1-1-1996

More Justice for Less Money

David D. Friedman

Santa Clara University School of Law, ddfr@daviddfriedman.com

Follow this and additional works at: http://digitalcommons.law.scu.edu/facpubs

MORE JUSTICE FOR LESS MONEY*

DAVID FRIEDMAN
Santa Clara University

ABSTRACT

In *Cimino v. Raymark,* a court resolved 2,298 asbestos cases with 160 trials by grouping cases, trying a random sample of each group, and awarding all members of each group the average verdict. I propose an improved "I cut, you choose" version of the procedure designed to correctly allocate damages among plaintiffs. The plaintiffs' attorney presents a claim for what each plaintiff should get. The defense selects cases for trial. The average ratio of award to claim for tried cases is calculated. Each untried case receives that ratio times its claim. The plaintiffs' attorney will set claims proportional to expected awards in order to prevent the defense from selecting overclaimed cases for trial and thus driving down the award for untried cases. Potential problems are examined through a formal model. Modifications are suggested to adapt the procedure to situations involving a very large number of small cases.

> It is apparent from the effort and time required to try these 160 cases, that unless this plan or some other procedure that permits damages to be adjudicated in the aggregate is approved, these cases cannot be tried. Defendants complain about the 1% likelihood that the result would be significantly different. However, plaintiffs are facing a 100% confidence level of being denied access to the courts. The Court will leave it to the academicians and legal scholars to debate whether our notion of due process has room for balancing these competing interests.
> [JUDGE ROBERT PARKER, *Cimino v. Raymark*]

IN *Cimino v. Raymark,*[1] Judge Robert Parker of the Eastern District of Texas implemented a radical solution to the problem of litigating mass torts. Instead of conducting individual trials for several thousand plaintiffs, he selected a random sample of 160, tried their cases, and based the awards given to the remaining plaintiffs on the outcome of those trials. In defending the procedure against the charge that it deprived the parties of due process, he argued that if he had instead required individual trials, most of the cases would never have been resolved.

---

[1] Claude Cimino *et al.* v. Raymark Industries, Inc., *et al.*, 751 F.Supp. 649.

211

The procedure of *Cimino* was explained and defended in a 1992 article by Michael J. Saks and Peter David Blanck.[2] The purpose of this essay is not to dispute either their views or those of Judge Parker but, rather, to suggest a further step along the same path. The procedure of aggregation and sampling implemented in *Cimino* does a reasonably good job of estimating the total damages that the defendants would have paid if every case had been tried separately[3] and does so at a cost much lower than that of individual trials. It does a much poorer job of allocating that total among the plaintiffs. My proposal is intended to solve that problem.

In Part I of this article I explain the procedure used by Judge Parker in *Cimino,* the improvements suggested by Saks and Blanck, and the limitations of the procedure, even with such improvements. Part II describes my proposal for generating an estimate of the relative claims of the plaintiffs and incorporating that estimate into the procedure. Part III considers the legal status of the modified procedure, arguing that it is in some ways more defensible than the version implemented in *Cimino.* Part IV discusses potential problems, both those implicit in the original idea of aggregation and sampling and additional ones created by my proposed modifications. Part V suggests ways in which the procedure I suggest could be extended beyond the context of class actions. Part VI describes the results of the application of the procedure to an explicit formal model; the mathematics are presented in appendix B. Part VII summarizes my conclusions.

I.   AGGREGATION AND SAMPLING: JUDGE PARKER'S SOLUTION TO MASS TORT LITIGATION

> If the Court could somehow close thirty cases a month, it would take six and one-half years to try these cases and there would be pending over 5,000 untouched cases at the present rate of filing. [JUDGE PARKER, *Cimino*]

Cimino v. Raymark went to trial as a class action with 2,298 plaintiffs and five defendants. The trial consisted of three phases. In phase 1, a set

---

[2] Michael J. Saks & Peter David Blanck, Justice Improved: The Unrecognized Benefits of Aggregation and Sampling in the Trial of Mass Torts, 44 Stan. L. Rev. 815 (1992).

[3] This assumes, of course, that the cases would have been tried. As Judge Parker pointed out in his opinion, the defendants "assert a right to individual trials in each case and assert the right to repeatedly contest in each case every contestable issue involving the same products, the same warnings, and the same conduct. The strategy is a sound one; the defendants know that if the procedure in *Cimino* is not affirmed, these cases will never be tried" (Cimino v. Raymark *651-52).

of issues common to all plaintiffs and defendants were resolved.[4] Phase 2 apportioned causation among the defendants and determined which plaintiffs had had sufficient exposure to asbestos, based on each plaintiff's workplace and craft, for such exposure to be a producing cause of an asbestos-related injury or disease.[5]

The purpose of phase 3 was to determine damages. Instead of trying all cases, the court divided the plaintiffs into five categories, according to the asbestos-related disease from which each suffered. A random sample was drawn from each category; the sample was larger for categories with more plaintiffs.[6] The sample cases were tried. Plaintiffs in the sample received the damages that they were awarded. Plaintiffs not in the sample received the average of the damages awarded to the tried cases in their category.

Judge Parker argued that the result was fair to the defendants since the total amount awarded was an accurate estimate of the total that would have been awarded if all cases had been tried. He cited confidence levels ranging from 95 percent to 99 percent but did not explain what those numbers meant or what assumptions were used to calculate them.

The situation is not quite so clear as Judge Parker apparently believed. The statistical conclusions reported, if correct, depend on assumptions about the distribution of the awards that would be produced by jury trial. It is possible to describe distributions consistent with the observed data for which the result of even as large a sample as was used in *Cimino* would be a very imprecise measure of the total damages that would be awarded.[7] With such a distribution, the expected result of the *Cimino* procedure would still be correct: if the procedure were repeated a large enough number of times, the average outcome would be very close to

---

[4] The issues were whether each asbestos-containing insulation product manufactured by each defendant, settling and nonsettling, was defective and unreasonably dangerous, the adequacy of warnings, the state of the art defense and the fiber type defense. The question of punitive damages in the entire case of the 2,298 class representatives was also submitted for jury determination.

[5] Phase 2 was resolved by stipulation by the parties.

[6] The increase in sample size was less than proportional, as one would expect if the objective was to get equally reliable results for each category. The opinion states that "[w]hen setting the sample size for each disease category, the Court sought a confidence level of 95%, in other words ±2.00 standard deviations" (Cimino v. Raymark *664). The numbers (samples of 50 each for two categories with 1,050 and 972 plaintiffs) suggest that the court did not apply any very precise statistical rule.

[7] One example is a distribution in which a very small number of plaintiffs have cases that, if tried, will generate enormous damage awards. If there is only one such plaintiff, and 10 percent of the cases selected at random are tried, there is a 90 percent chance that his case will not be selected and will thus have no influence on the observed results. But if the damage award he would get is large enough, his case may have a very large effect on what the total award to all plaintiffs would be if all cases were tried.

the result of trying every case separately. But the probability that the result produced by the *Cimino* procedure would be substantially different from the result of trying all cases might be much larger than implied by the confidence levels cited by the judge.[8] This suggests that it might be worth looking for a procedure superior to random sampling.

Even if, as Judge Parker argues, the procedure is fair to the defendants, there remains the question of whether it properly allocates the damage payment among the individual plaintiffs. The procedure used in *Cimino* does not do so, as Judge Parker himself conceded.[9] He dealt with that problem by obtaining the plaintiffs' assent in advance. In future litigation involving such procedures, however, the question will be important for at least five reasons.

1. Many people regard justice as part of what litigation is supposed to produce. If a procedure collects the right amount of damages but gives them to the wrong people, or to the right people but in the wrong amounts, it is not just.

2. One purpose of some of the legal rules that determine damages, such as contributory negligence, is to affect the incentives of potential plaintiffs. In *Cimino,* some plaintiffs whose cases were tried received no damages, possibly because their decision to smoke was regarded by the jury as contributory negligence.[10] The effect of such verdicts was to reduce the award given to all plaintiffs in the same disease category whose cases were not tried, smokers and nonsmokers alike. So the use of the procedure undercuts the effectiveness of such a legal rule.

3. In order for a class to be certified, the judge must find that the representative parties will fairly and adequately protect the interests of

---

[8] Judge Parker's statement that "[d]efendants complain about the 1% likelihood that the result would be significantly different" (Cimino v. Raymark *666) suggests that he interprets a 99 percent confidence level as a probability of 99 percent that the procedure will yield a result within some (unspecified) significant error—where "significant" means "important," not "statistically significant."

Whatever error he did use, what ought he to have used? One possibility would be to compare the procedure to the result of individual trials, taking account of the difference in litigation costs. Suppose, for example, that aggregation saves the defense a million dollars in legal expenses. One might then ask how likely it is that the award is more than a million dollars greater than what would have been awarded if all cases were tried. If the answer is .01, there is then only one chance in a hundred that the procedure has made the defendants worse off. While that approach solves the problem of picking an appropriate error, it still leaves the problem that statistics cannot generate such probabilities without making assumptions about the characteristics of the sample.

[9] "Individual members of a disease category who will receive an award that might be different from one they would have received had their individual case been decided by a jury have waived any objections" (Cimino v. Raymark *665).

[10] The opinion discusses under what circumstances smoking would constitute contributory negligence and notes that some plaintiffs received awards of zero, but it does not say whether any received zero awards for that reason.

the class.[11] A procedure that predictably awards some plaintiffs more than they would get from trying their case themselves and others less may not meet the requirement.

4. Even if the class is certified, individual members are free to withdraw. A procedure that predictably awards some plaintiffs less than they would receive at trial gives such plaintiffs an incentive to withdraw from the class, which reduces the benefit both of the class action and of the procedure.[12]

5. In awarding the right amount of damages to the wrong people, the *Cimino* procedure resembles fluid recovery. Under fluid recovery, where it is difficult to identify the members of the plaintiff class and determine how much of the award each is entitled to, money awarded to the plaintiffs is instead used to benefit a group of people similar to those who were injured. That approach has been seriously questioned by the courts.[13]

For all of these reasons, it is desirable to construct procedures that approximate the correct result among plaintiffs as well as between plaintiffs and defendants. In *Cimino,* Judge Parker attempted to do so in two ways. Phase 2 of the trial was designed to eliminate from the case plaintiffs whose exposure to asbestos was not a producing cause of an asbestos-related injury or disease. In phase 3, plaintiffs were grouped according to the particular sort of injury or disease they had suffered, presumably because individuals suffering from the same disease would have some tendency to be owed the same damages.

Both of these are very imprecise ways of allocating damage payments to individual plaintiffs. Saks and Blanck offer two additional possibilities. One is to use statistical analysis to define groups with common characteristics. The other is to construct a linear model relating damages to characteristics, use trial results to estimate the parameters of the model, and then use the estimated parameters to calculate damage awards for the untried cases.[14]

---

[11] Federal Rules of Civil Procedure 23a(4).

[12] Suppose the court uses an aggregation process that awards every plaintiff the average of what all plaintiffs in the class are entitled to. Plaintiffs who can expect an above average return withdraw from the class. That lowers the average that the remainder can expect to get, causing more plaintiffs to withdraw. Under some circumstances, the entire class may come apart in this way. This is a form of adverse selection, more familiar in the context of insurance. See George Akerlof, The Market for "Lemons": Quality Uncertainty and the Market Mechanism, 336 Q. J. Econ. 488 (1970).

[13] It was permitted in David Daar v. Yellow Cab Co., 67 Cal. 2d 695, 433 P.2d 732, 63 Cal. Rptr. 724, rejected by the Second Circuit in Eisen v. Carlisle and Jacquelin, 479 F.2d 1005, and has not been ruled upon by the Supreme Court.

[14] Saks & Blanck, *supra* note 2, at 851. Glen O. Robinson & Kenneth S. Abraham, Collective Justice in Tort Law, 78 Va. L. Rev. 1481 (1992), suggests and discusses several other statistical approaches to dealing with mass torts, using information from the outcomes of similar cases to determine, or at least affect, awards.

While these procedures can improve on the simple approach of giving every plaintiff the same amount or the slightly more complicated approach implemented in *Cimino,* they suffer from a common problem. It is neither obvious in advance nor uncontroversial what characteristics are relevant to the damage award or how they are related. Even if we knew the characteristics, there is no reason to assume the relation is linear.[15] As statisticians are aware, the same data can be fit with a multitude of different specifications. If, after trying a few thousand, the court finds one that happens to fit the tried cases fairly well, that should not give us much confidence that it will also fit the untried cases.

What we need is not a procedure for dividing the damage award among the plaintiffs—the best way of doing that will almost certainly vary from case to case. What we need is a procedure that makes it in the interest of someone to figure out, for any particular case, what the correct division among the plaintiffs in that case is. Part 2 describes one such procedure.

II.   PLAINTIFFS CUT, DEFENDANT CHOOSES:[16] AN INCENTIVE-COMPATIBLE PROCEDURE FOR LITIGATING MASS TORTS

I define the strength of a plaintiff's case as the average of what would be awarded if it were tried many times by many separate juries; I call this average verdict (for plaintiff i) d_i. The objective of the procedure is to produce a damage award of about d_i for each defendant i, at a cost much lower than the cost of trying every case many times or even trying every case once.

By examining the facts relevant to an individual plaintiff i, an investigator can estimate the value of d_i. The more resources are spent on the investigation, the more accurate the estimate will be. This is true both for an individual investigator and for a judge or jury calculating an award in the course of a trial. I assume the cost to a competent individual investigator of estimating d_i is much lower than the cost of a trial that produces an estimate, in the form of a verdict, with the same accuracy, since the investigator is a specialist in such investigations, is an individual rather than a committee, and is not limited by the elaborate procedural rules that control court trials.[17]

---

[15] As Saks and Blanck point out, average jury awards seem to increase less than linearly with the amount of injury suffered by the plaintiff (*id.* at 840).

[16] In explaining my proposed procedure, I assume that it is being applied to a case with many plaintiffs and one defendant; the application to the less common case of one plaintiff and many defendants should be straightforward.

[17] One reason such rules are necessary is that the decision maker in a trial has only weak incentives to reach the correct decision and can therefore not be trusted to do so unless severely constrained. Under the proposed procedures, it is in the private interest of the decision maker (the plaintiffs' attorney) to estimate the strength of claims accurately, making such constraints less necessary.

We start with a group of N plaintiffs represented by an attorney. The procedure is as follows:

STEP 1.   The plaintiffs' attorney produces, for each plaintiff i, a claim C_i, which is the amount the attorney claims that plaintiff i ought to receive in damages. For reasons that will become clear below, it will be in the attorney's interest to make C_i proportional to his estimate of d_i.

STEP 2.   The plaintiffs' attorney gives his list of claims C_i to the defendant's attorney.

STEP 3.   The defendant's attorney selects from the list a small number of cases to be tried. For simplicity in exposition, assume that 10 cases are to be selected and that the cases selected turn out to be those of plaintiffs 1–10.

STEP 4.   These cases are tried. The court awards damages D_i to each of the 10 plaintiffs.

STEP 5.   The court calculates $R = (D_1/C_1 + D_2/C_2 + \ldots + D_{10}/C_{10})/10$, and awards damages of $R \times C_i$ to each of the N plaintiffs.

Under this procedure, it should be possible to resolve the N cases much more cheaply than with N separate trials. Only 10 cases actually have to be litigated. All plaintiffs have their damages estimated, but the estimate is made for everyone else by the plaintiffs' attorney.

Why does the procedure generate actual damages for plaintiff i close to d_i? Consider the situation first from the standpoint of the defense attorney at step 3. He wants to select plaintiffs whose claims C_i are large relative to d_i, the amount a court would, on average, award them. By selecting plaintiffs who have overclaimed, he produces a low value of R and thus reduces the total amount ($R \, \Sigma_i C_i$) his client must pay in damages.

Next consider the situation from the standpoint of the plaintiffs' attorney at step 1. Because he knows that the defense attorney will try to select for trial plaintiffs with a high ratio of C_i to d_i, he maximizes the total payments his clients receive by trying to make the ratio the same for all clients. The simplest way of doing so is to set C_i equal to his estimate of d_i for each client.[18] So the amount claimed for each plaintiff will be equal to his attorney's estimate of what he can expect to get at trial.[19]

The award received by a particular plaintiff may deviate from what he

---

[18] The plaintiffs' attorney can achieve the same objective by attempting to set C_i proportional to d_i: $C_i = K d_i$, where K is some constant. The value of K has no effect on the outcome; R, on average, will be $1/K$, so plaintiffs will receive $R \times C_i = \langle d_i \rangle$, independent of K. I therefore assume $K = 1$ for simplicity in exposition.

[19] As we will see later, this statement is only approximately true. If some cases are harder to evaluate than others, the optimal strategy for the plaintiffs' attorney may deviate somewhat from that described here.

ought to receive for two reasons: the court may give the wrong verdicts
for the cases tried, or the plaintiffs' attorney may claim the wrong amount
for a particular plaintiff. Since 10 cases are tried separately and their
results averaged, the first source of error should be much smaller than if
each plaintiff's case had been tried by itself.[20] Since the attorney can
estimate d_i much less expensively than a court, the second source of
error can be made smaller than it would be with an actual trial, while
still keeping litigation expenses (including the expense of making such
estimates) well below those of individual trials. So it should be possible
to produce a more accurate verdict at lower cost under this procedure
than with individual trials. The cost is higher than with the *Cimino* proce-
dure since additional costs are born by the plaintiffs' attorney in making
claims and the defendant's attorney in choosing cases to litigate. But this
procedure, unlike that one, generates separate results for each plaintiff
proportioned to the strength of each plaintiff's case.

The procedure as I have described it makes sense for a hundred plain-
tiffs, for a thousand, or perhaps for more. In *Cimino,* the information
actually collected included medical evaluations for about 1,400 of the
2,298 plaintiffs who eventually went to trial, so much of the research
required by my suggested procedure had actually been done. But it makes
less sense for the sort of class action that involves a very large number
of plaintiffs, most with very small claims. In such a case, evaluating each
plaintiff's case in order to decide how much to claim for him might cost
more than the total damages awarded.

One approach to such a situation would be to allow the plaintiffs' attor-
ney to state C_i for classes of plaintiffs rather than for individual plaintiffs.
Thus he might claim that each heavy smoker born before 1960 was enti-
tled to $10, each light smoker born between 1960 and 1970 to $2, and so
on. The defendant's attorney would select classes for trial; individual
cases would be selected from those classes at random. Such a variant on
the procedure might be appropriate in situations where individual claims
are low and separate estimates for each case thus unreasonably ex-
pensive.

A more sophisticated approach would combine the procedure de-
scribed here with an idea suggested by Saks and Blanck.[21] Instead of
producing a claim for each plaintiff, the plaintiffs' attorney produces a
statistical model showing how he believes that the amount each plaintiff
is entitled to should depend on the characteristics of the plaintiff. The

---

[20] This is one of the central points made by Saks and Blanck in defending the *Cimino*
procedure. Saks & Blanck, *supra* note 2, at 833–36.

[21] *Id.* at 851.

defendant's attorney specifies a sampling protocol, describing how plaintiffs are to be selected for trial based on their characteristics. The court then selects plaintiffs for trial at random, subject to the constraints of the sampling protocol. The verdicts for those plaintiffs are used to estimate the parameters of the model, and awards for all plaintiffs are calculated accordingly.

In the simplest version of this, the plaintiffs' attorney would specify the entire model save for one multiplicative parameter. If, for example, he believed that the amount awarded ought to depend linearly on the age of the plaintiff and the number of years he had worked at a site using asbestos, he might offer the model

$$\text{Damages} = A \ (\$100,000 - \$1,000 \times \text{Age in Years}$$
$$+ \$10,000 \times \text{Years Worked on Site}).$$

The defense would then specify the range of ages and work histories that were to be sampled, and the court would choose plaintiffs within that range at random. Their cases would be tried, and the results used to calculate A.

In a more elaborate version, the plaintiffs' attorney would specify only the form of the model. An example might be

$$\text{Damages} = A - B \times \text{Age in Years} + C \times \text{Years Worked on Site}$$
$$+ D \times (\text{Years Worked on Site})^2.$$

The defense would again specify the characteristics of plaintiffs to be selected for trial, and the court would choose at random plaintiffs with those characteristics.

While both variants of this approach may sound complicated, especially to nonstatisticians, their logic is the same as that of the simpler version described earlier. The difference is that the plaintiffs' attorney is providing a description of how damages relate to characteristics, rather than a claim for each plaintiff. The same logic as before makes it in his interest to get the description right. If, for example, he erroneously claims that the amount plaintiffs are entitled to does not depend on their age, when a jury would actually award more to younger plaintiffs, the defense can specify a sample heavily weighted toward older plaintiffs—and the result will be to push down the total amount awarded.

The same logic applies to more subtle errors. Suppose the plaintiffs' attorney specifies a linear relation of the form

$$\text{Damages} = A + BL,$$

where L is (say) length of exposure to asbestos. Further suppose that the real relation, the one that correctly predicts jury verdicts, is a quadratic of the form

$$Damages = A + BL^2.$$

The defense, if it recognizes the error, can specify a sample containing only small values of L. Again the result will be to push down the total verdict.

In each of these situations, just as with the simpler version of my proposed procedure discussed earlier, an inaccurate specification by the plaintiffs' attorney of the relative claims of different plaintiffs gives the defense an opportunity to reduce the total amount awarded, which in turn gives the plaintiffs' attorney an incentive to do an accurate job of specifying the relative claims.

So far I have assumed that the cases we are considering are ones where the plaintiffs seek money damages. The procedure can be generalized to any case with a quantitative award—one describable by some cardinal measure. An example would be a suit where the plaintiffs were employees claiming seniority.

Another assumption I have been making is that tort litigation under my procedure is always resolved by trial. What is the effect on the analysis if we include the possibility of settlement?

Even if there is some possibility of settlement, the plaintiffs' attorney still has an incentive to estimate the relative claims of the plaintiffs accurately. If the case goes to trial, inaccurate estimates will result in lower total damages, since the defense will select the overclaimed cases for trial. If the case settles, it will settle on less generous terms if the defense believes that the estimates are inaccurate, and the plaintiffs thus likely to do badly at trial.

So even when litigation leads to settlement, the procedure still provides a mechanism for allocating damage payments to plaintiffs that reflects the relative strength of their cases. By doing so, it should reduce the conflict among plaintiffs over settlement terms and so make settlement easier.

III. The Legal Status of the Proposed Procedure

Suppose a judge wished to implement the procedure described in Part II above. What legal problems would he face?

To begin with, he would face the same problem faced in *Cimino:* the argument that due process required that each plaintiff have an opportunity to make his case in court and that the defendant should have the

opportunity to rebut each plaintiff's case. If, as in *Cimino,* the plaintiffs assented in advance to the procedure,[22] the argument should be no stronger here than there. If anything, the defendant's grounds for objection are even weaker under the procedure I have proposed. Insofar as the defense believes that some plaintiffs have weak cases—weaker cases, relative to other plaintiffs, than their claims indicate—the defense is free to select those cases for trial.[23]

If the plaintiffs, instead of or in addition to the defendant, object, the situation is somewhat more difficult. While the procedure saves the plaintiffs the cost of litigating every case separately, it also, for reasons I will discuss in Part IV, has some built-in bias against the plaintiffs. The plaintiffs might reasonably demand either that the procedure be modified to eliminate that bias (a possibility discussed below) or that they be compensated for accepting a biased procedure. Supposing that such objections were met, the plaintiffs under my procedure seem to be in the same situation as the defendants in *Cimino;* although their cases are not all being tried, they are being given an opportunity to get approximately the same awards they would get if they were tried, and at a much lower cost in litigation.

There is one respect in which the procedure is more defensible than that employed in *Cimino*—or, arguably, than the ordinary procedure for a class action. Rule 23(a)(4) of the Federal Rules of Civil Procedure requires that the representative parties in a class action will fairly and adequately protect the interests of the class. Under the procedure I have proposed, the representative parties have a clear interest in doing so. If they attempt to benefit themselves at the expense of other members of the class by arranging for their attorney to overclaim on their behalf, the defense will select their cases for trial.[24] The representative parties will

[22] Since the class was certified before the procedure was proposed, the assent was presumably by the representative plaintiffs controlling the litigation rather than by the unanimous decision of all plaintiffs. But the procedure created a conflict of interest among members of the class, which arguably called into question the ability of the representative plaintiffs to represent the interests of the remaining plaintiffs.

[23] For an extensive discussion of legal issues associated with aggregation, see Robinson & Abraham, *supra* note 14.

[24] If it is not obvious that they are overclaiming, the defense may miss some of their cases, in which case some of the overclaimed representative defendants will get more than they should. On the other hand, given that possibility, one would expect the defense to take special care in examining the claims made for the representative plaintiffs. I am assuming here that plaintiffs whose cases are actually tried get the amount awarded to them, rather than having their award calculated from their claim in the same fashion as plaintiffs whose cases are not tried. Without that assumption, representative plaintiffs gain by overclaiming even if they are sure their cases will be among those tried—although the attorney for the class of plaintiffs loses, if his recompense is an increasing function of the total amount awarded.

gain nothing, and their attorney will have a lower total award out of which to compensate himself.

IV. PROBLEMS WITH THE PROCEDURE

There are two fundamental problems with the procedure I have described. The first is that, while it could produce a more accurate result at a much lower cost than would individual trials, it is not entirely clear that it will; it might instead produce a much more accurate result at a higher cost. The second is that the procedure, as so far described, has a built-in bias in favor of the defense.

Does It Save Money?

> The method incorporated into phase III produces a level of economy in terms of both judicial resources and transaction cost that needs no elaboration. [JUDGE PARKER, *Cimino*]

At first glance it seems obvious that trying 160 cases costs a great deal less than trying 2,298 cases, but this is not quite so clear as it seems. Under the procedure employed in *Cimino,* the verdicts in the tried cases determined the outcome for all of the other cases. The result is that the amount at stake in each tried case was about 14 times as much as it would have been if each case had only determined the outcome for that plaintiff. With more at stake, we would expect both parties to spend more on trying to win.

Whether this eliminates the cost savings of fewer trials depends on how litigation expenditure varies with the amount at stake.[25] If the increase is

---

[25] I do not know of any definitive analysis of this question. One possible approach would be to assume Nash equilibrium. The amount at stake is S. The probability that the plaintiff will win the case depends on expenditures L_p (by the plaintiff) and L_d (by the defendant): $P(L_p, L_d)$. The parties increase their expenditures until they reach the point where a \$1 increase in expenditure by the plaintiff increases his expected return $(P(L_p, L_d)S)$ by \$1, and a \$1 increase in expenditure by the defendant increases his expected return $(-P(L_p, L_d)S)$ by \$1:

$$\frac{\partial P(L_p, L_d)}{\partial L_p}S = 1, \quad \text{and} \quad \frac{\partial P(L_p, L_d)}{\partial L_d}S = -1.$$

Under these assumptions, the question of how expenditure increases with amount at stake becomes the question of how rapidly

$$\frac{\partial P(L_p, L_d)}{\partial L_{p,d}}$$

decreases as L_p and L_d increase. If, for example,

$$\frac{\partial P(\alpha L_p, \alpha L_d)}{\partial L_{p,d}} = \frac{1}{\alpha^\beta} \frac{\partial P(L_p, L_d)}{\partial L_{p,d}},$$

proportional, the total cost of trials under either *Cimino* or the procedure I have suggested will be the same as if every case were tried separately; the only advantage of the procedure would then be the increased accuracy, due both to trying cases much more carefully and to using the average of the tried cases, rather than the result of one case, in calculating the amount to be awarded to each plaintiff whose case is not tried.[26]

Suppose, however, that expenditure rises less than proportionally with the amount at stake, everything else held constant.[27] Under that assumption, expenditure on the tried cases becomes less and less important as the number of plaintiffs increases since the larger the number of plaintiffs the smaller the fraction necessary to provide an adequate sample. In the limit of a very large number of plaintiffs, expenditure on trying the sample of cases is negligible compared to the cost of trying the cases individually. That is consistent with what actually happened in *Cimino*.

So far I have been considering a problem raised by both the *Cimino* procedure and the procedure I have proposed. There is an additional cost problem that applies only to the latter. Under that procedure, the plaintiffs' attorney spends resources estimating the relative claims of each plaintiff,[28] and the defendant's attorney then spends resources examining plaintiffs in order to decide which cases to select for trial.

Under our assumptions, the plaintiffs' attorney can produce his estimates of claims more accurately and less expensively than verdicts would be produced by individual trials. The same should be true for the defense attorney. In addition, if the number of cases is large, the defense need only examine a random sample of cases in order to do a reasonably good job of locating overclaimed cases to select for trial. It follows that the

---

then expenditure increases more (less) than in proportion to the amount at stake if $\beta < 1$ ($\beta > 1$).

One objection to this approach is that Nash equilibrium is not very plausible in a game involving only two parties, and it is still less plausible in a situation where the two parties can and do bargain with each other.

[26] In the case of the *Cimino* procedure, that must be balanced against the decreased accuracy from awarding plaintiffs whose cases are not selected for trial average verdicts even though the particular plaintiff may not have average characteristics.

[27] The comparison is between two cases whose only difference is the amount at stake; each of my 10 cases is simply one of the thousands of cases that might be tried individually. I am not assuming that the ratio of litigation cost to the amount at stake for the typical large case is smaller than for the typical small case; presumably the typical large case not only has more at stake but also a more complicated set of legal and factual issues than the typical small case.

[28] Or the plaintiff's attorney spends resources determining how the amount a plaintiff is entitled to is related to the plantiff's characteristics, under the alternative version that I proposed for cases with very large numbers of plaintiffs and small average claims.

attorneys can act in a way that, under the proposed procedure, produces more justice at a considerably lower cost than would individual trials.

It is not, however, clear that it is in their interest to do so. Each attorney's objective, at least in part, is to benefit his clients at the expense of the other party. By making a more accurate set of estimates, the plaintiffs' attorney not only produces a more just distribution among his clients, he also makes it harder for the defense to locate overclaimed plaintiffs for trial. The more he spends on improving the accuracy of his claims, the larger the amount his side will receive. He must balance that benefit against the associated cost. The defense attorney faces a similar situation.

Here, as elsewhere in the economics of litigation, there is no reason to assume that the level of expenditure that is privately optimal for one party to a legal dispute is also socially optimal. The amount spent on estimating claims and detecting overclaimed plaintiffs will depend on detailed assumptions about information costs and distributions of claims, as we will see in the formal analysis presented in Appendix B and discussed in Part VI.

It follows from these arguments that we cannot be sure the procedure as described will cost less than ordinary trial without aggregation and sampling. This suggests two further queries. The first is whether we can say anything interesting about the relation between the costs of alternative approaches and the number of plaintiffs. The second is whether, if experience suggests that expenditures associated with the procedure are undesirably large, there may be ways of modifying it to reduce such expenditures.

An increase in the number of plaintiffs reduces the percentage of cases that must be tried. If expenditure per case increases less than proportionately with the amount at stake, the result is that trial costs for my suggested procedure (or the *Cimino* procedure) decrease, relative to the cost of trying all cases, as the number of plaintiffs increases.

The same is probably true for the cost to the defense of selecting cases for trial. The larger the number of plaintiffs, the smaller the fraction that must be sampled in order to find 10 cases from (say) the most overclaimed 5 percent. We would expect defense expenditures to increase less than proportionally with the number of plaintiffs, and so become smaller and smaller, relative to the total amount at stake, as the number of plaintiffs increases. This result is demonstrated in Appendix B for the particular distributions assumed there.

The opposite result can be expected for the cost to the plaintiffs' attorney of calculating claims. The more plaintiffs there are, the easier it is for the defense to locate those who have overclaimed. The more accu-

rately the defense can locate overclaimed plaintiffs, the greater the incentive for the plaintiffs' attorney to make accurate claims. So an increase in the number of plaintiffs will tend to increase the amount spent per plaintiff by the plaintiffs' attorney. That is one reason why it might be desirable to shift from individual claims to statistical models when the number of plaintiffs becomes sufficiently large. The per-plaintiff cost of estimating the parameters of a model to a given accuracy will fall as the number of plaintiffs increases.

The size of the expenditures by the attorneys will depend on details of the distribution of claims and on the functions relating expenditure on investigating a claim to information produced. We cannot predict a priori how large it will be, any more than we can predict a priori, in the case of ordinary litigation, how much of the damages awarded will be eaten up in litigation costs. But if experience indicates that the attorneys are spending more than the improved accuracy their expenditure generates is worth, we can lower the amount they spend by a minor change in the procedure.

The incentive for the expenditures we (hypothetically) wish to reduce comes from their influence on the damages that will be awarded.[29] A court that wishes to reduce those expenditures can do so by selecting some cases for trial in the fashion I have described and some at random. The smaller the proportion of cases selected for trial by the defense, the lower the incentive that both attorneys have to spend more money estimating claims more accurately. Thus courts have a mechanism by which they can adjust the procedure to move its outcome closer to an optimal level of cost and accuracy. An alternative approach would be to try to impose limits on the amount each party was permitted to spend on evaluating claims.[30]

Bias in the Procedure: Who Cuts, Who Lies, and Other Fine Points

In the procedure as I have described it, the plaintiffs' attorney calculates claims, and the defendant's attorney selects which will be tried: the former cuts and the latter chooses, to take the obvious analogy from the incentive-compatible procedure for dividing a piece of cake. Is there any

---

[29] It is possible that the plaintiffs' attorney may have additional incentives, due to concerns with either justice or risk among his clients. They might prefer that claims be proportional to the actual injury each client has suffered, even if claims did not affect the total amount paid out.

[30] In a class action, a judge could limit expenditure by one side simply by limiting the expenses he was willing to permit the class attorney to claim. Limiting expenditure by the other party, or by both parties if the procedure was being used outside of a class action, would be more difficult.

good reason to do it this way, instead of requiring the defendant's attorney to list the amount he believes each plaintiff should receive and letting the plaintiffs' attorney choose which cases will be tried?

One reason is that the attorney who is calculating claims will need information from the plaintiffs that they might be reluctant to provide to the defense attorney, for fear that it would be used against them in trial. The procedure I have described does not eliminate this problem—the defense attorney still needs enough information to decide which cases to select for trial. But, if the group is large, he can do an adequate job by examining only a small subset of the plaintiffs and can thus afford to spend much more per case examined than the plaintiffs' attorney. That should make it possible for him to produce a reasonably accurate estimate even with less cooperation from the individual plaintiff.[31]

The defense attorney is not the only one who must worry about being misled by individual plaintiffs. Plaintiff i gains by increases in C_i above d_i, even though the plaintiffs as a group lose, so each plaintiff has an incentive, in dealing with his own attorney, to inflate his claim. A plaintiff's attorney would presumably specify in his contract with the plaintiffs, whether representative plaintiffs in a class action or joint plaintiffs in an ordinary joint action, their obligations to furnish information that he requires in estimating their claims. Thus the procedure yields a contractual equivalent of discovery rules between the plaintiffs and their attorney. Although each plaintiff gains by his own ability to mislead his attorney, he loses by the ability of all other plaintiffs to do the same, so plaintiffs and their attorney have a common interest in agreeing to rules that will allow the attorney to make an accurate estimate of the strength of each plaintiff's case.

One consequence of having the plaintiffs' attorney cut and the defendant's attorney choose is to give the latter a cost advantage, at least in situations where the number of plaintiffs is large. As discussed earlier, the party who chooses can use random sampling to identify overclaimed cases at a relatively low total cost. This advantage may or may not outweigh the advantage that the plaintiffs' attorney has, due to the fact that the plaintiffs, who possess private information relevant to the strength of

---

[31] Presumably there would be legal rules requiring some cooperation from the plaintiffs. "[M]ost courts have taken the view that reasonably necessary discovery against individual class members should be allowed as a matter of judicial discretion, but that discovery is not available of right as it would be against a party to a nonclass suit (see, e.g., Brennan v. Midwestern United Life Ins. Co., 450 F.2d 999 (7th Cir. 1971))." Fleming James, Jr., & Geoffrey C. Hazard, Jr., Civil Procedure 579 (3d ed. 1985).

their cases, are his clients and have agreed to make such information available to him.

A second consequence is to give the defense an advantage in the final verdict. As I show in Appendix A, the defense can produce an expected total damage payment equal to the expected payment under a system of individual trials by simply selecting cases for trial at random, with probabilities proportional to C_i. By examining cases and selecting those that appear to be overclaimed, the defense should be able to improve on that result.

How significant these advantages are will depend on the details of the underlying fact-finding technology—how accurately and at what cost each attorney can estimate d_i. If the net advantage to the defense turns out to be large,[32] and if we wish neither to change the tort system in a way that advantages defendants in mass torts nor to give plaintiffs an incentive to avoid the procedure in favor of individual litigation, we could compensate by altering other legal rules applicable to the procedure in ways that advantage plaintiffs.

An alternative approach would be to eliminate the bias by allowing both parties to cut and both to choose. Under such a system, the plaintiffs' attorney produces a set of claims C_i^p, and the defense produces a set of claims C_i^d. Each attorney selects a set of cases to be tried. The court calculates two values of R: R_p is calculated using the plaintiffs' claims and the verdicts of the cases selected by the defense; R_d is calculated using the defense claims and the verdicts of the cases selected by the plaintiffs' attorney. Each plaintiff i receives the average of $R_p C_i^p$ and $R_d C_i^d$. This version of the procedure will cost more to produce a given level of accuracy in the relative claims since each is being calculated twice. But it eliminates the bias in the outcome.[33] It may or may not increase the total cost of the procedure. Since each set of calculations plays only half the role it did before in determining the amount actually

---

[32] This does not require the defense attorney to be better at estimating $\langle d_i \rangle$ than the plaintiffs' attorney, as should be clear from the analysis above. If, for example, accurate estimates are very expensive and the number of plaintiffs is large, the plaintiffs' attorney will produce very inaccurate estimates and the defendant's attorney, spending much more per case on a small fraction of the cases, will be able to find cases that are greatly overclaimed, thus greatly reducing the total amount paid out in damages.

[33] That conclusion depends on assuming that both sides are equally able to generate the relevant information. If, as suggested earlier, the plaintiffs' attorney has better access to information about plaintiffs, the version of the procedure described here is biased in favor of the plaintiffs. If one knew how great the informational advantage was, one could compensate for it by using an appropriately weighted average of the awards calculated from the two different sets of claims.

awarded, the parties have an incentive to spend less than before on increased accuracy. That may or may not balance the increased cost of having each claim calculated twice and having each party try to identify cases that the other has overclaimed.[34]

V. APPLICATIONS OUTSIDE OF CLASS ACTIONS

My analysis so far has assumed that the procedure I am describing will be used, as the *Cimino* procedure was, in a class action. It might also be applied to an ordinary joint action with a large number of plaintiffs. The use of the procedure ought to make such a joint action easier to organize since it provides a mechanism for solving the problem of allocating damages among the joint plaintiffs. After a putative mass tort had occurred, one or more lawyers would announce that he was forming a group of plaintiffs to litigate under the procedure; his announcement would include the formula by which he would be reimbursed. Plaintiffs would be free to join his group, to join another group, or to litigate individually.

My discussion has focused on mass torts because the procedure requires a single agent representing the defense and a single agent representing the plaintiffs. A lawyer who assembles a group of plaintiffs for a joint action satisfies the second requirement; the fact that all of the plaintiffs are suing the same defendant satisfies the first. This raises the question of whether other ways of satisfying these requirements might make it possible to use the procedure to reduce litigation costs outside of the context of mass torts. Consider the following radical proposal:

A court bundles the cases before it into large groups, defined by common characteristics—a thousand intellectual property cases, a thousand personal injury cases, a thousand defamation cases. Each group is then auctioned off twice, with attorneys bidding for the right to represent all plaintiffs and for the right to represent all defendants. In the former auction, the attorney is offering to pay a sum in exchange for the right to represent the plaintiffs and collect all damages awarded to them: high bid wins. In the latter, the attorney is stating for what sum he will agree to represent the defendants and pay all damages awarded against them: low bid wins.

The two winning attorneys then go through the procedure I have described. When it is over, the defense attorney pays the plaintiffs' attorney

---

[34] If we assumed that each party aimed at the same level of accuracy as under the earlier version of the procedure, expenditure would be increased but not doubled by requiring each party to both cut and choose. The information generated in cutting can also be used in choosing. The defense attorney's first step in identifying overclaimed cases will be to compare the plantiffs' attorney's claims with his own.

the total damages awarded: $R \times \Sigma_i C_i$. Each defendant pays the defense attorney

$$\left(C_i \Big/ \sum_i C_i \right) \times B_d,$$

where B_d is the attorney's bid: the amount for which he agreed to be responsible for all costs and damages. Thus the defendants are dividing their total costs in proportion to the amounts owed to their respective plaintiffs. Similarly, the plaintiffs' attorney pays each plaintiff i an amount

$$\left(C_i \Big/ \sum_i C_i \right) \times B_p,$$

where B_p is the plaintiffs' attorney's bid: the amount he offered to pay for the right to collect all damage payments:

There are obvious problems to implementing this radical version of my proposal within our legal system since it deprives both plaintiffs and defendants of the right to choose their own attorneys. One solution would be to treat it as a form of alternate dispute resolution: cases go into groups subject to the procedure only if both plaintiff and defendant agree.

The general procedure could also be used in situations other than class actions where a single agent already controls what are really multiple cases. One example would be disputes between insurance companies, each of which controls a large number of legal claims for accidents involving its customers. In that context, the procedure would be a way of guaranteeing to each customer that the insurance company was fairly representing his interests in the litigation.

The procedure would be inappropriate if the agent who controlled multiple cases also fully owned them. Such an agent would care about the total awarded to all of the cases he owns, not the distribution among them. The *Cimino* procedure would give the correct total at a lower cost than the procedure discussed here. Such a situation could occur in the insurance context. It might also arise if, as some writers have suggested, tort claims were made fully marketable, allowing legal entrepreneurs to buy up large numbers of related claims and litigate them en masse.[35]

---

[35] For discussions of the idea of marketable claims, see Marc J. Shukaitis, A Market in Personal Injury Tort Claims, 16 J. Legal Stud. 329 (1987); David Friedman, Private Creation and Enforcement of Law: A Historical Case, 8 J. Legal Stud. 399 (1979); David Friedman, "What Is Fair Compensation for Death or Injury? 2 Int'l Rev. L. & Econ. 81 (1982); Jonathan R. Macey & Geoffrey P. Miller, The Plaintiffs' Attorney's Role in Class Action and Derivative Litigation: Economic Analysis and Recommendations for Reform, 58 U. Chi. L. Rev. 1 (1991).

Under such institutions the damage award would reach the victim in the form of the price for which he sold his claim, so the distribution among victims would be determined by the market rather than directly by the court. In this context, the radical version of the procedure described above can be seen as an alternative way of selling claims designed to eliminate the cost of separately bargaining over each transaction.

VI.   RESULTS OF THE FORMAL MODEL

Appendix B presents a formal model, based on an error distribution that is bounded and uniform. I demonstrate that, as the number of cases goes to infinity, the defense is able to perfectly identify overclaimed cases at a cost that is vanishingly small relative to the amount at stake. The plaintiffs maximize their net return by spending the same amount in investigating each case and claiming an amount equal to their estimate of the expected return at trial.

The result becomes more complicated if we assume that some cases are more difficult to evaluate than others. The optimal strategy is then to estimate those cases less accurately, insuring against the risk that the resulting estimate may be too high by deliberately claiming less than their estimated value.

Several further points are worth noting about this situation. The first is that cases that are difficult for the plaintiffs' attorney will also be difficult for the defense attorney, so the defense has an incentive not to examine those cases. The lower the probability that a certain sort of case will be examined, the less the risk of overclaiming for such cases, so this effect will work in the opposite direction from that demonstrated in the model.[36]

A second point arises if plaintiffs are risk averse. Cases that are difficult for the attorneys are also difficult for the court, so plaintiffs with hard cases face a bigger gamble if they go to court individually and thus gain more by replacing that gamble with the more certain outcome generated by the procedure I have proposed. In addition, hard cases are likely to be more expensive to litigate, again making the procedure particularly attractive as a substitute for individual trial to plaintiffs with hard cases. So even if the procedure gives plaintiffs with hard cases somewhat less than their expected return at trial, that may not make them less willing to join the class than plaintiffs with easy cases.

If, despite these considerations, the incentive to underclaim hard cases

---

[36] That is not true for the formal model of Appendix B in the limit of large N because, in that situation, the defense is able to perfectly identify overclaimed cases at negligible cost.

turns out to be a serious problem, it can be dealt with in the same way earlier suggested for dealing with the procedure's pro-defense bias. The analysis of strategies with regard to hard cases is symmetrical; if the defense cuts and the plaintiffs' attorney chooses, the defense has an incentive to overclaim hard cases. So if both parties cut and both choose, the biases will tend to cancel.

One important limitation of the formal model of appendix B is that its error distributions are bounded. The result is that, as the number of cases increases, the additional gain to the defense of more and more accurately identifying the overclaimed cases becomes less and less; there are no cases to be found that are overclaimed by more than a factor of $1 + \epsilon_p$. If the error distribution for the plaintiffs' estimates is unbounded, and if the defense can make the error of its estimate as small as it likes by spending enough money examining enough cases, it is in the interest of the defense to push R further and further down the larger the number of cases, so that in the limit of an infinite number of cases damages awarded would go to zero.

How serious a problem this is likely to be with plausible numbers of cases and error distributions is an empirical issue. If it does turn out to be a problem, it might be controlled by any of several modifications to the procedure suggested earlier.

VII. CONCLUSIONS

I have proposed a procedure that has the potential to settle mass torts at a cost much less than individually litigating each claim. Like the *Cimino* procedure, it produces about the same outcome for the defense as would individual trials. Unlike the *Cimino* procedure, it provides outcomes for the individual plaintiffs tailored to the strength of their individual cases; indeed, it may well produce a more accurate allocation of damage payments to plaintiffs than would individual trials.

One can imagine applying the procedure in a variety of different contexts. In a case such as *Cimino v. Raymark,* where there are a large number of plaintiffs each with a substantial claim, individual attorneys might compete to form groups to litigate under the procedure, thus avoiding some of the usual problems with class actions. Where individual claims were smaller, the class could be formed in the usual way; the procedure[37] would then provide a way of allocating damages among plaintiffs. By reducing the risk that the plaintiffs' attorney would sacrifice the interests of the absent plaintiffs to his own interest and that of the

[37] Possibly the statistical version discussed above.

representative plaintiffs, the procedure makes it more likely that a class would, and should, be certified in such situations.

The procedure is not perfect; it provides no guarantee of an optimal expenditure on evaluating cases in order to allocate damages. This is equally true of alternatives, including the alternative of litigating each case separately. Also, although the plaintiffs' attorney will find it in his interest to make his claims roughly proportional to the strength of the individual cases, the relation will not be exact; differences in the difficulty of evaluating cases may, as demonstrated in the formal model, make it in his interest to deliberately underclaim some cases relative to others.[38] Finally, the simpler versions of the procedure are to some degree biased in favor of the defense since the plaintiffs cut and the defense chooses. If such problems prove serious, there are ways in which the procedure can be modified to reduce them.

APPENDIX A

A SIMPLE STRATEGY FOR THE SECOND MOVER

Suppose we are dividing a cake under the conventional rule of "I cut, you choose." Further suppose that we have identical tastes; each of us prefers the larger slice. It seems obvious that, if there is any inaccuracy in cutting cakes, the party who moves second has the advantage. One way of seeing this is to note that if he selects his slice at random he will, on average, get half the cake; if he has any ability at all to recognize the larger piece, he will do better than that. An analogous argument implies that, under the procedure described in this article, the defense can always do at least as well as it would with individual trials, and may be able to do better. The analysis goes as follows:

Suppose that, instead of examining cases and trying to select the ones that are overclaimed, the defense simply selects cases by a random process, with a probability

$$p_i = \frac{C_i}{\sum_i C_i}$$

of selecting case i. We then have

Expected Total Damage Payment $= \langle R \rangle \times$ Total Claims

[38] I have presented this as a fault of the procedure, but it could be viewed as a desirable consequence. It may be desirable, on grounds of either efficiency or justice, for parties who insist on litigating difficult cases to bear part or all of the cost of doing so. Underclaiming difficult cases costs plaintiffs less than estimating the strength of their case as accurately as easy cases are estimated. Thus this feature of the procedure has consequences similar to those of the usual (American) rule that each party must bear his own litigation costs: plaintiffs with cases that are expensive to litigate take home less, net of litigation costs, than parties who have suffered similar damage but have easy cases.

$$= \left[\sum_i p_i \frac{\langle d_i \rangle}{C_i} \right] \sum_i C_i = \left[\sum_i \frac{C_i}{\sum_j C_j} \frac{\langle d_i \rangle}{C_i} \right] \sum_i C_i = \sum_i \langle d_i \rangle,$$

which is the expected total damage payment with individual trials.

So a random procedure, with no examination at all, produces as good a result for the defense as individual trials. By selecting cases that are overclaimed, the defense can get a better result than that—a lower expected total damage payment—at some cost. If the cost is less than the gain, the defense does better under this procedure than with individual trials. If the cost is greater than the gain for all levels of expenditure on examining cases, the defense follows the strategy described above and does as well as it would with individual trials.

APPENDIX B

THE FORMAL MODEL

There are N plaintiffs. Each plaintiff i has an expected result at trial d_i. The distribution of d_i is described by a probability density $\rho(d)$ and is the same for all i. Its expected value is $\langle d \rangle$. All parties are risk neutral, so the plaintiffs' attorney is trying to maximize expected damage payments net of expenditures on investigating cases (in order to decide how much to claim for each) and litigating them, while the defense attorney is trying to minimize expected damage payments plus defense expenditures on investigating cases (in order to select some for trial) and litigating them. A number N_t of cases will be selected for trial by the defense.

Either attorney can generate an estimate of d_i by spending an amount E_i investigating that plaintiff's case. The value of the estimate will be

$$\hat{d}_i = d_i(1 + e_i). \tag{B1}$$

Here e_i is a random error, uniformly distributed between $-\epsilon(E_i)$ and $+\epsilon(E_i)$.

The more the attorney spends on investigating the case, the more accurate the estimate

$$\frac{d\epsilon(E_i)}{dE_i} < 0. \qquad \text{(Assumption B1)}$$

Investigation is subject to diminishing returns—additional expenditures yield less and less reduction in error:

$$\frac{d^2\epsilon(E_i)}{dE_i^2} > 0. \qquad \text{(Assumption B2)}$$

There is no limit to how accurate the investigation can be if the attorney is willing to spend enough—he could, for example, stage repeated dummy trials. So,

$$\lim_{E_i \to \infty} \epsilon(E_i) = 0. \qquad \text{(Assumption B3)}$$

Finally, I assume that the prior distribution $\rho(d)$ is sufficiently flat, its support sufficiently wide, and $\epsilon(E_i)$ sufficiently small so that the conditional distribution

$\rho(d_i | \hat{d}_i)$ is essentially uniform between $d_i = \hat{d}_i/[1 + \epsilon(E_i)]$ and $d_i = \hat{d}_i/[1 - \epsilon(E_i)]$.[39]

My objective is to describe a Nash equilibrium, a pair of strategies such that each party's strategy is optimal against the strategy of the other party. I will start by deriving the plaintiffs' strategy in the limit of large N, then derive the defense strategy in the general case, then use that argument to derive the plaintiffs' strategy in the general case. My objective is to provide something more than a sketch but short of a full-blown proof; I will not, for example, demonstrate that the solution I offer is unique.

THE PLAINTIFFS' STRATEGY: THE LIMIT OF LARGE N

Consider the situation from the standpoint of the plaintiffs' attorney deciding how much to spend examining each case. He plans to spend an amount E_i examining each case i, then make a claim C_i based on $\hat{d}_i$, the estimate of d_i he produces. How much should he spend on each case, and how should he then calculate C_i?

As N goes to infinity, the defense, as we will see below, can perfectly identify the most overclaimed cases,[40] so the cases selected for trial will be those for which d_i/C_i is minimal.

The plaintiffs' attorney wishes to maximize the net gain to his clients. Since the number of cases being tried is determined by the rules of the procedure, not by the attorneys, we take expenditure for trial as fixed.[41] So the attorney minimizes

Damage Payment Received − Expenditure on Examination

$$= R \sum_i C_i - \sum_i E_i = \frac{1}{N_t} \left[\sum_{i \in \{t\}} \frac{D_i}{C_i} \right] \sum_i C_i - \sum_i E_i.$$

Here D_i is the verdict from a jury trial of case i, and $\{t\}$ is the set of cases tried. Averaging this over many trials, and taking advantage of the fact that $d_i = \langle D_i \rangle$, where D_i is the result of a single trial and $\langle D_i \rangle$ its expected value, we have

[39] The reason we need to assume a sufficiently flat prior distribution in order for this to be true is that the conditional probability will depend both on the distribution of the error and on the prior distribution, in a fashion described by Bayes's theorem. The reason we must assume that $e(E_i) \ll 1$ is that otherwise a uniform distribution of e will not give something close to a uniform distribution of $1/(1 + e)$. As should become clear, these are simplifying assumptions designed to keep the formal analysis manageable, not likely to have much effect on the (qualitative) results.

[40] More precisely, the defense can identify the most overclaimed cases that exist with positive probability. There could be (say) a single case that was overclaimed by more than any other but that the defense missed because it was not in the sample examined. As N goes to infinity, the probability and, hence, the effect on average damages collected of any single case goes to zero. Any kind of case with positive probability will be represented an infinite number of times in the total (as N goes to infinity) and thus will be included in the sample selected for examination by the defense.

[41] The strategy followed in selecting cases might have some effect on the two sides' incentives to spend money litigating them. Since I have no theory of litigation expenditures, and in any case expect them to become insignificant relative to the total amount at stake when N becomes sufficiently large, I ignore this possibility in the analysis.

Expected Value of Damages − Expenditure on Examination

$$= \frac{1}{N_t} \left[\sum_{i \in \{t\}} \frac{d_i}{C_i} \right] \sum_i C_i - \sum_i E_i = \text{net benefit to plaintiffs (nbp)}.$$

In the large number situation, we can think of the cases as grouped; each group contains cases with common values[42] of E_i, $\hat{d}_i$, and C_i. As N goes to infinity, the number of cases in each group expands without limit. The defense can then select, if it wishes, N_t cases from those within a group, each of which has the minimal value of (d_i/C_i). It follows that all that matters about a group, so far as its effect on R is concerned, is R_i, defined as the minimal value of (d_i/C_i) for cases in the group. Under our assumptions, the lowest value of d_i consistent with a given pair $(\hat{d}_i, E_i)$ is $\hat{d}_i/[1 + \epsilon(E_i)]$, so $R_i = (\hat{d}_i/C_i)\{1/[1 + \epsilon(E_i)]\}$.

Imagine that there are two groups, i and j, such that $R_i > R_j$. Since the defense is trying to make R as low as possible, cases from group i will not be selected for trial. So the plaintiffs' attorney can increase his clients' total damage payment by increasing C_i. It follows that, if nbp is being maximized, $R_i = R_j$ for every pair i, j. It follows that (still in the limit of large N)

$$R_i = (\hat{d}_i/C_i)\{1/[1 + \epsilon(E_i)]\} = R.$$

Since E_i is chosen and spent before $\hat{d}_i$ is observed, the plaintiffs' attorney does not have the option of making E_i depend on $\hat{d}_i$. He could, however, use different values of E_i for different cases and, for each observed $\hat{d}_i$, set

$$C_i = (\hat{d}_i/R)\{1/[1 + \epsilon(E_i)]\}, \tag{B2}$$

thus making $R_i = R$. Averaging over many repetitions of the trial, and taking advantage of the fact that the average value of $\hat{d}_i$ is d_i and the average value of d_i is $\langle d \rangle$, yields

$$\langle \text{nbp} \rangle = \sum_i \left\{ \frac{\langle d \rangle}{[1 + \epsilon(E_i)]} - E_i \right\}. \tag{B3}$$

There is some value of E_i that maximizes the expression

$$\frac{\langle d \rangle}{[1 + \epsilon(E_i)]} - E_i;$$

call that value E^*. We have, for $E = E^*$,

$$\frac{1}{\langle d \rangle} = -\frac{\dfrac{d\epsilon(E)}{dE}}{[1 + \epsilon(E)]^2}. \tag{B4}$$

The plaintiffs' attorney maximizes the net expected return for his clients by setting $E_i = E^*$ for all i. Having done so, $C_i = (\hat{d}_i/R)\{1/[1 + \epsilon(E^*)]\}$ for all i, so C_i is proportional to $\hat{d}_i$. As pointed out earlier, the optimal values for C_i are arbitrary up to a multiplicative constant, here represented by R. For simplicity we may set $R = 1/[1 + \epsilon(E^*)]$, making $C_i = \hat{d}_i$.

---

[42] Or, more precisely, a narrow range of values. As N goes to infinity, we can make the range as narrow as we wish.
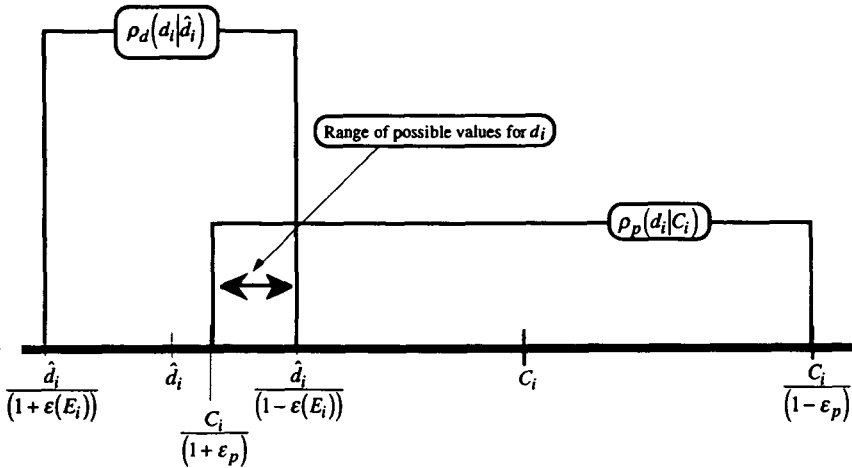
FIGURE B1

THE DEFENSE STRATEGY

We retain assumptions B1–B3 above. We further assume that the plaintiffs' strategy is as described earlier: $C_i = \hat{d}_i$, $E_i = E^*$, for all i. In order to avoid confusion between variables corresponding to plaintiffs and defense, we relabel E^* as E_p, $\epsilon(E^*)$ as ϵ_p. We are now free to use E_i, e_i to refer to defense expenditures and defense errors. Suppose the plaintiffs have made a claim C_i for case i. The defense spends E_i to estimate d_i and gets a value $\hat{d}_i$. Figure B1 shows the situation, including the probability distributions for d_i implied separately by the observations of plaintiffs and defense. The implication of these two distributions is that d_i is distributed uniformly between $C_i/(1 + \epsilon_p)$, the lowest value consistent with the plaintiffs' estimate, and $\hat{d}_i/[1 - \epsilon(E_i)]$, the highest value consistent with the defense estimate. It follows that the expected value of d_i/C_i, conditional on C_i and $\hat{d}_i$, is

$$\left\langle \frac{d_i}{C_i} \middle| C_i, \hat{d}_i \right\rangle = \frac{1}{2(1 + \epsilon_p)} + \frac{\hat{d}_i}{C_i} \frac{1}{2[1 - \epsilon(E_i)]} \tag{B5a}$$

if

$$\frac{\hat{d}_i}{[1 + \epsilon(E_i)]} \leq \frac{C_i}{(1 + \epsilon_p)},$$

and

$$\left\langle \frac{d_i}{C_i} \middle| C_i, \hat{d}_i \right\rangle = \left[\frac{\hat{d}_i}{C_i} \right] \left[\frac{1}{2[1 - \epsilon(E_i)]} + \frac{1}{2[1 + \epsilon(E_i)]} \right], \tag{B5b}$$
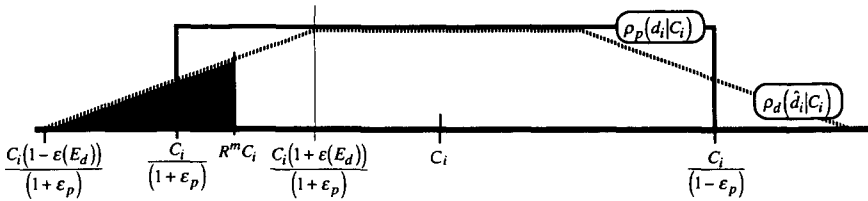
otherwise.

$$\frac{C_i(1-\epsilon(E_d))}{(1+\epsilon_p)} \qquad \frac{C_i}{(1+\epsilon_p)} \qquad R^m C_i \quad \frac{C_i(1+\epsilon(E_d))}{(1+\epsilon_p)} \qquad C_i \qquad \frac{C_i}{(1-\epsilon_p)}$$

FIGURE B2

Next consider the distribution of outcomes that the defense can expect from spending E_i examining a case i. The situation is shown in Figure B2. Here ρ_p is the probability distribution for d_i given C_i; ρ_d is the probability distribution for $\hat{d}_i$ given C_i. Suppose the defense wishes to examine a number of cases N_d in order to select N_t of them for trial, spending E_d on each case examined. In order to minimize $\langle R \rangle$, it is sufficient (from eqq. [B5a] and [B5b]) to choose the N_t cases for which $\hat{d}_i/C_i$ is lowest. This corresponds, on Figure B2 to picking the shaded region under ρ_d, with area N_t/N_d.[43] Here R^m is the maximum ratio of $\hat{d}_i/C_i$ selected for trial and depends on N_t/N_d and E_d.

Suppose that, as shown,

$$R^m < \frac{[1 + \epsilon(E_d)]}{(1 + \epsilon_p)}.$$

We then have, with a little manipulation,

$$\left\langle \frac{\hat{d}_i}{C_i} \right\rangle = \frac{[1 - \epsilon(E_d)]}{3(1 + \epsilon_p)} + \frac{2R^m}{3}, \tag{B6}$$

where the expected value, here and below, is taken over the cases selected for trial. Combining equation (B5a) with equation (B6), we have, for a given value of R^m,

$$\left\langle \frac{d_i}{C_i} \right\rangle = \frac{1}{2(1 + \epsilon_p)} + \left\langle \frac{\hat{d}_i}{C_i} \right\rangle \frac{1}{2[1 - \epsilon(E_i)]}$$

$$= \frac{2}{3(1 + \epsilon_p)} + \frac{R^m}{3[1 - \epsilon(E_d)]}. \tag{B7}$$

We also have, with a little more manipulation,

$$R^m = \frac{2\sqrt{\dfrac{N_t}{N_d}}\,\sqrt{\epsilon(E_d)}\,\sqrt{\dfrac{1 + \epsilon_p}{1 - \epsilon_p}} - 1 + 1 - \epsilon(E_d)}{1 + \epsilon_p}. \tag{B8}$$

---

[43] The defense is ignoring C_i in picking cases to examine since, as we see by eqq. (B6) and (B7), it can get the same $\langle R \rangle$ out of each level of C_i by picking the same fraction of each. The analysis is being done for a particular value of C_i, but applies to every value.

Substituting equation (B8) into equation (B7) gives us

$$\left\langle \frac{d_i}{C_i} \right\rangle = \frac{1}{(1 + \epsilon_p)} \left[1 + \frac{\sqrt{\epsilon(E_d)} \left[2\sqrt{\frac{N_t}{N_d}} \sqrt{\frac{1 + \epsilon_p}{1 - \epsilon_p}} - 1 \right]}{3[1 - \epsilon(E_d)]} \right] = \langle R \rangle.$$

The defense chooses N_d, E_d to minimize the expected damage payments plus expenditure examining cases:

$$\langle R \rangle \sum_i C_i + N_d E_d = \langle R \rangle N \langle d \rangle + N_d E_d = \frac{1}{(1 + \epsilon_p)}$$

$$\times \left[1 + \frac{\sqrt{\epsilon(E_d)} \left[2\sqrt{\frac{N_t}{N_d}} \sqrt{\frac{1 + \epsilon_p}{1 - \epsilon_p}} - 1 \right]}{3[1 - \epsilon(E_d)]} \right] N \langle d \rangle + N_d E_d.$$

(B9)

Setting the derivatives with respect to N_d and E_d equal to zero gives us

$$N_d = \frac{-\dfrac{d\epsilon(E_d)}{dE_d} N \langle d \rangle [1 + \epsilon(E_d)] \sqrt{\dfrac{N_t}{N_d}} \sqrt{\dfrac{1 + \epsilon_p}{1 - \epsilon_p}} - 1}{3(1 + \epsilon_p) \sqrt{\epsilon(E_d)} [1 - \epsilon(E_d)]^2}$$

(B10a)

and

$$E_d = N \langle d \rangle \sqrt{\frac{N_t}{N_d^3}} \frac{\sqrt{\epsilon(E_d)} \sqrt{\dfrac{1 + \epsilon_p}{1 - \epsilon_p}} - 1}{3[1 - \epsilon(E_d)](1 + \epsilon_p)}.$$

(B10b)

Combining the last two equations and solving for N_d yields

$$N_d = \frac{-\dfrac{d\epsilon(E_d)}{dE_d} [1 + \epsilon(E_d)] E_d N_d}{\epsilon(E_d)[1 - \epsilon(E_d)]},$$

from which it follows that

$$1 = \frac{-E_d \dfrac{d\epsilon(E_d)}{dE_d}}{\epsilon(E_d)} \left(\frac{1 + \epsilon(E_d)}{1 - \epsilon(E_d)} \right).$$

(B11)

The solution to equation (B11) is a value E_d^* independent of N. Since the plaintiffs must examine every case, it never pays them to spend more than $\langle d \rangle$ on each; so as N goes to infinity, E_p goes to some upper limit, and $\epsilon(E_p)$ goes to some limit greater than zero. Combining these facts with equation (B10a) implies that, for sufficiently large N, N_d increases as $N^{2/3}$. Hence,

$$\lim_{N \to \infty} \frac{E_d N_d}{N \langle d \rangle} = 0 \quad \text{and} \quad \lim_{N \to \infty} \langle R \rangle = \frac{1}{(1 + \epsilon_p)}.$$

As the number of cases goes to infinity, the defense perfectly identifies over-claimed cases at a cost that is vanishing small compared to the amount at stake. This confirms the verbal analysis earlier used to derive the plaintiffs' strategy in the limit of large N.

THE PLAINTIFFS' STRATEGY: FINITE N

We can use our results for the defense to learn more about the plaintiffs' strategy for finite N.[44] The equivalent of equation (B8), seen from the plaintiffs' side, is

Net Benefit to Plaintiffs

$$= \frac{1}{(1 + \epsilon_p)} \left[1 + \frac{\sqrt{\epsilon(E_d)} \left[2 \sqrt{\frac{N_t}{N_d}} \sqrt{\frac{1 + \epsilon_p}{1 - \epsilon_p}} - 1 \right]}{3[1 - \epsilon(E_d)]} \right] N \langle d \rangle - N E_p.$$

Here E_p is the expenditure by the plaintiffs on examining each case.

The plaintiffs maximize their net benefit by choosing E_p:

$$1 = -\frac{\dfrac{d\epsilon(E_p)}{dE_p} \langle d \rangle}{[1 + \epsilon(E_p)]^2}$$

$$\times \left[1 + \frac{2\sqrt{\epsilon(E_d)} \sqrt{\dfrac{N_t}{N_d}}}{3[1 - \epsilon(E_d)]} \sqrt{\frac{1 + \epsilon(E_p)}{1 - \epsilon(E_p)}} - 1 \left(1 - \frac{1 + \epsilon(E_p)}{2\epsilon(E_p)[1 - \epsilon(E_p)]} \right) \right].$$

In the limit as N goes to infinity, this gives us back equation (B4).

SOME COMPLICATIONS

We have assumed, so far, that all cases are identical ex ante. Suppose we instead assume that there are two sorts of cases: easy cases and hard cases. For easy cases, the distribution of error is $\epsilon_e(E)$; for hard cases it is $\epsilon_h(E)$. The condition $\epsilon_e(E) < \epsilon_h(E)$ for all E holds.

The argument implying that $R_i = R$ for all i still holds under this assumption, but the plaintiffs' attorney no longer maximizes his clients' expected net benefit by using the same value of E_i for each client and making claims C_i proportional to d_i. Instead, he applies equation (B4) separately to calculate E_e, E_h; C_i is then calculated from equation (B2), using E_e for the easy cases and E_h for the hard cases.

---

[44] This is not a full analysis since I have not redone, for finite N, the proof that the plaintiffs spend the same amount on every case and choose claims proportional to the estimated strength of each case.

In my earlier verbal analysis, I asserted that the plaintiffs' attorney would find it in his interest to make his claims proportional to his estimate of the strength of each case. We now see that this conclusion must be qualified. If some cases are known to be more difficult than others, meaning that it is more costly to estimate the average verdict if they are tried, the plaintiffs' attorney has an incentive to hold down his costs by making a less accurate estimate for those cases and making up for it by somewhat underclaiming them.

BIBLIOGRAPHY

Akerlof, George A. "The Market for 'Lemons': Quality, Uncertainty, and the Market Mechanism." *Quarterly Journal of Economics* 84 (1970): 488–500.

Friedman, David. "Private Creation and Enforcement of Law: A Historical Case." *Journal of Legal Studies* 8 (1979): 399–415.

Friedman, David. "What Is 'Fair Compensation' for Death or Injury?" *International Review of Law and Economics* 2 (1982): 81–93.

James, Fleming, Jr., and Hazard, Geoffrey C., Jr. *Civil Procedure*. 3d ed. Boston: Little, Brown, 1985.

Macey, Jonathan R., and Miller, Geoffrey P. "The Plaintiffs' Attorney's Role in Class Action and Derivative Litigation: Economic Analysis and Recommendations for Reform." *University of Chicago Law Review* 58 (1991): 1–118.

Robinson, Glen O., and Abraham, Kenneth S. "Collective Justice in Tort Law." *Virginia Law Review* 78 (1992): 1481–1519.

Saks, Michael J., and Blanck, Peter David. "Justice Improved: The Unrecognized Benefits of Aggregation and Sampling in the Trial of Mass Torts." *Stanford Law Review* 44 (1992): 815–51.

Shukaitis, Marc J. "A Market in Personal Injury Tort Claims." *Journal of Legal Studies* 16 (1987): 329–49.