

1-1-1996

Allocating the Burden of Proof in Sales Litigation: The Law, Its Rationale, a New Theory, and Its Failure

Alex Stein

Follow this and additional works at: <https://repository.law.miami.edu/umlr>

Recommended Citation

Alex Stein, *Allocating the Burden of Proof in Sales Litigation: The Law, Its Rationale, a New Theory, and Its Failure*, 50 U. Miami L. Rev. 335 (1996)

Available at: <https://repository.law.miami.edu/umlr/vol50/iss2/4>

This Essay is brought to you for free and open access by the Journals at University of Miami School of Law Institutional Repository. It has been accepted for inclusion in University of Miami Law Review by an authorized editor of University of Miami School of Law Institutional Repository. For more information, please contact library@law.miami.edu.

ESSAY

Allocating the Burden of Proof in Sales Litigation: The Law, Its Rationale, a New Theory, and Its Failure

ALEX STEIN*

I. INTRODUCTION	335
II. 1890, CAMBRIDGE, MASSACHUSETTS	336
III. 1994, NEW HAVEN, CONNECTICUT	336
IV. 1961, NASHVILLE, TENNESSEE	338
V. 1994, NEW HAVEN, CONNECTICUT	338
VI. 1961 TO PRESENT, TOO MANY PLACES TO BE MENTIONED	340
VII. 1996, CORAL GABLES, FLORIDA	343

I. INTRODUCTION

This essay examines the rationale for allocating the burden of persuasion in relation to the conformity of the tendered goods in sales cases. In a recently published article which analyzes this problem from an economic efficiency perspective, Professor Jody S. Kraus argues that this burden should be shouldered by the party with the best access to evidence.¹ By juxtaposing this argument with the principal writings on the subject (inexplicably ignored and consequently not confronted by Professor Kraus) I demonstrate that the argument and its supporting ideas are severely flawed. Economic efficiency would place the persuasion burden upon plaintiffs, even when their access to evidence is inferior relative to defendants'. As established long ago, access to evidence may only be invoked as a controlling factor in allocating the burden of *production*. Deviations from this conventional approach may be economically justified in special cases, but sales cases are not special.

* Copyright © (November 1995) Alex Stein. Visiting Professor, University of Miami School of Law (1994-95); Senior Lecturer, Faculty of Law, Hebrew University of Jerusalem (1991-); LL.B. (1983); LL.M. (1987) (Hebrew University of Jerusalem); Ph.D. (1990) (University of London).

I am grateful to Ronald J. Allen, A. Michael Froomkin, and Eyal Zamir for their helpful comments on an earlier draft.

1. Jody S. Kraus, *Decoupling Sales Law from the Acceptance-Rejection Fulcrum*, 104 *YALE L.J.* 129, 135-52 (1994).

II. 1890, CAMBRIDGE, MASSACHUSETTS

James Bradley Thayer publishes a pioneering article that clarifies the notion of "burden of proof" by distinguishing between the burden of persuasion and that of producing evidence.² The persuasion burden allocates risk of error in conditions of uncertainty, when the requisite standard of proof has not been satisfied by either side. This burden thus functions to mark out the bearer of the risk of error. The production burden is a duty to adduce evidence that needs to be brought before judges or jurors. This burden thus functions to identify the party who will be exposed to an adverse ruling when an issue pertaining to the case is left evidentially unsubstantiated. Consequently, the reasons for allocating each of these discrete burdens should also be different. For example, the mere fact that one party to a proceeding holds relevant information or has peculiarly good access to some important evidence cannot be a valid reason for shifting the persuasion burden to him. Once his evidence is produced for examination at the trial, his advantage evaporates. Bentham's idea of placing the burden of proof "on whom it will sit lightest"³ should accordingly only apply to the production burden.⁴ Allocation of the risk of non-persuasion should be grounded in other reasons, such as substantive legal preferences.⁵

Although well-known, Thayer's approach has not been unequivocally endorsed.⁶ This may partially explain, albeit not justify, Professor Kraus's thesis.

III. 1994, NEW HAVEN, CONNECTICUT

Professor Kraus publishes an article examining, *inter alia*, the allocation of the persuasion burden under the sales law provisions of the Uniform Commercial Code (UCC), in relation to the issue of whether tendered goods are "conforming."⁷ Under the UCC, this burden is placed on the seller of rejected goods and the buyer of accepted goods.⁸

2. James B. Thayer, *The Burdens of Proof*, 4 HARV. L. REV. 45 (1890). The same later appeared in Thayer's book *A PRELIMINARY TREATISE ON EVIDENCE AT THE COMMON LAW* ch.9 (Augustus M. Kelley 1969) (1898).

3. JEREMY BENTHAM, *An Introductory View of Rationale of Evidence; For the Use of Non-Lawyers as Well as Lawyers*, in 6 THE WORKS OF JEREMY BENTHAM 139 (John Bowring ed., 1962).

4. Thayer, *supra* note 2, at 59-70.

5. *Id.* at 64-65; see generally Charles V. Laughlin, *In Support of the Thayer Theory of Presumptions*, 52 MICH. L. REV. 195 (1953); James B. Thayer, *Presumptions and the Law of Evidence*, 3 HARV. L. REV. 141 (1889).

6. See 2 CHARLES T. MCCORMICK, *MCCORMICK ON EVIDENCE* § 337 (John W. Strong ed., 4th ed. 1992); CHRISTOPHER B. MUELLER & LAIRD C. KIRKPATRICK, *EVIDENCE* 128-30 (1995).

7. Kraus, *supra* note 1.

8. U.C.C. §§ 1-201(8), 2-607(4) (1994). See Kraus, *supra* note 1, at 136-37.

Confined to an economic analysis of the problem,⁹ Professor Kraus's examination starts with the access to evidence, which he considers to be a plausible ground for allocating the persuasion burden.¹⁰ He admits that "[t]he strength of this rationale . . . has diminished since the advent of modern discovery law,"¹¹ arguing at the same time that it "might still explain why the Code's burden-of-proof rule is efficient."¹²

Thoroughly discredited by Thayer (some decades before the advent of modern discovery law), this rationale is manifestly weak. An attempt at revitalizing it can therefore properly be made only by acknowledging and confronting Thayer's argument. Unfortunately, Professor Kraus ignores this argument altogether.

Another rationale examined by Professor Kraus is closely related to the access to evidence. It places the persuasion burden upon the cheapest producer of cost-effective evidence.¹³ This rationale (to the best of my knowledge) was pioneered by Bentham.¹⁴ However, as explicated by Thayer, Bentham used the term "burden of proof" generically, without distinguishing between the two burdens. Thayer's argument makes it clear that the cheapest producers of cost-effective evidence should carry only the production burden. Regrettably, Professor Kraus fails to explain why they should also absorb the risk of non-persuasion.

Professor Kraus's analysis of the problem focuses on incentives for a self-interested litigant: one that would produce evidence favorable to her case and try to suppress any unfavorable evidence.¹⁵ Would unfavorable evidence be produced by a litigant if she carried the persuasion burden? Certainly not. Production of unfavorable evidence would entail a certain loss on the issue; to bear the risk of non-persuasion is to be exposed only to a probable loss.¹⁶ Truth lovers do not require the incentives focused upon by Professor Kraus; as for Holmes's "Bad Man," he will be left thoroughly unimpressed.¹⁷ The "access to evi-

9. Kraus, *supra* note 1, at 132-35, 138-41.

10. *Id.* at 142-46.

11. *Id.* at 142; see Ronald J. Allen, *Presumptions in Civil Actions Reconsidered*, 66 IOWA L. REV. 843, 860 (1981) ("[I]t is difficult to imagine a case today in which the sanction of possible dismissal would generate more evidence than discovery schemes."); see also Laughlin, *supra* note 5, at 220.

12. Kraus, *supra* note 1, at 143.

13. *Id.* at 146-48.

14. BENTHAM, *supra* note 3, at 136-39.

15. In some cases, evidence favorable to a litigant will not be produced because its benefits are simply outweighed by the costs of production. From the utilitarian perspective, such outcomes will be desirable.

16. Professor Kraus is apparently aware of this. See *supra* note 1, at 147 n.55.

17. As for ambiguous evidence, it will be produced only when it is expected to be more beneficial than harmful. Allocation of the persuasion burden cannot affect such calculations. As for hopelessly ambiguous evidence, I can see no economic justification for its extraction.

dence" rationale would thus work against Bad Men and truth lovers alike. Because the ratio of Bad Men and truth-lovers amongst buyers and sellers of goods is unknown—and, indeed, unknowable—allocation of the persuasion burden by this rationale would produce more disutility than utility. It will not produce more correct verdicts than erroneous. At the same time, it will exert a chilling effect on at least some truth lovers. They will attempt to suppress evidence not unambiguously supportive of their allegations, which they believe to be true. This chilling effect will result from fear that every factual ambiguity will, justifiably or not, benefit the opponent.¹⁸ As demonstrated below, efficiency concerns place the persuasion burden upon *plaintiffs*, regardless of their position in litigated sales transactions.¹⁹

IV. 1961, NASHVILLE, TENNESSEE

Professor Vaughn C. Ball publishes an influential article that will become regarded as a precursor of the "New Evidence Scholarship."²⁰ One of its insights was the following warning against double-counting:

The risk of non-persuasion is allocated (a great part of the time, at least) upon the basis of the probability of the existence of the fact in the run of cases of the particular kind, absent any specific evidence. Since the evidence and the jury's consideration of it have come to naught, we will make the fewest mistakes if we let the case fall back into the general class, to be decided on those original probabilities. But the jury, unless it lacks the common knowledge we ascribe to it by definition, has begun its own deliberation with those probabilities in mind, and it is the combination of both those and the probabilities drawn from the specific evidence, that the jury says are at a balance. If we then use the initial probabilities to remove the balance, we are in some sense counting them twice.²¹

V. 1994, NEW HAVEN, CONNECTICUT

The third rationale examined by Professor Kraus "seems to underlie

Inferential progress generated by such evidence would be minimal, while its examination at trial would be costly. *Cf.* FED. R. EVID. 102 & 403.

18. See Steven Shavell, *Optimal Sanctions and the Incentive to Provide Evidence to Legal Tribunals*, 9 INT. REV. LAW & ECON. 3, 4 (1989).

19. But the UCC's allocation of the burden may still be justified on noneconomic grounds. See discussion *infra* note 40.

20. Vaughn C. Ball, *The Moment of Truth: Probability Theory and Standards of Proof*, 14 VAND. L. REV. 807 (1961). The "New Evidence Scholarship" is briefly described below. See *infra* notes 32, 33 and accompanying text.

21. Ball, *supra* note 20, at 817-18 (footnote omitted). An example of this fallacy, referred to by Professor Ball, can be found in Julius Stone, *Burden of Proof and the Judicial Process*, 60 L. Q. REV. 262, 278-84 (1944); *cf.* Laughlin, *supra* note 5, at 212.

the practice of assigning the burden of proof to the party asserting conduct that is out of the ordinary."²² The party most likely to be asserting a false claim should thus carry the risk of non-persuasion:

Like the access-to-evidence and evidence-production rationales considered above, this rationale seeks cost-effectively to increase the likelihood of accurate adjudication. If particular kinds of claims are known to be probably false, then assigning the burden of proof to the party asserting such claims might increase the accuracy of adjudication. Whatever reasons justify the *ex ante* belief that the sort of claim being advanced is probably false would also justify a rebuttable legal presumption that the claim is false. The allocation of the burden of proof to the party asserting such a claim creates that presumption. As a result, the party most likely to be correct will be more likely to prevail.²³

Juxtaposition of this rationale with Professor Ball's argument can be made without my assistance. Fortunately, Professor Kraus discards this rationale in the context of his discussion. The prior probability of finding liars amongst sellers of rejected goods is unknowable. This conclusion also holds true with regard to buyers of accepted goods.²⁴ As for the probability of conformity, it is similarly unknowable in relation to both accepted and rejected goods.²⁵

Professor Kraus thus concludes:

What is the best rule to substitute in place of the Code's burden-of-proof rule? The rationale of assigning the burden to the party likely to be asserting a false claim has no obvious application in sales cases, and is misguided. Only the "access-to-evidence" rationale remains. Therefore, the choice is between the two rules suggested by the two different interpretations of the "access-to-evidence" rationale: the "access-to-the goods" [sic] rule, which requires the party in possession to bear the burden, and the "access-to-inspection-evidence" rule, which requires the buyer to bear the burden following tender. In the absence of empirical data, it is difficult to choose between them; fortunately, they are likely to converge in most cases. Whenever the buyer alleges nonconformity, the allegation is likely to be made after tender and after the buyer has taken possession of the goods. I would therefore support a rule allocating the burden of proof to the buyer upon tender.²⁶

Appropriate as it may be for allocating the production burden, this rationale is thoroughly misguided when employed in placing the risk of

22. Kraus, *supra* note 1, at 148 n.56.

23. *Id.* at 148.

24. *See id.* at 149-50.

25. *Id.*

26. *Id.* at 151-52.

non-persuasion. As stated above, this was demonstrated by Thayer in 1890.²⁷

Despite this objection, Professor Kraus's conclusion could still survive, if the access-to-evidence rationale were the only rationale to survive scrutiny. This, however, is not the case. As Professor Kraus himself acknowledges, other candidates include:

- (1) allocation of the persuasion burden in accordance with the pleadings;
- (2) allocation of the burden to a party asserting a disfavored claim or defence;
- (3) allocation of the burden to a party attributing wrongdoing to her opponent;
- (4) allocation of the burden to a party attempting to change the status quo.²⁸

The first two candidates can be dismissed as begging the question.²⁹ But what about the remaining two? Professor Kraus argues that they can also be dismissed "because they offer no efficiency-based rationale for allocating the burden of proof" and "because they are inapplicable to sales cases."³⁰ As far as the economics of the wrongdoing rationale are concerned, Professor Kraus is probably right.³¹ As for his parenthetical remark about the status quo rationale, it surprised me even more than his previous omissions.

VI. 1961 TO PRESENT, TOO MANY PLACES TO BE MENTIONED

The New Evidence Scholarship emerges and develops.³² It produces numerous interdisciplinary inquiries into various fact finding

27. Thayer, *supra* note 2, at 64-65.

28. Kraus, *supra* note 1, at 141 n.43.

29. See, e.g., Dale A. Nance, *Civility and the Burden of Proof*, 17 HARV. J. L. & PUB. POL'Y 647, 663-69 (1994).

30. Kraus, *supra* note 1, at 141 n.43.

31. On this point, however, I fail to understand Professor Kraus's logic. According to his own argument, a regularity-based allocation of the persuasion burden can, in principle, be justified on efficiency grounds. Kraus, *supra* note 1, at 148-50. Most people assume, in proceeding in daily affairs, that conformity with legal standards is far more common than wrongdoing. See Nance, *supra* note 29, at 648-55. If so, the wrongdoing rationale can, perhaps, be justified on efficiency grounds as well.

Professor Nance supports this rationale from another angle. He argues that the existing burden-allocating framework should be justified by the moral principle of civility. From this principle he derives a rebuttable presumption in favor of compliance with the legally prescribed standards. See *id.* at 655-72. Professor Nance's theory would allocate the persuasion burden to the party alleging breach of contract by her opponent, and thus have an obvious impact in sales cases.

32. The term was coined by Professor Richard Lempert, *The New Evidence Scholarship: Analyzing the Process of Proof*, 66 B.U. L. REV. 439 (1986).

problems,³³ incoming economic analyses of the persuasion burden as applied in civil litigation. According to one of these analyses:

- (1) D_d represents the disutility a plaintiff will sustain as a result of a wrong decision for the defendant, and D_p represents the disutility a defendant will sustain if the error favors the plaintiff.
- (2) P represents the probability of the allegations brought by plaintiff.
- (3) As the case is one of uncertainty, $0 < P < 1$, where 1 represents absolute certainty and 0 impossibility.
- (4) Plaintiff should thus recover whenever:

$$PD_d > (1-P) D_p.$$

- (5) This can be reformulated as follows:

$$P > \frac{1}{1 + \frac{D_d}{D_p}}.$$

- (6) Subject to special cases, plaintiffs' and defendants' wrongful losses are equally inefficient (and equally inequitable): $D_p = D_d$.
- (7) Hence, a plaintiff should recover whenever $P > 0.5$, and a defendant should win whenever $P < 0.5$ (the "P>0.5 rule").³⁴

The optimality of the P>0.5 rule has been demonstrated with an even greater precision:

- (1) D denotes the value of the litigated goods;
- (2) p_1 and p_2 denote, respectively, the probabilities of plaintiff's and defendant's conflicting allegations. As is always the case: $0 < p_1 < 1$ and $0 < p_2 < 1$, where 1 stands for certainty and 0 for impossibility.
- (3) The following decisions thus become available:
 - d_1 = plaintiff loses (the risk of error is imposed on plaintiff);
 - d_2 = defendant loses (the risk of error is imposed on defendant);
 - d_3 = a compromise reflecting the expected value of each allegation: plaintiff recovers from defendant p_1D ; p_2D goes to defendant by not allowing plaintiff to recover this amount.³⁵
- (4) S_1 and S_2 will now respectively denote the actual states of affairs, favorable to either plaintiff or defendant.
- (5) The average damage to be incurred by each of the available decisions will be as follows:

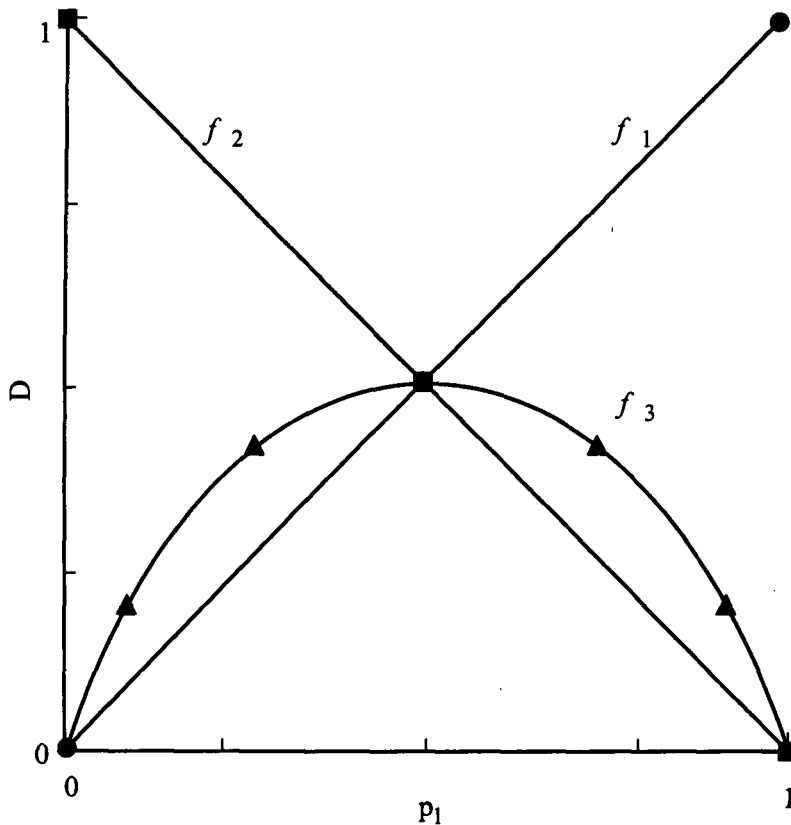
33. For a brief survey see William L. Twining & Alex Stein, *Introduction to EVIDENCE AND PROOF* at xxi-xxiv (William L. Twining & Alex Stein eds., 1992).

34. John Kaplan, *Decision Theory and the Factfinding Process*, 20 STAN. L. REV. 1065, 1071-72 (1968); see also David Kaye, *Naked Statistical Evidence*, 89 YALE L.J. 601, 605 n.19 (1980) (book review).

35. The parties' trial expenses are ignored for the sake of simplicity.

DECISION	DAMAGE IF S_1	DAMAGE IF S_2	TOTAL DAMAGE
d_1	p_1D	0	p_1D
d_2	0	p_2D	p_2D
d_3	p_1p_2D	p_2p_1D	$2p_1p_2D$

(6) The long run probability-damage relationships can thus be represented and assessed as follows:



f_1 = plaintiff absorbs the damage function corresponding to d_1 ;
 f_2 = defendant absorbs the damage function corresponding to d_2 ;
 f_3 = the expected value function corresponding to d_3 .³⁶

In balanced cases, i.e., when $p_1=p_2=0.5$, each of the available decisions will result in the same overall damage. From an economic view-

36. See David Kaye, *The Limits of the Preponderance of the Evidence Standard: Justifiably Naked Statistical Evidence and Multiple Causation*, 1982 AM. B. FOUND. RES. J. 487, 493.

This theory has not gone unchallenged. It has been argued that under the diminishing utility of wealth assumption—which would require the legal system to minimize large errors— d_3 would be the best decision. Indeed, by denoting the average large error as D^2 , we will arrive at the following:

point, however, defendant should prevail in such a case. This decision rule would eliminate the enforcement costs that would be incurred if plaintiff recovers. In addition, to allow plaintiff to recover when $p_1=0.5$ would raise the number of unmeritorious claims, thus incurring greater litigation costs.³⁷ When $p_1=p_2=0.5$, decision d_1 would be optimal also because *taking* is generally perceived as more harmful than *not giving*. This perception can be justified by the diminishing utility of wealth.³⁸

VII. 1996, CORAL GABLES, FLORIDA

My discussion has arrived at its concluding point. From an economic efficiency perspective, a plaintiff should be denied recovery whenever the probability of her allegations is below 0.5. This decision rule should apply in all civil cases, including sales cases. In exceptional circumstances, costs avoidable by this rule may be offset by other benefits. If such circumstances are both recurrent and amenable to advance doctrinal formalization, an appropriate replacement of the $P>0.5$ rule would be in order.³⁹

Typical sales cases do not involve such circumstances. The burden

DECISION	DAMAGE IF S_1	DAMAGE IF S_2	TOTAL DAMAGE
d_1	p_1D^2	0	p_1D^2
d_2	0	p_2D^2	p_2D^2
d_3	$p_1(p_2D)^2$	$p_2(p_1D)^2$	$p_1p_2D^2$

See Neil Orloff & Jerry Stedinger, *A Framework for Evaluating the Preponderance-of-the-Evidence Standard*, 131 U. PA. L. REV. 1159, 1165-68 (1983).

37. See Ralph K. Winter, Jr., *The Jury and the Risk of Nonpersuasion*, 5 LAW & SOC'Y REV. 335, 337 (1971).

38. See RICHARD A. POSNER, *ECONOMIC ANALYSIS OF LAW* 552 (4th ed. 1992). Incidentally, this echoes the Jewish Law principle "המורצ'א מוזברו עליו הר'א'ד" ("one who purports to *take* from his fellow man should bear the burden of proof."). See BABA KAMMA 6a, 46b (E.W. Kirzner trans.) in 1 THE BABYLONIAN TALMUD: SEDER NEZIKIN (I. Epstein ed. & trans. 1952).

Empirical research has shown that giving away a psychologically "vested" right is perceived as more painful than not seizing upon an identically valuable but psychologically "unvested" right. See Robert C. Ellickson, *Bringing Culture and Human Frailty to Rational Actors: A Critique of Classical Law and Economics*, 65 CHI-KENT L. REV. 23, 35-40 (1989); Amos Tversky & Daniel Kahneman, *Rational Choice and the Framing of Decisions*, 59 J. BUS. S251, S260 (1986). This can also explain allowing plaintiff to recover from defendant only when $P>0.5$.

39. Leaving such decisions to judges would involve exorbitant enforcement costs. Decomposition of the doctrine into a multitude of case-specific rules would involve unjustifiably high promulgation costs. Determination of each individual rule would be undesirably costly, while its enforcement benefits, attainable only in a few cases, would be very modest. The proposed law-making strategy would therefore be optimal. See generally Louis Kaplow, *Rules Versus Standards: An Economic Analysis*, 42 DUKE L.J. 557 (1992). For a different proposal, not motivated solely by efficiency concerns, see Ronald J. Allen & Robert A. Hillman, *Evidentiary Problems in—and Solutions for—the Uniform Commercial Code*, 34 DUKE L.J. 92, 105-08 (1984) (proposing allowing judges to reallocate the persuasion burden in special circumstances).

of proof rules should therefore be decoupled from the acceptance-rejection fulcrum, but not in a way suggested by Professor Kraus. The persuasion burden in sales litigation should always be placed upon the plaintiff, irrespective of whether she is a buyer or a seller, and regardless of whether the goods of disputed conformity were accepted or rejected.⁴⁰

40. A rights-based, nonutilitarian approach would, however, lead to an entirely different conclusion. Thus, under the equality principle, the persuasion burden should be distributed evenly over issues rather than in a fashion that systematically favors defendant over plaintiff. This principle may justify the UCC's allocation of the burden. See Alex Stein, *The Refoundation of Evidence Law*, 9 CANADIAN J. L. & JURISPRUDENCE (forthcoming 1996). This justification should be distinguished from the idea of equalizing the overall error rate between plaintiffs and defendants as groups. See MICHAEL O. FINKELSTEIN, *QUANTITATIVE METHODS IN LAW* 68 (1978). Finkelstein's idea is untenable because members of his groups are not mutually associated and do not share their gains and losses. See Kaye, *supra* note 34, at 607-08.