2010

Rough Justice

Alexandra D Lahav, Fordham Law School

Available at: https://works.bepress.com/alexandra_lahav/6/
ROUGH JUSTICE

Alexandra D. Lahav
Professor of Law, University of Connecticut
Alexandra.lahav@law.uconn.edu

ABSTRACT

This Essay offers a new justification for rough justice. Rough justice, as I use the term here, is the attempt to resolve large numbers of cases by using statistical methods to give plaintiffs a justifiable amount of recovery. It replaces the trial, which most consider the ideal process for assigning value to cases. Ordinarily rough justice is justified on utilitarian grounds. But rough justice is not only efficient, it is also fair. In fact, even though individual litigation is often held out as the sine qua non of process, rough justice does a better job at obtaining fair results for plaintiffs than individualized justice under our current system. While rough justice also has its limitations, especially to the extent it curbs litigant autonomy, in the end it is the most fair alternative currently available for resolving mass tort litigation.

INTRODUCTION

What is rough justice? In many mass tort cases (as in many ordinary tort cases) it is impossible to bring all cases to trial. Even if the judge were to try cases for one hundred years only a fraction of the cases in the typical mass tort litigation would be heard. To deal with this
problem, judges have begun using informal statistical adjudication techniques to determine more or less what damages, if any, plaintiffs ought to be awarded. Often courts will try “informational” bellwether cases, taking the verdicts of those cases and assisting the parties in extrapolating them over the entire population in an aggregate settlement. The key feature of rough justice in mass torts is the attempt to resolve large numbers of cases by giving plaintiffs some recovery within the range of compensation in comparable cases.

This is a defense of rough justice. Most defenses of rough justice begin with the assumption that rough justice is second best, “good enough” justice in cases where there are limited resources. When each person cannot have an individual trial, the argument usually goes, rough justice is an acceptable fallback because it is efficient. But rough justice is not merely an acceptable alternative to “real” individualized justice. As demonstrated below, rough justice is justifiable on its own terms because in mass tort cases rough justice is more fair than the alternatives. In fact, individualized justice is rougher than justice on a mass scale. The reason is that in mass tort or aggregate litigation cases can be resolved using rigorous, transparent methods that will result in like cases being treated alike. In individual litigation, cases are resolved through loose methods of comparison that are hidden from view and are unlikely to result in equality between similarly situated litigants.

The Essay begins with a description of how participants in the tort system assign value or monetize tort cases. It then turns to emerging methods for resolving mass torts, especially bellwether trials, the most popular statistical adjudication procedure. This procedure brings to light a fact that has been hidden in the shadow of the lawsuit: the extent to which valuations of injury are contextual and subjective. Finally, the first part lays out the requirements for a fair and rigorous sampling method.
The second part of the Essay presents the justifications for rough justice. Equality between similarly situated litigants is a central value in our procedural law and a moral imperative. In support of this position, the second part presents three arguments. First, rough justice better realizes the ideal of equality and evenhandedness in the assignment of value to tort cases than the alternatives and is therefore more fair. Second, statistical methods also allow judges to control case management and timing in ways that are more fair to litigants, particularly to the most harmed litigants. Third, because the only fair way to value tort cases is comparatively, in order to fairly value mass tort cases, the methods of statistical adjudication used must be rigorous and transparent.

The conclusion addresses some of the limitations of rough justice. Two competing values undergird our adjudicative system: liberty and equality. Rough justice realizes the ideal of equality at the expense of litigants’ liberty and autonomy. In using statistical adjudication, participants in the tort system must come to terms with the face that not every litigant will have an opportunity to participate in a public trial of her case. For individualists, this loss of autonomy may be unacceptable. But the loss of autonomy is already an integral characteristic of our tort system. The question ought to be, given that not everyone will get their day in court, what is the fairest way to resolve cases?

1. CASE VALUATION

In order to understand why it is that the rough justice of statistical adjudication is less rough than ordinary individualized litigation, it is important to understand what I mean by case valuation and how cases are in fact valued by lawyers on the ground. This Part first discusses
what it means to say that case valuation is “subjective.” It then contrasts the tools lawyers use to value cases in individualized litigation with emerging methods of statistical adjudication in mass torts, particularly court engineered sampling.

1.1 The Concept of Case Valuation

The normative ideal of case valuation is that plaintiff ought to receive what she is entitled to under the substantive law. For example, if the defendant is not liable then the plaintiff is entitled to nothing under the substantive law. If the defendant is liable then the plaintiff is entitled to some measure of damages. Tort damages are usually understood to consist of three components: (1) economic or pecuniary damages, (2) noneconomic damages such as compensation for pain and suffering, disfigurement, emotional distress and loss of quality of life, and (3) punitive or exemplary damages. When a case is valued for settlement, at least in theory, the plaintiff is supposed to receive the amount of her damages under each of these three categories discounted by the possibility that defendant will be found not liable in the course of litigation. The ideal, then, is that settlement is to approximate the outcome that litigation would reach and that litigation approximates the actual damages suffered by plaintiff as a result of defendant’s misconduct.

The problem with this understanding of tort law is that with respect to damages, the tort system does not approximate the actual damages suffered by the plaintiff. What the tort system does is assign a value to the damages suffered by the plaintiff. Why is this so? The tort system is an institution that monetizes injuries, yet injuries are not readily monetizable. Instead, the amount of money damages the system assigns to injuries is contextual and cultural. This is the reason that the amounts awarded in tort cases are sometimes controversial. And it is why critics
of the tort system are able to say that the system is unpredictable. Furthermore, this problem of value is not limited to the litigation context. In settlement, even if one is able to accurately discount the amount of damages by the probability of the defendant being found liable, the damages assigned to a plaintiff (the amount that is to be discounted) will still be contested.

Some scholars have proposed administrative methods for resolving this issue. For example, Eric Posner and Cass Sunstein note that administrative agencies value loss of life based on what people would be willing to be paid to accept risk.¹ Such valuations are contested. And the authors themselves admit that the act of pricing human life is contextual.²

There are two ways that this assertion could be read. First, it could be read to say that it is impossible to know what an injury is actually worth, even from a god’s eye point of view. Second, it could be interpreted as an epistemic problem: we lack the tools to accurately monetize injuries. For the purposes of this Essay, it does not matter which of these explanations is right. We do not have the tools to accurately monetize damages, either because the task is impossible or because it is beyond our current capabilities.

Despite the limitations of its tool kit, the tort system nevertheless needs to monetize injuries in order to compensate plaintiffs. It does this by providing a set of guidelines to juries, who hear evidence regarding the plaintiff’s damages and assign a monetary value to them. A


² Id. at 554. Douglas Kysar has made a trenchant and convincing criticism of this approach. See DOUGLAS KYSAR, REGULATING FROM NOWHERE: ENVIRONMENTAL LAW AND THE SEARCH FOR OBJECTIVITY 110-119 (2010).
closer look demonstrates how it is that tort damages are varied and contextual rather than fixed and objective.

Of the three kinds of tort damages (economic, pain and suffering and punitive damages), economic damages are generally considered the most easily ascertainable. After all, economic damages concern things we are accustomed to measuring, such as wages. Yet measuring economic damages is not easy or obvious and the measure of economic damages is perhaps as contested as other forms of tort damages. Although noneconomic and punitive damages are most often criticized for being outrageous, excessive or emotionally driven, economic damages suffer from many of the same difficulties of valuation. 3

First, predicting future lost income – a substantial part of any economic damages award – requires the exercise of a great deal of judgment and leaves plenty of room for argument. Consider the simple example of predicting the future lost income of a law student in May 2007 as compared to the same law student in May 2009, after the melt down of the financial markets in 2008 and the predictable effects of the Great Recession on the legal profession. In another example, a recent study of the 9/11 Victim Compensation Fund demonstrated that economic loss awards were influenced by testimony from forensic economists and that economic awards varied

across similarly situated claimants. This indicates that with respect to individuals there is a range of acceptable awards that are justifiable, rather than a specific amount that reflects their entitlement. Furthermore, that range can shift over time.

Second, the category of economic damages is an evolving one that courts have stretched to include damages that once were considered noneconomic. In wrongful death cases, for example, recovery was historically limited by statute to pecuniary damages such as lost wages, which are often quite low. Recovery for grief and loss was not permitted. To ameliorate the harshness of the law, judges permitted plaintiffs to demonstrate the value of services the decedent provided, even if these were not compensated and did not have a market value. These types of additional damages included both loss of services and loss of consortium. In this way, survivors were permitted to collect damages for the loss of wives who were uncompensated for work in the home and children who did not work at all.

Third, the lines between the three different categories of tort damages are fluid. Award amounts can often reasonably be assigned to more than one legal category, rendering the measure of damages malleable. For example, an empirical study by Catherine Sharkey demonstrated that when noneconomic damages are capped by statute, there is no overall effect

---


on compensatory damages in medical malpractice cases. She showed that lawyers can use expert testimony and innovative theories to expand categories of damages such as future wages and the costs of future medical care. Thus economic losses replace other forms of damages, with the overall total held roughly constant.

Sociological studies of juries support these findings. Jurors interviewed by the sociologist Neil Vidmar about their participation in a medical malpractice case considered the effects of emotional trauma and disfigurement on a plaintiff’s likelihood of obtaining a promotion. Damages for disfigurement should ordinarily be considered as part of the category of non-economic or pain and suffering damages, whereas loss of future income ought to be considered economic damages. That jury’s treatment of disfigurement demonstrates how porous these categories can be.

The idea that categories of damages are fluid and therefore malleable is consistent with the premise that many participants in the legal system think of damages, as a single all-inclusive number as opposed to discrete legal categories. Studies have shown that jurors consider damages holistically, despite the law’s mandate to parse damages out into the three familiar

---


7 *Id.* at 438-440 (stating that “experts can exploit controversies surrounding calculations of lost wages and future medical costs, thereby breathing life into the crossover effect.”).

types. Lawyers have a holistic perspective as well. As one plaintiff’s attorney told a researcher: “in most instances a jury has a figure in mind, and when you have a figure in mind, it can come in the guise of compensatory damages or in the guise of punitive damages.” It makes sense for lawyers to look at cases holistically because most cases settle, and they settle for a single number. Lawyers on both sides may argue about what types of damages they think are likely to be awarded in a given case. At the end of the day, however, the overall number is more important to lawyers and juries than legal damages categories.

1.2 How Values Are Assigned in Individual Tort Cases

In ordinary tort litigation, cases are monetized by juries or lawyer negotiation. Both methods are flawed. Studies of jury verdicts demonstrate unpredictable variations. This is consistent with the theory that methods for assigning value to injuries do not produce objectively verifiable results. Settlements are often based on comparable cases, but the methods lawyers use to create a set of comparable cases are not rigorous and their results ought to be viewed with suspicion.

9 Ibid.


11 The tendency to use the most easily accessible data or anecdote is sometimes referred to as the “availability heuristic,” first developed by the psychologists Amos Tversky and Daniel Kahneman. See, e.g., Daniel Kahneman, Paul Slovic, and Amos Tversky, eds, *Judgment Under Uncertainty: Heuristics and Biases* (Cambridge 1982).
When juries assign damages there is sometimes unexplained variability in awards. Studies have found that jury verdicts vary for reasons that researchers do not understand or cannot measure. These variations are not always related to the severity of the underlying injury. For example, research using simulated juries suggests that when all known relevant variables are held constant, jury verdicts vary in ways that cannot be accounted for by the severity of harm to the plaintiff.\(^\text{12}\) Research based on historical jury awards reaches similar conclusions.\(^\text{13}\) What does this unexplained variability in jury awards mean? There may be differences between plaintiffs that are legally relevant but difficult to measure.\(^\text{14}\) If so, we may want juries to take into account immeasurable facts, as long as they are not illegal ones such as a party’s race. But variation in simulation research is not caused by differences in plaintiff’s case because the simulated facts are the same every time. Some scholars hypothesize is that this variance is due to

\(^{12}\) Michael J. Saks et al., *Reducing Variability in Jury Awards*, 21 LAW & HUM. BEHAVIOR 243-56 (1997) (simulation research suggesting that when variables are held constant jury verdicts vary).


\(^{14}\) “Even precisely the same injury can have vastly different implications for the compensation of different plaintiffs that a jury might reasonably consider. A classic example is the difference between the loss of a finger for a professional violinist versus a professor or a factory worker.” Shari Diamond & Mary Rose, 1 ANN. REV. OF L. & SOC. SCI., 255-284, 264 (2005).
the fact that juries are not given comparable cases to assist them in the process of valuation.\textsuperscript{15} Variability, especially in simulation cases, raises the possibility that juries’ assignments of value are not moored to an objective measure.

Because tort cases are rarely tried, most comparable values come from settled cases informed by the outcome of pre-trial motions that are either partially or wholly dispositive of a case.\textsuperscript{16} In cases resolved through negotiation, settlements are based largely on the lawyers’ sense of the “market” in settlements and verdicts.\textsuperscript{17} Lawyers develop a sense of the “going rate of settlement” with respect to a particular set of cases, of which their client’s case is but one.\textsuperscript{18} This “going rate” is determined by comparison to other case values. By comparing the outcomes of similar cases, lawyers can evaluate what the outcome should be in their client’s case. I call this latter method “anecdotal sampling.” Anecdotal sampling is a form of inductive reasoning based on establishing consistency of a given case with other cases. It bears some rough similarity to qualitative social science research methodology. Unlike good social science methods, the pool of cases lawyers use as a point of comparison is often a sample culled from whatever case results


\textsuperscript{18} This has been the case for a long time. See Samuel Issacharoff & John Fabian Witt, \textit{The Inevitability of Aggregate Settlement: An Institutional Account of Tort Law}, 57 Vand. L. Rev. 1571 (2004).
are easily available to the lawyer. Statisticians call this a “convenience sample,” which means it is a nonrandom (or biased) sample readily available to the researcher. 19 By contrast, good social science methods require random sampling. In cases where there are consistent repeat players in the particular area of law, it is likely that the sample of settled cases will be more comprehensive. There is little empirical evidence regarding the data sets possessed by lawyers or the methods lawyers apply to those data sets.

Sometimes settlement amounts are not based on jury verdicts or comparable settlements but on the limits of defendant’s insurance coverage. Insurance in these cases acts as a cap on damages, so that the assigned value of a case will be more closely tied to the defendant’s coverage than the plaintiff’s condition. 20 This method seems arbitrary.

The methods participants in the tort system use to monetize injuries present three key problems. First, they are not transparent. Instead, the tort system is a complex, private, largely hidden system of compensation. 21 The dearth of empirical evidence about how cases settle is

---

19 1 SAGE ENCYCLOPEDIA OF SOCIAL SCIENCE RESEARCH METHODS 197 (Michael S. Lewis Beck, Alan Bryman & Tim Futing Liao, eds., 2003).


symptomatic of this problem. The best data sets of case outcomes are those owned by insurance companies, the entities that pay for litigation and settlement. Insurance companies do not ordinarily make this data available. To the extent that insurance companies are willing to sell their data to researchers, this would be a fantastic resource. The legal system needs good qualitative empirical studies of case valuation outside the context of juries.

Second, the process of comparison among outcomes may be skewed by a biased sample. A sampling process based on anecdotal evidence culled from whatever cases the lawyer may have come across is especially vulnerable to bias.

Third, the process of comparison that is the basis for valuation may entrench existing inequalities or errors because it compares the new case to past cases. If some types of cases have been undervalued in the past for illegal reasons, such as the race of the plaintiff, then using those outcomes as a benchmark for current or future cases will perpetuate that undervaluation. Furthermore, valuing cases based on past outcomes produces a static value. But if the value of cases is dynamic and evolving, then the previous valuations may not reflect the current consensus on the value properly assigned to that case.

---


23 One recent qualitative study aims to remedy this. See Engstrom, supra note 17 (qualitative empirical study of settlement practices of personal injury lawyers).

Mass torts expose this private, hidden system in a way that “ordinary” cases do not. In mass torts, judges oversee a resolution process which some have called a “quasi-administrative agency.” 25 This process mimics on a grand scale the process lawyers use to resolve ordinary cases. The history of tort law in the U.S. demonstrates that this type of “scheduling” of tort cases has been the norm since the Industrial Revolution. 26 The awards given out by mass tort claims administrators can be publicized and centrally available, in contrast to the ordinary case where comparable values are not easily accessed. This more transparent process requires judges and policy makers to think more thoroughly about the problems raised by assigning values based on sampling. It also highlights the often ignored valuation problems inherent in “ordinary” litigation.

1.3 Court Engineered Sampling In Mass Torts

Recently, judges overseeing large numbers of tort cases collected in a single forum under the auspices of the Multi-District Litigation statute and similar procedural devices have started using

---


26 Scheduling is the determination of compensation for an individual case based on a preset amount according to various factors, such as injury severity. See generally Issacharoff & Witt, supra note 14 (a history of settlement practices in personal injury cases).
sampling techniques such as non-binding bellwether trials for “informational” purposes.\textsuperscript{27} I call this process court-engineered sampling. It is in some ways similar to what lawyers ordinarily do when they compare similar cases to determine value.

In the late 1990’s, a few trial courts experimented with binding statistical adjudication procedures. In \textit{Hilao v. Marcos}, a federal court used statistical methods to adjudicate a class action brought on behalf of persons who suffered human rights abuses under the regime of Ferdinand Marcos in the Philippines.\textsuperscript{28} A special master conducted on-site depositions in the Philippines, and based on these he recommended a recovery schedule to a jury, which then adopted his recommendations (for the most part). The Ninth Circuit upheld this procedure. Around the same time, a U.S. District Court judge in Texas tried 167 asbestos cases and was prepared to use these verdicts to extrapolate to the remainder of asbestos cases before him.\textsuperscript{29} The Fifth Circuit squashed his efforts, holding that the extrapolation of the results of the sample

\begin{flushright}
\begin{minipage}{0.95\textwidth}
\footnotesize

\textsuperscript{28} Hilao v. Estate of Marcos, 103 F.3d 767 (9\textsuperscript{th} Cir. 1996).

\end{minipage}
\end{flushright}
verdicts violated the defendant’s due process right and the Seventh Amendment. No trial court has followed in the footsteps of these innovators.

Sampling has resurfaced in the last five years, albeit as an informal method for encouraging aggregate settlements rather than a binding method for resolving cases in the class action context. It has been used in both state and federal forums. The most often cited example is the Vioxx litigation, but there are many others. Sampling was used to encourage settlement in the September 11th Litigation. A sampling process was instituted in the related World Trade Center (WTC) Disaster Site Litigation. Sampling has been proposed in the litigation over formaldehyde laden FEMA trailers, in the litigation arising out of the presence of methyl tertiary butyl ether in the water supply, and in the Fosamax litigation. It has also been

---

30 Cimino, 151 F.3d 297.


32 See Order Amending Case Management Order No. 8, In re World Trade Center Disaster Litigation, 21 MC 100 (Feb. 19, 2009).

33 FEMA Trailer Formaldehyde Prods. Liab. Litig, MDL No. 1873, E.D. La.

34 In re Methyl Tertiary Butyl Ether (“MTBE”) Products Liability Litigation, MDL No. 1358, S.D.N.Y.

35 In re Fosamax Products Liability Litigation, MDL No. 1789, S.D.N.Y.
proposed in the gender discrimination class action against Wal-Mart to solve the problem that the case required individual trials and was therefore not manageable as a class action.  

The WTC Disaster Site Litigation presents a particularly intriguing and sophisticated approach to statistical adjudication. Approximately 9,900 plaintiffs filed lawsuits against nearly 200 defendants alleging injuries arising out of their exposure to harmful chemicals in the aftermath of the tragedy of September 11, 2001. The plaintiffs include New York City employees such as firefighters and police officers, as well as civilian volunteers and others. The court appointed two special masters, both law professors, to set up a procedure for sampling the cases to encourage settlement. They developed a method for allocating the plaintiffs into groups, which would then be sampled in preparation for bellwether trials. Plaintiffs were required to fill out questionnaires regarding types of diseases and severity of injury and the information was entered into a database. The groups were then organized based on type of illness and severity of harm. Out of the first group of 2,000 cases, the special masters collected 200 of those alleging the most severe injuries, 25 additional cases of other diseases that had not been included in the severity chart and 400 cases chosen at random. Of these, the judge picked two cases, the defense picked two cases and the plaintiffs’ lawyers picked two cases, for a total of six cases set to proceed through pretrial and trial. Judge Hellerstein explained that this “allows the parties to get a good sense of the strengths and weaknesses of all the cases” and

36 *See* Dukes v. Wal-Mart, 509 F.3d 1168, 1191 (9th Cir. 2007).

37 *See* Order Amending Case Management Order No. 8, In re World Trade Center Disaster Litigation, 21 MC 100 (Feb. 19, 2009)

38 The special masters are two very prominent torts professors: James Henderson of Cornell Law School and Aaron Twersky of Brooklyn Law School.
presumably would lead to settlement.\textsuperscript{39} The Judge later increased the number of bellwether trials to twelve, but before any cases were tried an aggregate settlement was reached.\textsuperscript{40}

In all these cases, court engineered sampling proceeds more or less as follows. Among a large set of similar cases, the judge slates several for pretrial practice and trial. Lawyers litigate these sample cases through discovery and summary judgment, pretrial motions and even trial. When these selected cases have settled or reached judgment, lawyers can use the experience culled from them to settle all the other cases in the group because they will have developed a sense of the value of similar cases from the process of litigating the sample cases.

Aggregate litigation presents opportunities to develop more rigorous sampling methodologies, bellwether trials among them. Unlike the settlement of the stand-alone case, courts can verify that the reference class is appropriate, that the sample is chosen randomly and is sufficiently large to yield reliable results. The ordinary methods of case valuation do not require—and are virtually never based on—anything approaching a rigorous methodology. As a result, these methods do not account for the always-present potential that the settlement deviates considerably from the average value the group of similar cases would have been assigned if all were to proceed to trial. It is very difficult to sample in ordinary tort cases, but mass tort cases provide opportunities for developing a more rigorous method for assigning value to tort cases.

\begin{itemize}
\item \textsuperscript{39} See Mark Hamblett, \textit{Plan Implemented to Resolve Suits in World Trade Center Cleanup,} N.Y. L. J. (Feb. 25, 2009).
\item \textsuperscript{40} Mireya Navarro, \textit{Settlement Plan Drafted for Sept. 11 Lawsuits,} N.Y. TIMES (Feb. 4, 2010) (reporting that twelve cases were scheduled to go to trial).
\end{itemize}
Sampling raises two critical issues that courts must address: sample bias and unexplained variability in the distribution of results. I will address each of these in turn.

1.3.1. Sample Bias

The most important issue in sampling – whether done on an anecdotal basis by an individual lawyer negotiating a settlement on behalf of a client or engineered by a court in the context of aggregate litigation – is whether the sample is skewed. We must always suspect that any non-random method of picking sample cases will be skewed and therefore will be an inaccurate estimate of the population average. Second, even if the sample is an accurate estimate, verdicts may vary for some reasons that cannot be accounted for. In such a case the question becomes what is the significance of the variation – are they pure noise or is there some component of the cases that needs to be accounted for.

Sample bias is a substantial problem in the current system. Consider for a moment how the ordinary lawyer is likely to obtain data for determining the going rate of settlement. The lawyer may be familiar with values in a set of “comparable” cases because they went to trial and the verdict is publicly available, because the lawyer settled those cases him or herself or had access to settlement data for similar cases through informal channels. Thus the lawyer will either be able to use public records or friends and colleagues to obtain a “data set.”

Trial verdicts are particularly likely to provide unreliable samples for comparison. There are two reasons for this. First, cases that go to trial are aberrations.\textsuperscript{41} A case is likely to reach trial when the parties are very far apart in evaluating the case. As a corollary, any case where the result is predictable by both sides is very unlikely to reach trial. After all, if the parties felt

\textsuperscript{41} See Judith Resnik, \textit{Trial as Error, Jurisdiction as Injury}, 113 HARV. L. REV. 924, 925 (2000).
comfortable in their ability to predict the outcome, they are much better off settling and avoiding the transaction costs of an expensive trial.\textsuperscript{42} Second, parties decide whether to go to trial. Repeat players can therefore systematically skew the sample of publicly available verdicts in order to shape an end result that is most favorable for them. Defendants have an easier time doing this because they can offer settlements to plaintiffs they believe have strong cases and let weaker cases go to trial.\textsuperscript{43}

As an example of the selection bias in trials, consider the case of \textit{Rhone Poulenc v. Rohrer}.

\textsuperscript{44} A group of hemophiliacs who had contracted AIDS from tainted blood products brought a class action against the manufacturers of those products. Judge Richard Posner held that the case could not proceed as a class action, reasoning (among other things) that the class action would exert such extreme pressure on that defendant that it would be forced to settle. In support of the claim that such pressure was unwarranted, Judge Posner relied on the fact that twelve of the thirteen cases that had gone to trial resulted in defense verdicts.\textsuperscript{45} His opinion did not note how it came to be that these particular thirteen cases went to trial. It is possible, for

\textsuperscript{42} Much has been written on the theory of settlement in the field of law and economics. For a somewhat dated but very useful review of the literature, see Robert D. Cooter & Daniel L. Rubinfeld, \textit{Economic Analysis of Legal Disputes and Their Resolution}, 27 J. OF ECON. LIT. 1067 (1989).


\textsuperscript{44} In the Matter of Rhone Poulenc Rorer, Inc., 51 F.3d 1293 (1995) (Posner, J.).

\textsuperscript{45} Id. at 1299-1300.
example, that the defendant settled all or most trial ready cases prior to trial and only permitted those cases it was likely to win to go forward. If that were the case, any conclusions drawn from this sample of thirteen blood products cases would have been biased in favor of the defendant. For those watching litigation from afar, it is quite difficult to tell why some cases reached trial and others settled, leaving ample possibility for sample bias. This means that conclusions based on trial outcomes without more information are likely to lead to indefensible results. By this I mean results that are not justifiable by reference to outcomes in similar cases that never reached trial because of the machinations of one side or the other.

Random selection is critical to obtaining a useful sample. Convenience samples based on the lawyer’s personal experience or the experiences of colleagues suffer from potential bias because they are not randomly selected. If the lawyer could conduct rigorous qualitative research, by collecting a broader set of cases on which to base his or her evaluation and using them to develop a fine-grained theory of which variables in that sample are relevant to case outcomes, there is a greater chance that this evaluation will accurately reflect the going rate of settlement. Even so, qualitative methodology requires recognition of its own limitations, such as the potential of sample bias and the difficulty of finding correct points of comparison. It is most beneficially used in conjunction with quantitative methods that can verify findings. Similarly, even when qualitative methods are rigorous, some type quantitative analysis is still useful. In any event, there is no evidence that lawyers use this type of rigorous qualitative study in setting settlement amounts. If such rigorous methods are used anywhere, it is likely in the insurance context where companies collect data on settlements or perhaps where well-funded lawyers are
repeat players. Even in the case of lawyers who are repeat players, it seems likely they would be satisfied with anecdotal sampling given that there is no incentive to use more rigorous methods.\textsuperscript{46}

1.3.2. Unexplained Variability

The second most important issue in sampling is unexplained variability of the distribution of results. One of the key differences between the type of anecdotal methodology based on convenience sampling used by lawyers in the ordinary course of litigation and a court-engineered sampling methodology is the potential for the latter method to give an explicit account of variability. In situations where case outcomes are very heterogeneous, assigning case values is possible if we believe that the reasons for variation are “noise” rather than the effect of important variables that ought to have been taken into account.

One solution to the variability problem is to more rigorously set the parameters of the reference class. If the court is sampling from a reference class of cases that are similar to one another with respect to the key variables, then the result of sampling should be sufficiently homogeneous to be useful in valuing other cases in the reference class. That is, we can extrapolate the results of the sample to the rest of the reference class if all the cases are reasonably similar. But in order to decide the parameters of the appropriate reference class, the court will need to identify the variables that are relevant to case outcomes. In other words, determining the parameters of the reference class requires taking a normative position regarding which variables are important. Furthermore, these variables must be not only relevant but

\textsuperscript{46} See supra note 11.
Variables that are not objectively verifiable – such as a person’s mental state – require time consuming individualized hearings to identify in individuals. For this reason, they are not very useful in procedures that seek to extrapolate from a sample to a larger population of plaintiffs.

Comparing the case at hand to a convenience sample may create the illusion that we know its value with certainty. In convincing clients to settle, lawyers are likely to be too sure that the client’s case is comparable to other cases they have in mind, even when those cases evidence a selection bias. By contrast, a transparent, rigorous sampling method engineered by the court is less likely to suffer from such failures. Inherent in the task of developing a sampling methodology for aggregate litigation ought to be a process for taking a hard look at the problems of sample bias and the significance of unexplained variability of distribution. A rigorous methodology must recognize that the model will not be perfectly determinative. The question for the participants in the statistical adjudication procedure (and the court system more generally) is how much variance is too much? When does variance become normatively unappealing? My argument here – discussed at greater length below – is that unexplained variation due to noise is normatively acceptable so long as the overall distribution among plaintiffs is fair.

To the extent that decision makers utilize good sampling methodology, aggregate litigation may in fact provide better individualized justice than “individualized” settlements based on anecdotal sampling. Both in the individual case and the aggregate case plaintiffs receive settlements based on comparable cases. The better the methodology used to determine

---

which cases are comparable, the closer the outcome of any individual case to the average value assigned in a series of trials of similar cases. This average is the most defensible method of assigning a value to an injury.

Since our system monetizes injury in tort cases by comparison to outcomes in similar cases, especially in settlement, the key to a reliable valuation will be the quality of the sample used. At its best, a systematic, rigorous approach to sampling produces a sample that can be analyzed to assign a defensible value to a group of mass tort cases before the court. At a minimum, a rigorous approach to sampling avoids some of the biases and concomitant inequality in the assignment of case values that plague the anecdotal sampling method.

1.3.3. Requirements for a Rigorous Sampling Methodology

What are the requirements of a rigorous sampling technique? A reliable sample requires that the selection process be free from bias and that the sample be sufficiently large to produce reliable results given the variance of outcomes within the group.

Making sure that the sample is not biased is best achieved by collecting a random sample. Randomization has not generally been the practice in court engineered sampling, but it should be. Courts seem to prefer a sample constructed by permitting defendants’ and plaintiffs’ attorneys to each choose an equal number of cases, with perhaps a few additional cases thrown in by the court. \(^{48}\) This method gives the parties the illusion of control. It has the merit of signaling the nature of the bias inherent in the sample. We can predict that the defendants’ attorneys will try to pick their best cases – that is, cases that will minimize recovery – whereas the plaintiffs’ attorneys are likely to pick the cases that maximize plaintiff recovery. This knowledge can help

\(^{48}\) Fallon & Grabill, *supra* note 27 at 2349.
us determine how much we ought to discount the judgments in the cases selected by the parties. Nevertheless, litigant driven sampling encourages the selection of outlier cases and is not likely to produce a data set that provides reliable information about the distribution of the larger set of cases. If that distribution is already known and it is possible to situate the biased sample cases within that distribution, the limits of this method may be surmountable. But that is impossible without a preliminary procedure using random sampling to determine the characteristics of the larger population. In any event, if a larger study has been done, there is no need to follow it with biased sampling in order to obtain the information the court and litigants need to settle cases fairly.

A second requirement is that the sample be sufficiently large to provide reliable results. The size of the population of cases being measured does not dictate the size of the sample. Instead, the size of the sample will depend on the variability of the group. In the mass tort context, the special master ought to determine what observable and relevant criteria exist that can place individuals within relatively homogenous sub-groups and then sample from each of those sub-groups. The samples need not be large. The more homogeneous the group, the smaller the necessary sample.

Sampling is a very useful method for estimating damages awards in mass torts or class actions because smaller numbers of cases can be tried and the results extrapolated to the larger population of cases from which the sample is drawn. This extrapolation process saves considerable public and private resources as compared to individual trials. But we know that in tort cases there are often numerous relevant variables, and that cases can differ from one another

considerably. For this reason, courts need a good estimate of the variations within the group in order to determine what the size of the sample ought to be.

A preliminary estimate of the sub-groupings within a mass tort is best obtained though the collection of data from the parties. For example, in the WTC Disaster Site Litigation, Judge Hellerstein ordered the parties to complete questionnaires to determine the variability within the group. The judge determined the size of the sample to be tried prior to the litigants actually completing the questionnaires. He did this largely because the timing of completing the questionnaires would have held off sample trials for too long, delaying justice for individuals and leaving defendants in limbo. But this approach can result in a very rough estimate of damages indeed. In a rigorous sampling method, the variability in the population will determine the size of the sample. It is not possible to determine the appropriate sample size without first knowing a sense of that variability. If the group is relatively homogeneous, then a small sample will be enough. But if there is substantial variation within the group, the sample size will need to be larger. To know the margins of variation, courts need to obtain information about the population to be sampled, through surveys such as that used in the WTC Disaster Site Litigation or past experience.

So why try a sample that is likely too small? First, the court might move a few cases towards trial on the theory that the momentum will result in settlements. The object would not be to construct the most reliable procedure to determine value. Instead, this method relies on anecdotal sampling and the belief that litigant consent is sufficient to justify outcomes. This seems to have been the approach in the September 11th Litigation. In that case, the judge slated cases for trial explicitly in order to encourage settlement, on the theory that even a single verdict
would bring the settlement offers on both sides closer together.\textsuperscript{50} No case has yet been tried, but a number were settled as they approached trial.\textsuperscript{51}

Second, as cases proceed through pretrial litigation – discovery, summary judgment and motions \textit{in limine} – lawyers narrow their claims and develop a keener sense of the story that they will be able to tell at trial. These developments might be called the “soft benefits” of sampling. The questions presented in a given case are framed more precisely and often some of the claims initially included in the complaint fall away. Furthermore, summary judgment decisions, especially those on questions of law where individual variables are not likely to matter, dispose of issues that are similar in the larger population of cases. Even if such decisions are not preclusive as a formal matter, they serve as an indicator of what is likely to occur in the other cases presenting similar issues before the judge. When the court decides such dispositive or key issues, lawyers use them to develop a finer sense of the possibilities in other cases. But the reliability of these predictions is not the same for all issues facing litigants. The closer the decided issue is to a question of law applicable across cases, the more likely the judge’s decision will have an impact on other cases in the reference class.

Permitting a small sample of cases to go forward on a limited basis, even if the results cannot be reliably extrapolated across cases, can be very useful in case coordination and issue refinement. Nevertheless, we should recognize the limits of an approach that does not use a

\textsuperscript{50} \textit{See Opinion Supporting Order to Sever Issues of Damages and Liability in Selected Cases, In re September 11 Litigation, 21 MC 97, at 4 (AKH) (S.D.N.Y., July 5, 2007).}

\textsuperscript{51} \textit{See Order, In re September 11 Litigation, 21 MC 97 (AKH) (S.D.N.Y., Sept. 17, 2007) (ordering fourteen cases closed due to settlement).}
reliable sample. If the results of a very limited convenience sample are used to determine outcomes in a broad range of cases without attention to the variables that differentiate those cases, the result will not reflect a reasonable assessment of case value based on comparable cases. For this reason it will be unfair to the litigants.

Even if the larger population is meticulously studied and grouped into more homogenous categories based on observable and relevant criteria, there will still be some noise. This noise will be caused by variables that are not observable although legally relevant or that are not legally relevant but that nevertheless alter the outcome in a given case. Furthermore, there may be variables that are observable and relevant, but are so rarely present it is difficult to take them into account through sampling.

To illustrate, consider a group of persons who suffered injuries to their hands. One of them is a concert pianist, and the rest are white collar workers. We may all agree that the concert pianist should receive greater compensation than a lawyer for the same hand injury. But she presents a significant problem for the sampling procedure. It is hard to predict the presence of the concert pianist within the group. If the concert pianist is within the sample, then her presence will skew the results and the rest of the group will be overcompensated. But if the fact that she is a concert pianist is not taken into account for her individual case, then she will be undercompensated. The court will need to realize that the presence of a pianist is a relevant factor that should be included in the model. If the costs of determining variables like this one are very great, this presents a problem for implementing a sampling procedure in the real world.

Rigorous sampling forces litigants and the court to face the issue of variance in the distribution of case valuation. It requires courts and litigants to think systematically about both
the generic case of the white collar worker in our illustration and the outlier case of the concert pianist. Courts must either justify treating the concert pianist the same as a white collar worker or create a procedure to fairly distinguish her case. Such systematic consideration is the first step to a fair and transparent resolution of large-scale litigation. The alternative is not likely to be an accurate determination of each case through individualized litigation, but instead the even rougher justice of anecdotal sampling.

The trend towards informal sampling as a method for encouraging settlement brings comparison, the traditional method of case valuation, out of the shadows. Sampling illustrates, counter-intuitively, that justice administered at the wholesale level may be less rough than that at the individual level precisely because high-quality sampling is better than anecdotal sampling. But it also raises questions as to what justifies such systemic rough justice. Is it enough to argue that because we tolerate rough justice at the retail level we should tolerate it at the wholesale level? To the extent that sampling is more expensive than informal methods of case resolution, what principle justifies the additional cost? The next Part presents the justifications for rough justice in mass torts.

2. JUSTIFICATIONS FOR ROUGH JUSTICE

Ordinarily the arguments in favor of statistical adjudication are efficiency arguments, which boil down to the idea that the overall benefits of sampling to litigants, the legal system and society is worth the cost to individuals (in other words, sampling provides the greatest good to the greatest number). This Essay takes a different approach, arguing that sampling is desirable
not only because it is efficient but also because it is fair. This section explains why it is that a system of statistical adjudication is justified on this basis.

Sampling is fair because it better achieves equal treatment of litigants than existing alternatives. Equality is a fundamental principle of the rule of law and of our procedural system. It is also a requirement for fairness. To be fair towards litigants, a procedural regime ought to be evenhanded; it ought to treat like cases alike. This principle supports a sampling regime over the current regime that privileges individuation over equality. Currently litigants are not treated equally when it comes to the timing of case resolution (when cases go through the pretrial process and get to trial) and case valuation. At a minimum, thinking about the effect of sampling on case management and valuation ought to raise serious questions about the fairness of the current system. More than this, I hope to convince the reader that sampling is desirable because it better realizes the ideal of equality than the current system.

Secondly, sampling is beneficial because it requires a level of transparency currently not available in civil litigation. As discussed above, our tort system has evolved into a private and largely hidden system for resolving disputes. Sampling requires transparency in the form of overt comparison between outcomes and an open discussion of which variables are relevant to outcomes and which are not. This discussion will likely expose existing disagreements between parties and between parties and judges. It is an important discussion to have, one that implicates fundamental values of the tort system and our legal system more generally.
Recognizing these benefits, it is also important to acknowledge that statistical adjudication involves trade-offs. Statistical adjudication consciously substitutes settlements for individualized adjudication. There is a loss associated with statistical adjudication; it does not permit each individual to have her case tried to a jury. For this reason, the use of statistical adjudication highlights the classic tension between liberty and equality in the context of civil litigation. The concept of liberty translates into an autonomy principle in civil litigation. Liberty requires that each case be treated individually rather than lumped together and that each individual have a right to participation in their case. Sampling admittedly runs roughshod over this ideal, resolving cases (whether formally or informally) without the possibility of an individual adjudication for most plaintiffs.

To say that the tension between liberty and equality is not easily resolvable is an understatement. Our system seems to be built on an uneasy balance between the two principles. For those who believe that liberty is fundamental and ought to trump equality at every turn—especially in adjudication—statistical adjudication will be an unattractive option. But those who recognize that both principles are important to a system of adjudication in a democracy will admit that it is necessary to get beyond the reflexive preference for participation and look with a skeptical eye at the inadequate way our system currently promotes both the principles of liberty and equality. Given the current state of process scarcity, it is unlikely that we will move to a system of widespread individualized trials that a strong autonomy norm requires. We can, however, promote the kind of transparent, evenhanded system that is required by a robust principle of equality.
It should be noted at the outset that this Essay is not an efficiency-based defense of statistical adjudication or a cost-benefit analysis of statistical adjudication as compared with other methods, including both the current widely used lax method of anecdotal sampling and the day in court ideal. How much society ought to be willing to pay for fairness is an important question but one that is beyond the scope of this Essay.\textsuperscript{52} It is nevertheless necessary to keep in mind these two other options as we consider the fairness arguments in favor of statistical adjudication laid out in this section, because the argument here focuses on demonstrating that statistical adjudication is \textit{more fair} than these alternatives.

2.1 Equality

Equality is a central component of the rule of law. To the extent that similar cases have inconsistent outcomes, the principle of equality is being violated. Such inconsistency is not only the result of differing jury verdicts, but can also be true of settlement outcomes. Because settlements are difficult to study as a systemic matter and empirical evidence is scarce, we happen to know more about inconsistency in jury verdicts. But there is good reason to believe that settlements share this problem. The discussion below explains that equality is a value in our legal regime and demonstrates that statistical adjudication is better at realizing this principle than individualized adjudication under current conditions. Because it is more evenhanded, statistical adjudication is fairer than the alternative.

2.1.2. Consistency and the Rule of Law

One of the fundamental requirements of the rule of law is equality or the requirement that like cases be treated alike. The requirement of equality has been justified on fairness grounds and legitimacy grounds.\(^{53}\) Each side of the political spectrum espouses the view that inconsistency is unfair, raises concerns about bias and thereby undermines the legitimacy of the legal system. Critics of the tort system in particular often focus on the variability of verdicts, particularly on the high end.\(^{54}\)

There are a variety of procedural doctrines that attempt to instill the discipline of equality in our legal system. Preclusion prevents the inconsistent outcomes that would result from relitigation of cases. Remittitur similarly disciplines outcomes by allowing judges to reduce outlier verdicts. Both of these doctrines limit litigants’ day in court in favor of equality and consistency.

For example, claim preclusion forbids the relitigation of the same case. The doctrine of non-mutual collateral estoppel can prevent a defendant from relitigating its defenses in subsequent proceedings against new plaintiffs.\(^{55}\) These doctrines ensure that case outcomes are

\(^{53}\) “Like cases should reach like outcomes for purposes of fairness. If a dispute resolution system processes similar cases to disparate outcomes, there is something wrong with the process. Litigants will lose faith in adjudication as a means of dispute resolution if outcomes appear to be random, or worse, if they appear to be biased.” William Rubenstein, The Concept of Equality in Civil Procedure, 23 CARDOZO L. REV. 1865, 1893 (2002).


consistent. But claim preclusion only applies to the same claim being relitigated. Similarly, issue preclusion (collateral estoppel) applies only to the same issue being relitigated by the same party. How broadly or narrowly courts define the claims or issues at stake in a particular litigation will determine whether one outcome supersedes future litigation. Even if a court were to preclude subsequent litigation and thus prevent inconsistent outcomes in a particular case, there is no guarantee that a particular outcome was consistent with the value assigned to the underlying injury in comparable cases. Because the concept of value in tort law is comparative, we can only determine value by reference to a portfolio of similar cases. We are still searching for an acceptable method of determining what makes cases similar for purposes of assigning value. Preclusion raises the fundamental problem of value in tort law: alike in what way?

The doctrine of remittitur is another way the procedural law enforces the principle of equality. Remittitur permits the court to reduce jury verdicts that it believes are unreasonably high with a lower value verdict from comparable cases.\textsuperscript{56} This type of truncation is similar to what statisticians do with outlier data points. Statisticians will sometimes recode the outlier datum or drop it altogether on the theory that it represents some kind of mistake.\textsuperscript{57} But the existence of the outlier may be an important clue to failures of the model being used.\textsuperscript{58} In fact,\textsuperscript{56} Dimick v. Schiedt, 293 U.S. 474, 486-487 (1935) (recognizing that remittitur withstands Seventh Amendment attack, but rejecting additur as unconstitutional). See also Joseph Kadane, \textit{Calculating Remittiturs}, 8 LAW PROBABILITY & RISK, 125 (2009).

\textsuperscript{57} See Wissler et al., supra note 15.

an outlier may indicate that the model is wrong. For example, omitting the outlier becomes a problem when the outlier represents not an aberration but the direction in which valuation is moving. The value assigned to a particular kind of injury may increase over time. If that is the case, realignment through remittitur will systematically prevent increases in the value of injury. Or if remitturs are randomly granted, such changes in case value may increase the difficulty of assigning value in future cases because this practice increases unpredictability. Furthermore, remittitur favors lower valuations of injury. Because remittitur is constitutional but additur is not, in the federal courts and at least some state courts remittitur is a one way ratchet.\textsuperscript{59} It lowers awards that are extraordinarily high, but never raises awards that are too low. If awards are being averaged after remittitur and that average is extrapolated across a population of cases, the average will be consistently lower than it ought to be if rigorous sampling methods are used.

Court engineered sampling tackles the problem of variability in case outcomes while avoiding the systematic reduction of damages awards associated with remittitur. Sampling techniques assist in solving the inequality problem by choosing observable and relevant variables across which cases can be compared. Unlike remittitur, court engineered sampling standardizes results over a large number of cases where the underlying similarities have already been measured. It equalizes outcomes across cases at either end of the distribution curve, rather than lopping off outliers only on the high end and ignoring inconsistency on the low end of the curve.

\textbf{2.1.2. Evenhandedness in Compensation}

If it is done well, sampling is consistent with our moral intuition that like cases ought to be treated alike. Two basic insights underpin the discussion that follows. First, equality between similarly situated litigants is a moral imperative as well as a central value to our procedural regime. A second, corollary insight is that cases that are different in legally relevant ways ought to be treated differently.

Plaintiffs in a mass tort are linked to one another, whether they like it or not. Because many cases settle and settlements are determined on a comparative basis, any individual’s outcome is dependent upon the manner in which other cases are resolved. This is true even in cases where there is no limited fund. For example, it has become something of a trend to require a most plaintiffs – as many as 95% – to accept the aggregate settlement in order for the settlement to go forward. Both the Vioxx and the WTC Disaster Site Litigation settlements included such a requirement.61

---

60 While the general intuition among participants in the legal system seems to be that most cases settle, it is actually difficult to determine what settlement rates actually are because of coding problems in the data sets maintained by court systems. See Gillian K. Hadfield, Where Have All The Trials Gone? Settlements, Non-Trial Adjudications And Statistical Artifacts In The Changing Disposition Of Federal Civil Cases, 1 J. OF EMP. LEGAL STUD. 705 (2004):

Inherent in the concept of comparative valuation is the fact that all the cases in the reference class are linked to one another. One settlement affects the price of another. Yet these cases also differ from one another. Some individuals suffer more because their injuries are greater or because injuries of similar severity have caused more damage to them and their family. In other words, the contingencies of social life and luck affect the extent to which an individual plaintiff is harmed. Yet correcting preexisting inequality is not a part of the doctrine of damages in tort law. Litigants ought to be treated equally, but they are not in fact the same and the law compounds some of these preexisting inequalities. Recognizing this, we ought to consider what procedural devices will provide true equality between litigants and to what extent the law enables or limits such considerations in the process of assigning value to individual cases.

Inequality is a critical issue in mass tort cases. Renowned special master Kenneth Feinberg confronted the question of social inequality when he oversaw the September 11th Victim Compensation Fund. Numerous people perished in the terror attack of September 11th and Feinberg was charged with compensating families and individuals who had opted out of the tort system and agreed to have their compensation determined by the Fund. Some were very rich and others very poor, but all experienced terrible losses. In his book on the experience, Feinberg concluded that he would have preferred to give identical amounts to all claimants than to have to quantify the value of human life in the aftermath of such a disaster, especially when those valuations reflected existing economic inequality.62

---

At first blush, a sampling regime would seem to violate the ideal of equality because in the process of extrapolating from a sample to the greater population, individual cases will by necessity be lumped into groups according to variables that are tested in the sampling process. When the groups are relatively homogeneous, then a simple extrapolation process such as averaging should yield a similar result for each plaintiff, which will be fair because they are all similarly situated. The greater the variation between group members, however, the greater the redistribution from high value to low value cases.

Where there is substantial variation among plaintiffs, the averaging will benefit the least harmed at the expense of the most harmed, which is unfair. For example, imagine a case where half the plaintiffs in a sample received $50 and the other half received $100 in damages. The average across the entire sample is $75. In that case, the plaintiffs who would have received $100 are subsidizing those who would have received $50 because the averaging process awards them $25 less than they would have gotten at trial. This is a problem from a moral point of view because the most harmed (the $100 plaintiffs) are required to subsidize those who are less harmed (the $50 plaintiffs). A procedure such as this is too unfair to be sustained. If the system grants plaintiffs the autonomy to exit, one would predict that the pool will suffer from adverse selection.\(^{63}\) The plaintiffs suffering greater harm will choose to opt out of the procedure and leave only the lowest value claims to participate because these severely injured plaintiffs will predict that they will be systematically undercompensated in the averaging process.

Where sampling yields a broad distribution of values, the case for averaging is very poor because of adverse selection on the part of plaintiffs. Plaintiffs with higher value claims will not want to participate in a settlement that averages their anticipated compensation with lower values because they will be systematically undercompensated. Each plaintiff can anticipate that other plaintiffs with higher value claims will not participate, causing the average compensation to be reduced, until finally only claims not otherwise worth litigating will be left. This unraveling only occurs, however, when plaintiffs know more about their own claims than the court knows. This may or may not be true in a given mass tort, as in many cases the outcome largely depends on the court’s rulings.\textsuperscript{64}

In some cases, sampling can solve this problem by finding observable and relevant variables that distinguish among groups of plaintiffs. If the court can create sub-groupings that are acceptably homogenous, then even the best off plaintiffs will be better off participating in the procedure. If the outcome variance is not accounted for by measurable criteria, however, the court and litigants have no basis for sub-grouping and a mass exodus is the predictable result (assuming of course that plaintiffs know more than the court about their claims). It is also probably the correct result. Too much meaningful variance in the sample means an extrapolation process based on averaging will not yield a fair outcome for all plaintiffs and is likely to cause greatest harm to the most damaged plaintiffs. This is because a system of averaging that includes both claims with low value (roughly speaking persons who were not harmed very much)

and high value (persons who were greatly harmed) will result in a redistribution of awards from high value to low value cases. To use the numbers previously proposed, in an averaging regime a case valued at $50 will receive $75 and a case valued at $100 will receive $75. This benefits the $50 plaintiff at the expense of the $100 plaintiff. Accordingly, if the percentage of plaintiffs participating is too low, the court will need to modify or scrap the sampling procedure.

Will plaintiffs consent to a sampling procedure? Whether a plaintiff will want to participate depends on the extent of his or her risk version, the rate of time preference and the amount of discount from expected value. A good sampling regime will reflect something very close to the expected assigned value of a case. In statistical adjudication, a plaintiff gives up the chance to receive a higher payment in exchange for a guarantee of an average payment (and quickly). If the variability of distribution within the categories is not too great, sampling will benefit the most harmed those who otherwise would receive less than they ought. Risk averse plaintiffs may benefit because, even if their expected payout may be lower, the reduction in risk is enough to offset the smaller expected award. Even risk neutral plaintiffs may find this in their interest because the process will get them earlier payment. The greater the reductions in future payouts, the less likely plaintiffs are to participate.

The current popularity of informal sampling as a method for reaching settlement despite inconsistent verdicts indicates that there is a lot of potential for sampling procedures to fairly resolve large-scale litigation despite this objection. This is because plaintiffs are not likely to know more about the probable outcome of their case than the court does. Even if lawyers were equipped to reliably predict jury verdicts, judges retain the power to reduce awards through motions for judgment as a matter of law, remittitur and on appeal. For example, the verdicts in the Vioxx cases that were tried to a jury varied substantially. Some cases resulted in fifty million
dollar verdicts, others in findings of no liability. Even cases decided in the same forum under the same legal regime were split. Of the five cases tried in New Jersey state court, three were defense verdicts and two resulted in multimillion dollar verdicts for plaintiffs.\textsuperscript{65} There appears to have been some reasonable means of distinguishing between these verdicts, as the parties came to a settlement in which Merck agreed to pay over $4 billion to plaintiffs according to a “grid” agreed to by lawyers for the parties.\textsuperscript{66} A calculator available at the Vioxx settlement website will even estimate settlement amounts based on the criteria developed by the lawyers.\textsuperscript{67} In sum, if the sampling procedure is well designed and the reference class appropriate, adverse selection will not defeat the procedure.

Averaging across a large plaintiff population where the results of the sampling exhibit variation will not lead to an unfair result because the variation in awards is not always justifiable. Accepting that valuation in tort cases is assigned (rather than discovered) and that therefore the valuation of tort cases should be conceived of as a reasonable range and not a specific number, averaging is justified. Averaging will reduce variation in compensation and truncate both the high end and the low end of awards, rendering a fair final result for all, especially the most harmed plaintiffs.


\textsuperscript{66} The grid provides a set amount of compensation to plaintiffs depending on various factors, such as injury type, injury severity and length of time taking the drug.

If the variability of the distribution within the reference classes is not too great, sampling will benefit the most harmed plaintiffs. The most harmed plaintiffs at the end of the hypothetical sampling process described above are those who are damaged $100 but receive an award of $50. Sampling will reduce the risk that a plaintiff will be assigned an outlier award that is lower than comparably situated plaintiffs. This same plaintiff gives up the chance to receive a higher award than similarly situated plaintiffs. This loss of the chance to obtain an award at the high end of the range is not unfair, however. Fairness requires that like cases be treated alike. The evenhandedness principle does not support an entitlement to a lottery for potentially greater but unjustifiable awards. Instead, equality favors the idea of limiting unjustified high and low awards in order to treat cases equally.

The fact that sampling mitigates unjustified variation in monetization of harms brings us to the central question of value in tort law: why do some cases receive greater compensation than others? If there is some observable and justified reason that some cases are in the $100 category and others are in the $50 category, then we must find a way to distinguish between the two types of cases. But what if the group suffered the same injury severity: suppose everyone suffers a $75 harm? If the difference between the two outcomes represents variability for reasons that we cannot measure or for reasons that we think are morally irrelevant, such as plaintiff’s physical attractiveness or race \(^{68}\) (i.e. “noise”), then redistributing the difference among plaintiffs is the fairest approach. This redistribution benefits the least well off (a $75 plaintiff who in fact

received $50) at the expense of the best off (a $75 plaintiff who got $100). When differences are attributable to noise, the process of averaging does not redistribute from most harmed to least harmed. Instead, this system distributes awards equally among those (more or less) equally harmed. The key here is to accept the sufficiency of treating equally plaintiffs who are “more or less” equally harmed, rather than requiring complete identity between plaintiffs. In other words, it requires our legal system to accept that the values assigned to tort cases can be extrapolated across populations of similarly situated individuals because those values are assigned rather than inherent.

The problem courts face in this context is uncertainty. This uncertainty is due to variables that are not readily observable. The problem of valuation in tort law is one of unknown unknowns. As demonstrated earlier, studies have found that tort jury verdicts vary for reasons that we do not understand. The unavoidable uncertainty in jury verdicts is an argument in favor of extrapolation if one thinks that variation is not warranted by the individual circumstances of the cases, such as when the quality of one party’s lawyer or her race or gender are the reasons for differences in compensation, in contradistinction to reasons that track the requirements of the substantive law.

69 “As we know, there are known knowns. There are things we know we know. We also know there are known unknowns. That is to say we know there are some things we do not know. But there are also unknown unknowns, the ones we don't know we don't know.” Donald Rumsfeld, Defense Department Briefing (February 12, 2002).

70 To the extent that experts can predict the effects of these variables on verdicts, they need to be included in the model.
So far we have two paradigmatic cases, the outlier with a legitimate claim to greater damages (e.g. the concert pianist) and the outlier without a legitimate claim to greater damages (e.g. the plaintiff with a superior lawyer or who is the subject of prejudice). In real life things are a lot more complicated and cases are largely found in the middle ground, such as people with preexisting conditions who may be awarded a lower amount either because of the reduced life expectancy or because jurors have a tendency to be less sympathetic to the underlying condition. The answer to these more complex problems can probably be found in qualitative research into the basis for jury decisions. But in the absence of data some middle ground factors which have unknown or suspected effects on verdicts may have to be considered to be “noise.”

In sum, a sample picked randomly from the correct reference class will yield fair results, so long as it does not systematically devalue certain categories of claims for socially undesirable and legally impermissible reasons. It cannot correct for underlying social inequality because this is not a concern of the substantive tort law. The ideal is that a procedural system should obtain for plaintiffs what they are entitled to under the substantive law. That said, the law is not static. Adoption of more formal procedures of statistical adjudication may lead to changes in the substantive law, such as a relaxation of the individual proof rule, but is unlikely to result in seismic shifts in the purpose of the tort system.

2.1.3. Fairness in Timing and Case Management

The order in which courts hear cases is a critical component of procedural justice. The Federal Rules of Civil Procedure are to be “administered to secure the just, speedy, and inexpensive determination of every action and proceeding.”71 These considerations are in

---

tension with one another. One can easily imagine a very inexpensive summary proceeding that would not pay any attention to individual issues in cases. Such a regime would be unjust not only because it limits participant autonomy, but also because it does not even attempt to give plaintiffs what they are entitled to under the substantive law. Similarly, with enough resources it is possible to hold an individual trial in every case. Our system has resolved this tension by strengthening the procedures available for settling cases in the pretrial phase and giving judges additional discretion to resolve cases before trial, through motions to dismiss, summary judgment, judgment as a matter of law or orders for a new trial.  

Even cases resolved prior to trial can take a long time to reach a pretrial stage where they can be settled or resolved by summary judgment. Such delays can have negative effects on both parties. A prominent theorist explains: “Delay may have two different effects on a decision: it may undermine accuracy in the sense that it increases the risk of error, and it may undermine the practical utility of judgments for the purpose of redressing rights.” As time passes, events recede into the distance; witnesses forget, disappear or pass away. The passage of time compounds the injury for plaintiffs awaiting compensation. It also makes it more difficult for lawyers working on a contingency fee, who must fund the litigation going forward.

For defendants, the effect of delay is more mixed. On the one hand, defendants in mass tort cases may suffer as investors wonder what the effects of large-scale litigation will be on the

---

72 See Fed. R. Civ. P. 12(b) (motions to dismiss), 56 (summary judgment), 50 (judgment as a matter of law) and 59 (new trial).

company. On the other hand, delay often favors defendants who benefit from any difficulties plaintiffs may have in proving a case years after the fact. Defendants also benefit from putting off payment of damages, should they be found liable. Finally, in some cases the passage of time permits changes in scientific thinking to coalesce, clarifying the causation inquiry.\footnote{See, e.g., \textit{In re Breast Implant Cases}, 942 F. Supp. 958, 961 (E. & S.D.N.Y. 1996) (finding scientific evidence inadequate to prove breast implants caused systemic injuries alleged by plaintiffs).}

Docket management is one of the main reasons for the development of informal methods of aggregate litigation. The traditional approach to docket management is first come, first served: litigants are heard in the order that their cases were filed. But a sampling procedure allows the court to change the order that cases are heard because of the wide discretion it affords the judge. Courts may choose to favor litigants who have suffered the most. In the WTC Disaster Site Litigation, for example, Judge Hellerstein ordered bellwether trials of plaintiffs who suffered the greatest damages.\footnote{Order Amending Case Management Order No. 8, \textit{In re World Trade Center Disaster Litigation}, 21 MC 100 (Feb. 19, 2009).}

Sampling allows the judge to ameliorate the twin problems of delay. First, by getting some cases underway, evidence is brought forward while it is fresher in witnesses’ minds. Second, by favoring the most damaged plaintiffs in the order of cases tried or resolved, it limits the most egregious effects caused by the wait for compensation.

Arguably, the decision to resolve cases on a first come, first served basis as our system currently does is fair in much the same way a lottery is fair. Each litigant, no matter their
importance outside the judicial system, the subject matter of their case or the extent of their injury, will be heard in turn. Furthermore, to the extent that the timing of filing can be very roughly correlated to the timing of the injury, it makes sense to allow those injured first to be heard first. This approach (very roughly) solves the problem of fairness timing by moving individuals’ cases forward in the order that they were injured. Thus all injuries are treated equally in a formal sense. It does not take special account of those who suffer most by waiting, either because of evidentiary problems caused by delay or because the cost of waiting is higher for them than for other litigants. For the most harmed, our system allocates the timing of litigation poorly. This problem is aggravated in cases where individuals suffer injuries in and around the same time period. In that case, there does not seem to be a good justification for privileging the savviest litigants or the fastest filers when other individuals, who are not as quick or savvy but have been harmed to a greater extent, languish. Instead, the court should approach such litigation as if all cases are filed at once.

If the court treats all cases as simultaneously filed, it then faces the question of what other criteria ought to be used to determine priority. Courts ought to order the cases in a way that most comports with the ideal of litigant equality and with the principle that social institutions such as the courts should be used for the common benefit. One option is to choose the cases heard first

---

76 As a formal matter, jumping the line requires a showing of some kind of imminent harm. See Fed. R. Civ. P. 65 (preliminary injunctions and temporary restraining orders). As a practical matter, judicial docket control adds some discretion into the order in which cases are heard.

77 First filed would be a fair approach in cases where the timing of manifestation of injury correlates to harm (that is, the most injured file first). It is less fair in cases where the weaker cases are filed first.
on a random basis. This would give everyone an equal opportunity to be heard in a timely manner. A random selection would compound the harm of those who were severely damaged by delaying compensation. A lottery, therefore, would be a detriment to the least advantaged plaintiffs.

There are serious difficulties with the most-harmed-first approach. First, there is the difficulty of determining who is the most harmed. Is it one who would be entitled the highest award under the substantive law or one who perhaps is entitled to a lesser award but enters the legal system with the greatest need for their award? This question returns us to the problem of social luck which goes largely unaddressed by the tort system. There is an opportunity for judges to take such considerations into account in case management because it is largely discretionary. Whether this is justifiable is debatable. But if we set aside this problem for a moment, there is no justification for treating those with less severe harms more favorably than those with more severe harms. Doing so increases the inequality already suffered by the severely harmed. Accordingly, a court is justified in choosing the most severe cases to be heard first, with the court’s challenge being how to define the meaning of “most harmed.”

A second challenge to the most harmed first approach is that it appears to conflict with the principle of randomization. If the most harmed are disproportionately selected for trial, then the sample is by definition not random. The reason there is no conflict is that severity of harm ought to be one of the parameters of the appropriate reference class. A rigorous sampling process would not average the results of sample trials of substantially harmed individuals with those of individuals who were only minimally harmed. Instead, the population of similarly harmed individuals would constitute a reference class and the sample would be randomly chosen from that group. To use our earlier example, the court would survey the litigants to determine
who were concert pianists and sample from that group first. The results of that sample would be extrapolated to all the concert pianists. Then the court would randomly sample from the population with lesser harms, the white collar workers, and the results from those trials would be extrapolated to the other white collar workers.

The “most harmed first” approach was adopted by Judge Hellerstein in the WTC Disaster Site Litigation. Instead of hearing cases on a first come, first served basis, he singled out the most harmed plaintiffs and sampled from that group. This required substantial data collection, which is necessary in any event in order to determine the appropriate reference classes. In the WTC Disaster Site litigation, this data came largely from plaintiffs’ themselves (and their attorneys). Self-reporting creates some possibility for sloppy, mistaken or even fraudulent reporting and the court must design incentives to prevent this behavior and systems for monitoring lawyers. The design of such systems if beyond the scope of this Essay, but it is important to recognize that successful policing of misrepresentation is an important part of fair statistical adjudication procedures.

2.2 Transparency

Transparency is an essential value in litigation for two reasons. First, as an independent principle, transparency in civil litigation is a necessary part of a thriving democracy. The civil justice system is a public good because it is the primary system of law enforcement in our society. Citizens should be informed of the workings of the court system, the manner in which
cases are resolved and the ultimate resolution of those cases.\textsuperscript{78} For this reason, the resolution of all cases, not only by judicial action but also by settlement, ought to be publicly available.

Second, transparency – at least between litigants – is a necessary condition for fairness. Sampling is inherently comparative. Any individual’s case is determined by reference to the other cases and in order to move forward participants need access to the universe of similar cases. This is markedly different from individual litigation where a repeat player may be, even ought to be, familiar with many cases, but there is no systemic necessity to know the universe of cases as a matter of institutional design.

A statistical adjudication procedure will lead to greater transparency to the public for several reasons. First, to the extent that sampling leads to trials, the process and the results of those trials will be publicly available. Second, at a minimum court rulings with respect to docket management and sampling procedures will be published or available at the courthouse. Already most decisions in high profile multi-district litigation may be accessed free of charge on court websites, although transcripts of hearings and expert reports are ordinarily not available online.

Third, because the judge, special masters or experts must articulate the reasons for pursuing a particular sampling regime, choosing reference classes and determining relevant variables, the reasoning behind these decisions will be available to litigants (at a minimum) and ought to be available to the public as well. As mentioned earlier, the Vioxx settlement claims

administrator created an online calculator that shows how a plaintiff with the relevant characteristics will be compensated under the settlement regime. The settlement in the WTC Disaster Site litigation is also publicly available, although there is no calculator for damages as of yet and the judge initially expressed concern that the valuation process for individuals was not sufficiently transparent so that litigants could not make informed decisions about whether joining the settlement would be beneficial for them. In that case the judge even held a fairness hearing, although there is no legal mandate to hold one in aggregate litigation.

Because judges must justify sampling regimes, sampling brings to the forefront and makes transparent usually unarticulated assumptions about what does and what ought to matter in evaluating compensation in tort cases. Although lawyers and judges may have an informal sense of these assumptions, they are rarely, if ever, publicized, and the informal senses may well

---


80 Mireya Navarro, Federal Judge Orders More Talks on 9/11 Deal, N.Y. TIMES (March 19, 2010) (writing that “Judge Hellerstein also said that the terms of the settlement were too complicated for the plaintiffs to be able to reach an “intelligent decision” on whether to accept it.”).

81 In re World Trade Center Disaster Site Litigation, Order Setting Public Meetings on Settlement, 21 MC 100, 102, 103 (AKH) (S.D.N.Y. July 19, 2010). The American Law Institute has proposed holding such hearings as part of its evaluation of the Principles of Aggregate Litigation. See ALI, PRINCIPLES OF THE LAW OF AGGREGATE LITIGATION (2009).
be wrong, even though the participants are experts.\textsuperscript{82} Sampling therefore leads to greater rigor in methodology, creates accountability through publicity and encourages the type of openness and dialogue that ought to be the hallmark of a civil justice system in a democracy.

**CONCLUSION: The Limits of Rough Justice**

This Essay has argued that the rough justice administered in mass tort cases can be *more fair* than the justice available in individual litigation, despite the fact that individualized litigation is often held up as an ideal. The reason that rough justice is more fair is that all tort cases – one-off cases as well as mass torts – are valued comparatively in our system. Statistical adjudication, for all its roughness, is a better method for comparing cases with one another than individualized litigation because it requires rigorous and thoughtful system design.

There are many problems in tort law that sampling cannot solve. Sampling cannot completely solve the problem of value in tort law. How to assign a value to an injury and what that value ought to be are issues on which people disagree. Where disagreement is substantial, a sampling procedure can only reflect that disagreement. The method by which the court encourages agreement as to valuation in itself reflects contested norms. Nevertheless, tort cases

\textsuperscript{82} W.M. Grove, W. M. & P.E. Meehl, *Comparative Efficiency Of Informal (Subjective, Impressionistic) And Formal (Mechanical, Algorithmic) Prediction Procedures: The Clinical-Statistical Controversy*, 2 *Psychology, Public Policy, and Law* 293–323 (1996) (demonstrating that “expert” opinions are not better (and are often worse) than very crude “statistical” predictions). Thanks to Peter Siegelman for this point.
must be valued. In that process, equality and consistency are fundamental norms which rigorous and transparent sampling methods can help achieve.

Sampling does not solve the fundamental problem of uncertainty that pervades tort law. Once we have determined which variables are observable and relevant, there are still unobservable variables that affect valuation: the known unknowns. And then there are variables that are both unobservable and unaccounted for: the unknown unknowns. Whether these are noise that ought to be ignored or legitimate bases for differentiating among plaintiffs is a matter for serious debate. Experts can minimize the perceived importance of these variables by finding some proxy for them, and that is one of the key requirements of the design of a good sampling procedure. The process of constructing a sampling procedure requires a theory of what relevant variables can be relied on to produce acceptable outcomes, and it is precisely this requirement that forces the parties to make public their beliefs, values, and assumptions.

Nor can sampling realize the autonomy values that also form an important part of our tradition. Statistical adjudication quite consciously favors equality over liberty. The right to participate in the resolution of one’s case is severely limited for many litigants under such a regime. For individualists, tipping the balance in favor of equality values in this way may be unacceptable.

Individualism ought not to be a trump against statistical adjudication for two reasons. First, liberty and equality are competing values in our civil litigation system. As I have argued above, there are both legal and fairness arguments in favor of equality as a significant consideration which ought not to be ignored in the process of institutional design. It is also true that equality is not a trump against liberty. This brings us to the second justification for rough
justice. The argument in favor of individualism often assumes that the current system is individualistic, in contrast with proposals such as statistical adjudication which are understood to be utilitarian and collectivist. But this is not an accurate depiction of our current system. Instead, the system we have now is a complex, private and largely hidden system where most plaintiffs obtain a second-order version of statistical adjudication through informal methods of sampling in any event. If rough justice is what our system does anyway, it should at least do it right.

To the extent that consent is the basis for justifying the outcomes of the current system, given its flawed results, consent requires transparency. Statistical adjudication is transparent, in contrast to individual litigation which is not. Consider the WTC Disaster Site Litigation as an example of the limits of consent as a justification for individualism to trump equality. The parties proposed a settlement that required approval of 95% of the plaintiffs in order to go forward. They presented this settlement to the judge, impliedly seeking his approval. The judge believed that the settlement did not provide sufficient protections for the litigants and told the parties so at a public hearing. Initially the parties balked at the judge’s actions, accusing him of overstepping his bounds and even appealing his decision to hold a fairness hearing. It was not long before the parties renegotiated the agreement and returned to the judge with a settlement that he ultimately found to be fair. Would most of the plaintiffs have consented to the initial settlement, which had a smaller settlement fund and where a larger portion of the plaintiff’s recovery was to go to their lawyers? It is by no means clear that in this context consent is

---

83 See Mark Hamblett, City Asks Circuit to Override Judges Rejection of 9/11 Pact, NAT’L L. J. (April 15, 2010).
sufficiently robust such that it justifies allowing collective settlements like this one without judicial oversight. A private settlement reached without judicial oversight and rigorous methods for determining value may be too rough to provide justice to individual plaintiffs.

The type of rough justice I have described in this Essay is a departure from the hidden world of tort settlement. It may be rough, but well designed statistical adjudication can offer a transparent and evenhanded justice.