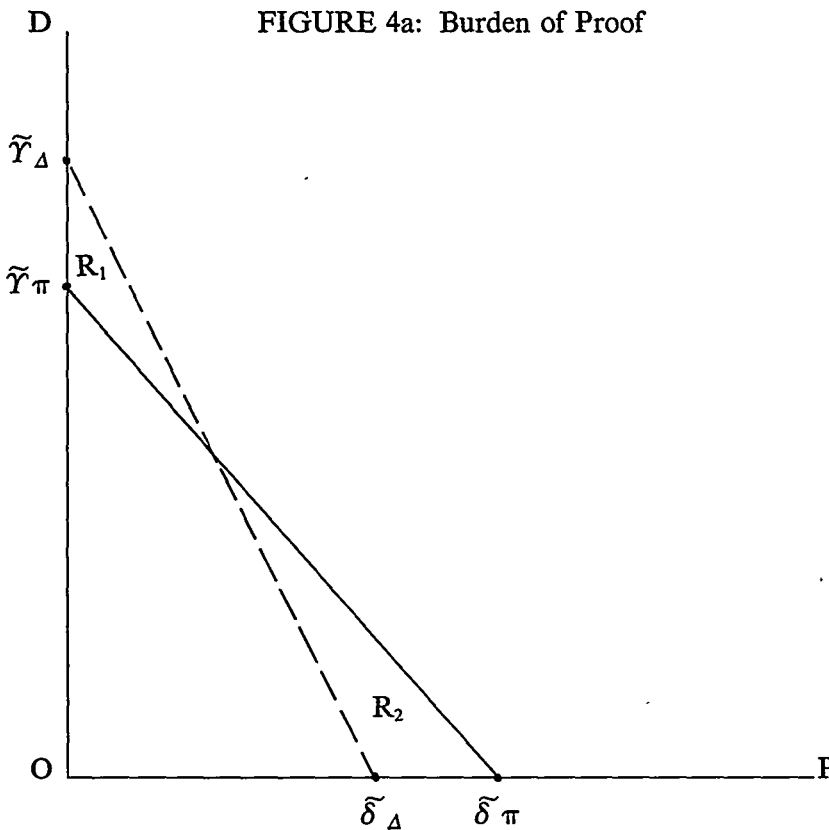
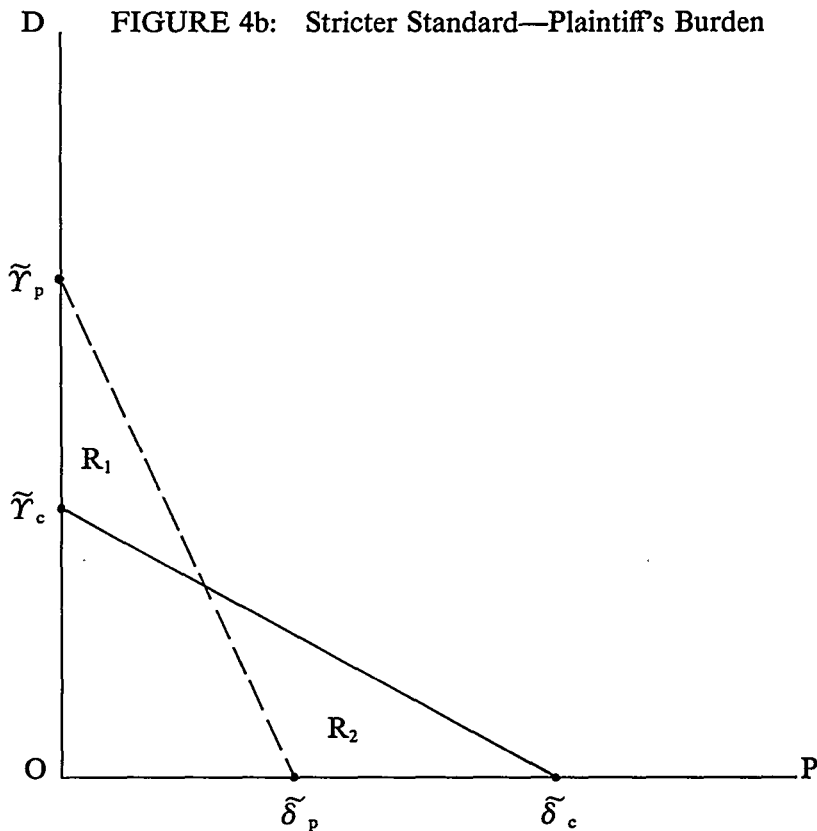


Figure 4a illustrates the effect of shifting the burden of persuasion from defendants to plaintiffs. Initially, the error frontier is the broken line  $\tilde{\gamma}_\Delta \tilde{\delta}_\Delta$ ; the new frontier is the solid line  $\tilde{\gamma}_\pi \tilde{\delta}_\pi$ . Shifting the burden reduces the risk of Type I error,<sup>137</sup> but increases the risk of Type II error by an equivalent amount.<sup>138</sup> Overall risk of error is diminished by the area of triangle  $R_1$ , but increased by area  $R_2$ ; the total area beneath the error frontier remains constant. Shifting the burden from plaintiffs to defendants follows the same analysis in reverse. No matter who wins the ties, the total risk of error is unchanged.

from Wittman's Mistakes, 14 J. LEGAL STUD. 215, 221 (1985); Rubinfeld, *Econometrics in the Courtroom*, 85 COLUM. L. REV. 1048 (1985). The effect is somewhat analogous to shifting the liability for court costs and litigation fees. See Posner, *supra* note 24, at 428; Cooter & Marks with Mnookin, *supra* note 133, at 244-45; cf. Ordoover, *supra* note 133, at 277-79; Priest, *supra* note 133. These indirect effects cannot be modeled without incorporating a settlement theory, which is beyond the scope of this Article.

137. Type I risk is reduced by  $(\gamma_1 - \gamma_2) \sum_i D[i]$ .

138. Type II risk is reduced by  $(\delta_1 - \delta_2) \sum_i P[i]$ .



In contrast, changes in the standard of proof will directly affect the total risk of error. Error is minimized by the preponderance standard; requiring more stringent proof inevitably increases Type II error by more than it reduces Type I error.<sup>139</sup> Figure 4b shows the effect of changing from the preponderance standard ( $\tilde{Y}_p \tilde{\delta}_p$ ) to the clear-and-convincing standard ( $\tilde{Y}_c \tilde{\delta}_c$ ) when the burden of proof is on the plaintiff. Requiring plaintiffs to prove their case by a clear and convincing margin decreases the chance that blameless defendants will be mistakenly held liable. But this reduction is not as great as the increased risk that guilty defendants

139. See, e.g., Ball, *supra* note 13, at 822-23; Brook, *supra* note 30, at 86; Kaye, Book Review, *supra* note 30, at 605 n.19; Cullison, *Probability Analysis of Judicial Fact-Finding: A Preliminary Outline of the Subjective Approach*, 1969 TOLEDO L. REV. 539, 569; Lempert, *supra* note 121, at 1133-34; Finkelstein & Fairley, *supra* note 111, at 508. Although the preponderance standard minimizes total errors, it does not necessarily distribute errors equally between plaintiffs and defendants. See, e.g., Brook, *supra* note 30, at 106-08. Nor does it necessarily minimize the number of very costly errors. See Orloff & Stedinger, *supra* note 68.

will avoid liability.<sup>140</sup> Graphically, total error is increased to the extent that area *R2* exceeds *R1*.

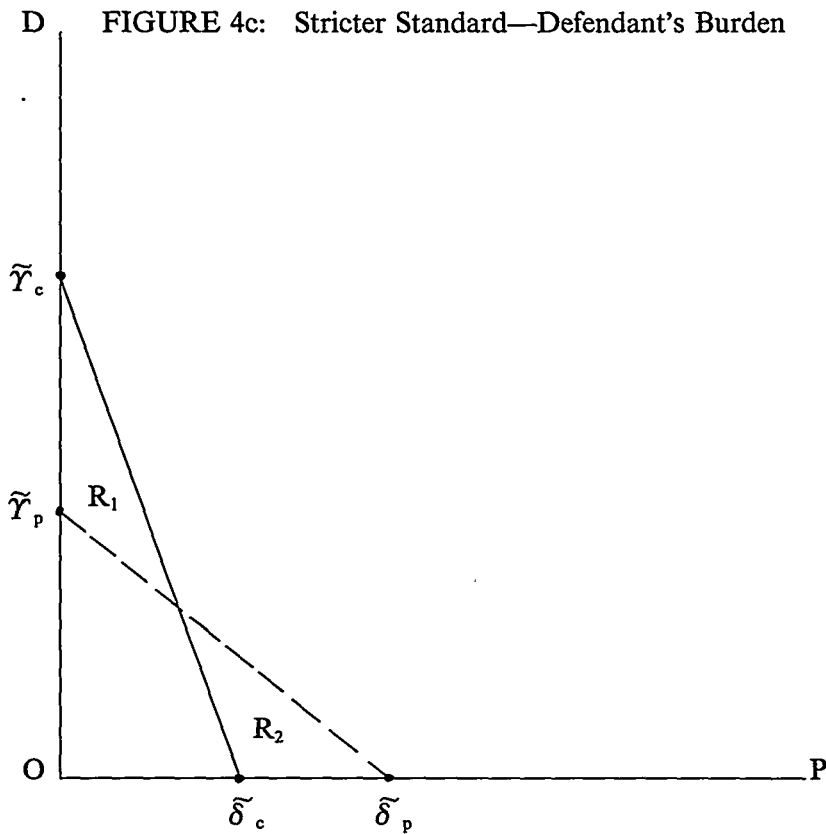


Figure 4c shows the effect of the same change of standards when the burden of persuasion rests with the defendant. In this case, Type I risk is increased by more than the reduction in Type II risk, again resulting in greater total risk of error.

C. Uncertain Rules

Uncertainty in the law is partly a function of the specification of legal rules. Uncertainty can be aggravated by imprecise wording in a statute or decision, or by unsettled judicial doctrines. If the standard of liability for a given offense is frequently revised, the effect on uncertainty

---

140. I.e.,  $(\gamma_p - \gamma_c) \sum D[i] \leq (\delta_c - \delta_p) \sum P[i]$ .



will be much the same as if the standard were vaguely drafted.<sup>141</sup> When rules are not clearly specified, neither litigants nor the courts may know whether certain behavior will result in liability. This will generally have the effect of increasing both  $\Upsilon$  and  $\delta$ ,<sup>142</sup> shifting the error frontier outward (see Figure 5). Increased uncertainty may leave the slope of the frontier unchanged, or it may increase the risk of one type of error by more than the other.

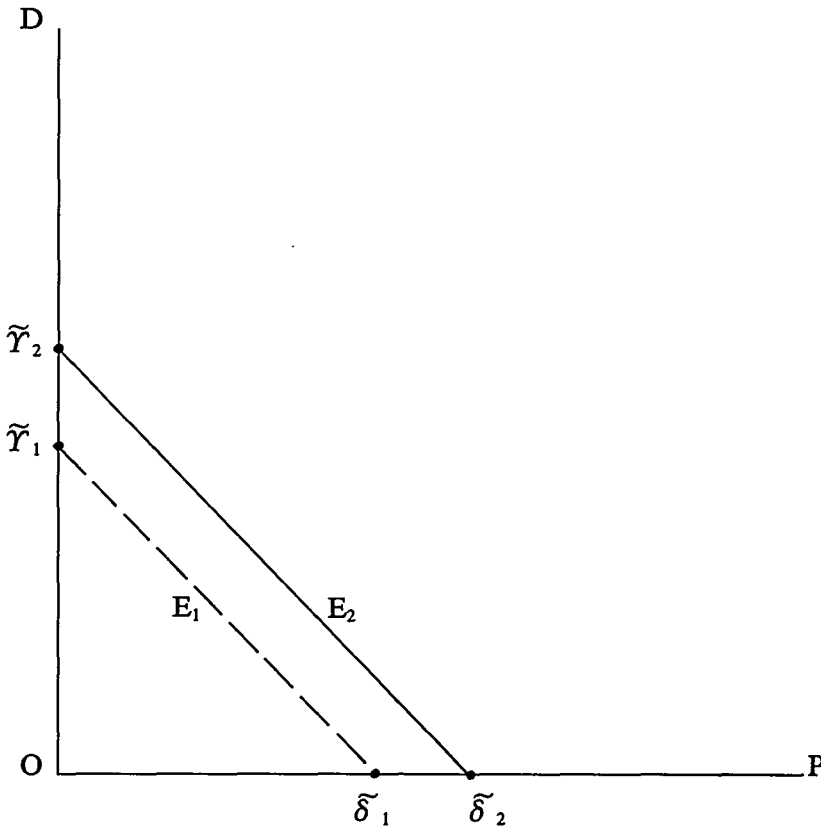


FIGURE 5: Effect of Increased Uncertainty

#### *D. The Example of Defamation Law*

Applying the foregoing analysis to defamation law leads to surpris-

141. One important distinction exists. Vague drafting is ultimately constrained by the void-for-vagueness doctrine; the vagueness of unsettled rules is not.

142. Uncertain standards will increase the variance of the distributions in Figure 2. If the mean apparent liability remains the same for both classes of defendant, the two distributions will overlap to a greater extent, increasing  $\alpha$  and  $\beta$ . *Ceteris paribus*, this will also raise the values of  $\Upsilon$  and  $\delta$ . See *supra* notes 128-29.

ing conclusions. Under the common law, publishing a written defamation was a strict liability tort.<sup>143</sup> Truth and privilege were the only defenses.<sup>144</sup> Plaintiffs had to prove injury; defendants had to prove justification. In 1964 the Supreme Court began a piecemeal expansion of the class of defamations that could be published without liability. *New York Times Co. v. Sullivan*<sup>145</sup> was the landmark case where the Court held that public officials could not recover for defamatory falsehoods concerning their official conduct, absent a showing of "actual malice."<sup>146</sup> Actual malice was defined as knowledge of or reckless disregard as to a statement's falsity.<sup>147</sup> The burden of proving actual malice was placed on the plaintiff.<sup>148</sup> The apparent rationale for this change was that requiring defendants to prove the truth of their statements imposed too high a cost, in the risk of erroneous judgments for plaintiffs.<sup>149</sup>

---

143. Sheer & Zardkoohi, *supra* note 70, at 369 n.4.

144. *Id.* at 370. The privilege defense can be ignored for the present analysis.

145. 376 U.S. 254 (1964).

146. *Id.* at 280.

147. *Id.*

148. *Id.* at 281.

149. The risk of erroneous verdicts for plaintiffs is implicit in the Court's concern with possible "chilling" effects and the need to provide "breathing room" for constitutionally protected speech. See, e.g., *Hepps*, 475 U.S. 767, 772; *Dun & Bradstreet, Inc. v. Greenmoss Builders, Inc.*, 472 U.S. 749, 762 (1985); *New York Times Co. v. Sullivan*, 376 U.S. 254, 272 (1964); *NAACP v. Button*, 371 U.S. 415, 433 (1963).

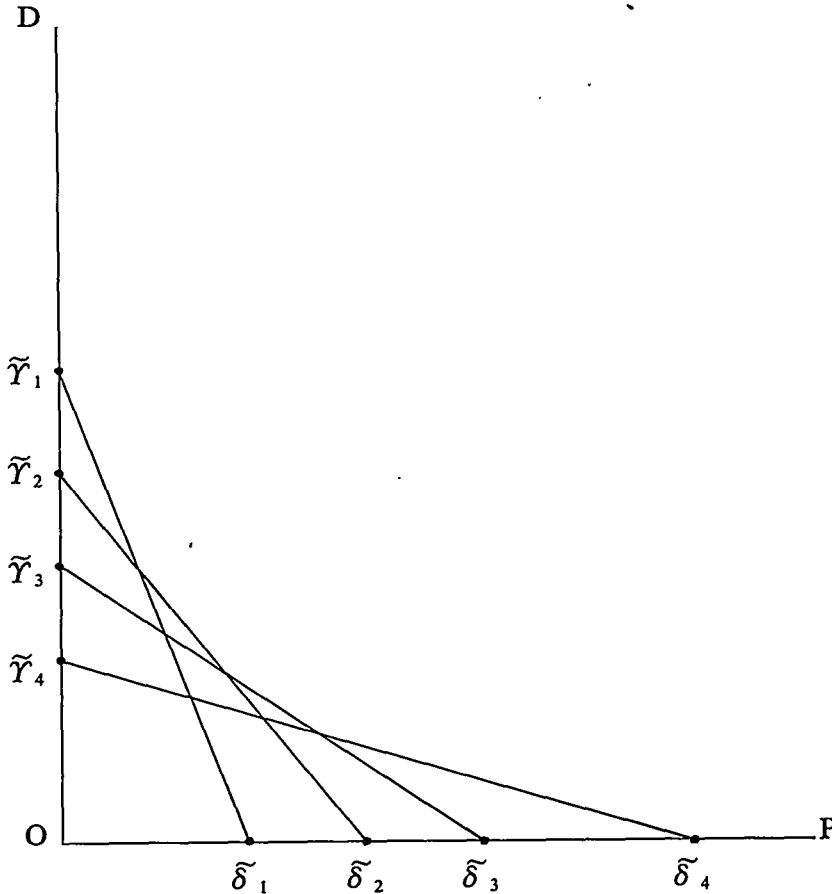


FIGURE 6a: Component Effects of *Sullivan*

In adopting the actual malice standard, the Court presumably intended the effect diagrammed in Figure 6a. Under the common-law standard,  $\gamma_1$  gave the probability of mistakenly holding a defendant liable for publishing a true statement;  $\delta_1$  gave the risk that a defendant would not be held liable for publishing a falsehood. Under the actual malice rule, some true statements that are mistakenly judged false will be held non-malicious, thereby decreasing the risk of Type I error to  $\gamma_2$ .<sup>150</sup> Correspondingly, some false statements published with actual malice will be ruled non-malicious, increasing Type II risk to  $\delta_2$ . Giving the plaintiff the burden of proving actual malice causes a further reduction in Type I risk, to  $\gamma_3$ , with an equivalent rise in Type II risk, to  $\delta_3$ .<sup>151</sup> Finally, raising the standard of proof to clear and convincing causes yet another re-

150. Outcomes that are correct for the wrong reasons are not counted as errors.

151. See *supra* text accompanying note 136.

alignment in the frontier, to  $\tilde{\gamma}_4 \tilde{\delta}_4$ .<sup>152</sup> The net effect, as shown in Figure 6b, is a significant reduction in the risk of error in favor of plaintiffs ( $R_1$ ), but a more than proportionate rise in the probability of mistaken verdicts for defendants ( $R_2$ ).

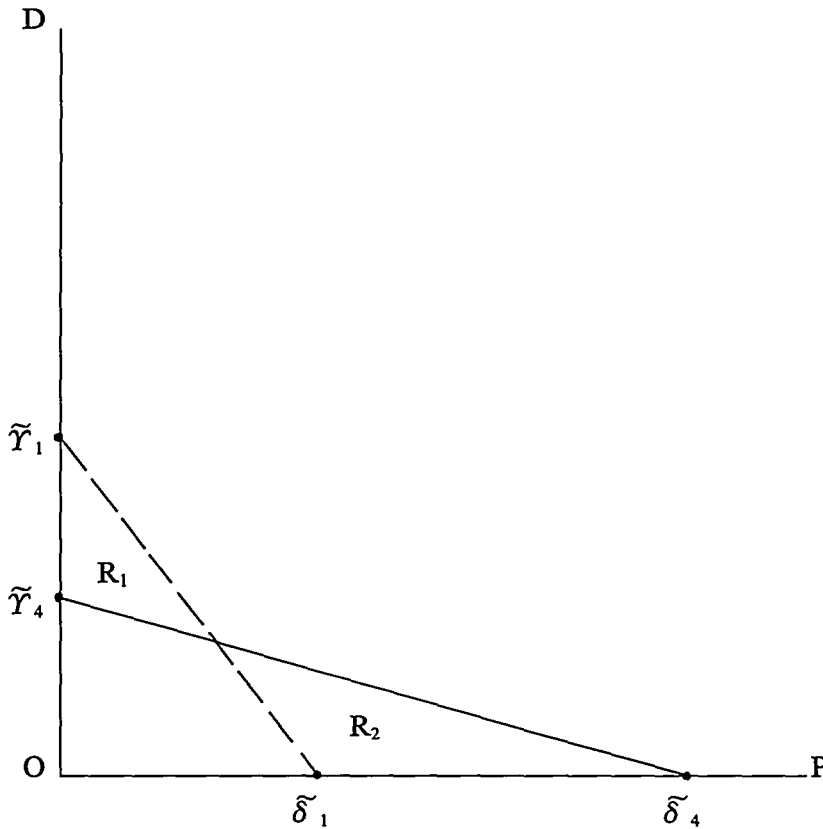


FIGURE 6b: Net Effect of *Sullivan*

Of course, some cases that would have been decided correctly under the common law will result in Type I error under *Sullivan*. These cases involve falsehoods published without malice that are mistakenly held to be malicious. The erroneous appearance of malice was irrelevant under the old rule, but now results in liability where, under the actual malice standard, none should accrue. Conversely, some non-malicious falsehoods that would have escaped liability at common law by being adjudged true (Type II error) will count as correct decisions under the

152. See *supra* text accompanying note 150.

malice standard.<sup>153</sup> These effects will rotate the error frontier back toward its initial position, as shown in Figure 6c.

153. The effect of adding an "absence of malice" defense is to create a third category of defamatory statements in addition to those that are simply true or false. The effect of this change on the relative incidence of error can be seen from the following table:

STATE OF THE WORLD	JUDICIAL CHARACTERIZATION			
		TRUE	FALSE (Non-Malice)	FALSE (Malice)
	TRUE	CL: Correct M: Correct	CL: Type I M: Correct	CL: Type I M: Type I
	FALSE (Non-Malice)	CL: Type I M: Correct	CL: Correct M: Correct	CL: Correct M: Type I
	FALSE (Malice)	CL: Type I M: Type II	CL: Correct M: Type II	CL: Correct M: Correct

CL: Common Law Standard      M: Malice Standard

The offsetting effects of introducing the malice standard can be highlighted simply:

		CHARACTERIZATION		
		True	False (Non-M)	Malice
A C C U R A C Y	True		REDUCE TYPE I	
	False (Non-M)	REDUCE TYPE II		INCREASE TYPE I
	Malice		INCREASE TYPE II	

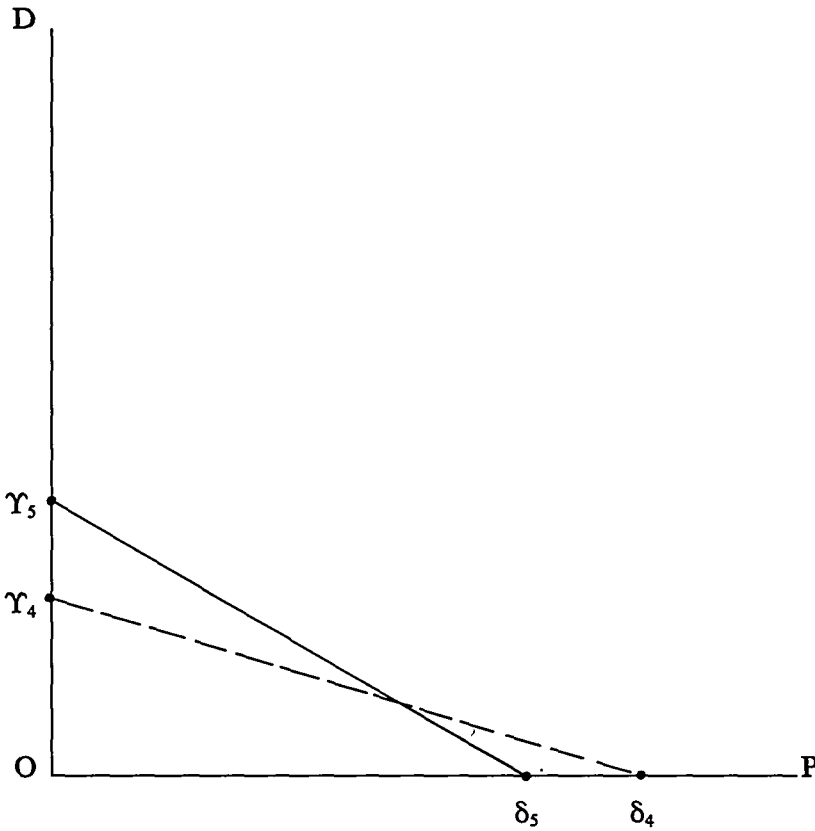


FIGURE 6c: Offsetting Effects in *Sullivan*

The extent of this offsetting rotation is indeterminate. The heightened Type I risk of holding innocently false statements to be malicious ( $\tilde{\gamma}_5 - \tilde{\gamma}_4$ , Figure 6c) may arguably be greater than the reduced risk of holding some true statements to be false ( $\tilde{\gamma}_1 - \tilde{\gamma}_2$ , Figure 6a). This is because malice is a mental state, proof of which is inherently subject to greater uncertainty than the fact of publishing a false statement.<sup>154</sup> This raises the possibility that, by itself, adopting the actual malice standard may have actually increased the risk of Type I error in libel actions by public officials. Any net decline in the risk of mistaken liability stems entirely from the accompanying changes in the burden and standard of proof.

#### IV. CONCLUSIONS, CONFLICTS, CAVEATS

Tracing the incidence and consequences of legal error can quickly

---

154. See *supra* text accompanying note 141.

become a very complex matter. The first step in advancing our understanding is to lay aside the myth that systematic legal errors do not occur. The second step is to incorporate predictable error into our models of legal decision-making. This Article has been addressed mainly to these preliminary issues.

It seems clear that error minimization often conflicts with other values in the legal system. A fully developed model should help us evaluate these trade-offs by identifying the actual costs and benefits involved. Consider again the recurring example of defamation law. One response to my analysis of the malice standard<sup>155</sup> is to observe that this test was intended to advance values other than accuracy. For example, although the malice standard may have increased the proportion of Type I errors in defamation cases,<sup>156</sup> it has presumably reduced the proportion of defendants found guilty.<sup>157</sup> This follows from the qualitative difference between the Type I errors eliminated by adopting the new standard, and those created by it. The reduction in risk given by  $\tilde{Y}_1 - \tilde{Y}_2$  in Figure 6a represents a decrease in the expected proportion of convictions. The offsetting increase ( $\tilde{Y}_4 - \tilde{Y}_5$ , Figure 6c) consists only of a change in classification; a certain number of convictions that would have been counted as correct under the old standard will be counted as errors under the new one. If the Court's goal was merely to reduce the likelihood of conviction, the malice standard could therefore be considered successful regardless of its effect on judicial accuracy. However, this leaves open the question of whether this objective could have been accomplished by other means entailing a lesser cost of error (say, by simply elevating the standard of proof). It is always appropriate to examine the effect of a rule on the magnitude and distribution of error. Without such an evaluation, it is impossible to judge how closely a given standard approaches a genuine social optimum.

As a guide to policy, error minimization generally points toward clearer rules, fewer rules, or no rules at all.<sup>158</sup> Thus, error minimization may directly conflict with the law's deterrence and social incentive values. As Craswell and Calfee have shown, increasing the clarity of rules can lead to either over- or underdeterrence of proscribed behavior.<sup>159</sup>

155. See *supra* text accompanying notes 150-54.

156. See *supra* text accompanying note 154.

157. This would have been the case if the amount of defamation and litigation remained constant. But the uncertainty created by the malice standard may have increased either or both of these variables, leaving the overall effect indeterminate.

158. See D'Amato, *supra* note 134, at 45-51.

159. Craswell & Calfee, *supra* note 18; cf. Johnston, *Uncertain Fact-Finding and Efficiency: Toward a General Economic Theory of Liability*, 61 S. CAL. L. REV. 137 (1987); Calabresi &

The social costs of these effects must be balanced against the benefits of reduced error in determining the optimal level of rule uncertainty.

Finally, future models should consider endogenous effects and feedbacks that have not been treated here. For example, the quality and quantity of evidence may often respond to shifts in the burden of proof—especially the burden of production. This may either increase or decrease the total risk of error, in contrast to the assumptions of this Article.<sup>160</sup> Such effects must ultimately be included in comprehensive models of the legal process.

---

Klevorick, *Four Tests for Liability in Torts*, 14 J. LEGAL STUD. 585, 617-19 (1985); Calfee & Craswell, *Some Effects of Uncertainty on Compliance With Legal Standards*, 70 VA. L. REV. 965 (1984); Farago, *Intractable Cases: The Role of Uncertainty in the Concept of Law*, 55 N.Y.U. L. REV. 195 (1980).

160. See *supra* text accompanying notes 135-38.



