

9-19-2012

How to Think About Errors, Costs, and Their Allocation

Ronald J. Allen
rjallen@northwestern.edu

Follow this and additional works at: <http://scholarship.law.ufl.edu/flr>

 Part of the [Civil Procedure Commons](#), and the [Legal Profession Commons](#)

Recommended Citation

Ronald J. Allen, *How to Think About Errors, Costs, and Their Allocation*, 64 Fla. L. Rev. 885 (2012).
Available at: <http://scholarship.law.ufl.edu/flr/vol64/iss4/1>

This Symposium is brought to you for free and open access by UF Law Scholarship Repository. It has been accepted for inclusion in Florida Law Review by an authorized administrator of UF Law Scholarship Repository. For more information, please contact outler@law.ufl.edu.

SYMPOSIUM ARTICLE

HOW TO THINK ABOUT ERRORS, COSTS, AND THEIR ALLOCATION

*Ronald J. Allen**

There is an ongoing, robust debate about the structure of litigation, and in particular, about access to the courts. For a considerable period of time, the mantra that the courts should be readily available to all the people so that people may present claims that their rights have been violated has dominated academic discourse and has, perhaps, significantly influenced the structure of litigation.¹ This conventional view—that the courts should be freely open to all—was dealt a blow by the *Iqbal*² and *Twombly*³ decisions, which imposed greater gatekeeping responsibilities on the federal district courts. These decisions predictably provoked a storm of protest, in large measure because they may indeed make it more difficult for many petitioners to have their petitions considered on the merits.⁴ But whether that result is a social harm or a social good depends on matters aside from simply winnowing the field of potential disputants—a point neglected by much of contemporary civil procedure scholarship. That scholarship has placed a laser-like focus on facilitating the bringing of claims, and in doing so, has made two serious errors: first, the scholarship fails to consider that litigation is but one small part of a larger social optimization problem; and second, the scholarship has a peculiar conception of errors and costs, including how to allocate those errors and costs. This brief Article provides the analytical background to these assertions.

“Primary behavior” and “litigation behavior” are conventionally thought of as distinct spheres with internal logics of their own. The former articulates rules governing everyday actions—from social interaction to structuring efficient economic behavior—and the latter governs the peculiar set of actions involved in litigation. Facilitating appropriate primary behavior is the overriding goal of social organization, and one of its main tools is substantive law. Litigation behavior is the effort to resolve disputes about inappropriate primary behavior or to reestablish the status quo following disruptions of the social fabric.

* John Henry Wigmore Professor of Law, Northwestern University; President, Board of Foreign Advisors, Evidence Law and Forensic Science Institute, China University of Political Science and Law.

1. See Arthur Miller, *From Conley to Twombly to Iqbal: A Double Play on the Federal Rules of Civil Procedure*, 60 DUKE L.J. 1, 3–5 (2010).

2. *Ashcroft v. Iqbal*, 129 S. Ct. 1937 (2009).

3. *Bell Atl. Corp. v. Twombly*, 550 U.S. 544 (2007).

4. See, e.g., Symposium, *Reflections on Iqbal: Discerning Its Rule, Grappling with Its Implications*, 114 PENN. ST. L. REV. 1143 *passim* (2010).

Resources devoted to litigation appear to most legal commentators as wasted resources—adding no value to society. Since litigation itself does not produce any useful good, litigation should obtain correct results as efficiently as possible. These aspirations are reflected in Rule 1 of the Federal Rules of Civil Procedure, which states that the rules of civil procedure “should be construed and administered to secure the just, speedy, and inexpensive determination of every action and proceeding.”⁵ These aspirations also appear in Rule 102 of the Federal Rules of Evidence: “These rules should be construed so as to administer every proceeding fairly, eliminate unjustifiable expense and delay, and promote the development of evidence law, to the end of ascertaining the truth and securing a just determination.”⁶ The principle animating these provisions is equal access to justice, particularly that even the indigent should not be disadvantaged when facing a wealthy adversary. Regrettably, life, as always, is complicated. Costs cannot be eliminated, and thus, the most important question is how to allocate those costs.

Primary behavior does not produce goods cost-free; producing goods creates waste products and risks harm to others. Naturally, reducing production costs encourages production, whereas raising costs has the opposite effect. Thus, if the producer can externalize some of its cost (for example, by dumping waste in a river or on a neighbor’s property), the producer’s actual cost of the good will not reflect its true social cost, which means—with regard to social utility—there will likely be overproduction of the good in question. By contrast, optimal production of social goods is facilitated by ensuring production at true social cost. This equilibrium is why it is important for the substantive law to align costs with behavior.

Litigation costs are generally believed to be socially perverse because they act as a tax on productive behavior. To some extent, this belief is true, but a costless legal regime would stimulate the production of its product—litigation—and possibly result in too much litigation. Although this may appear counterintuitive, one should remember that parties must make decisions as to how to dispute—in simple terms, whether to sue or negotiate. Everyone comes into contact and potential conflict with numerous people. Perhaps a neighbor plays music too loudly or neglects to dispose of trash correctly. If litigation were costless one would simply sue rather than negotiate.⁷ The costs of litigation affect the manner in which people relate, and those effects can be beneficial or perverse. The costs of litigation, in short, may counterintuitively

5. FED. R. CIV. P. 1.

6. FED. R. EVID. 102.

7. When transaction costs for private negotiation and settlement are relatively low, parties will negotiate with each other to resolve conflicts. See R. H. Coase, *The Problem of Social Cost*, 3 J. L. & ECON. 1 (1960).

produce social goods through the incentive effects they create for alternative modes of disputing.

The precise policy prescriptions following a deeper understanding of the problem of social cost are ambiguous, because they depend, in part, on the relative values of resolving different kinds of disputes in different ways. It may be sensible to guide certain types of disputes toward formal dispute resolution and to guide others away from it. Maybe commercial disputes differ from family disputes, and maybe discrete commercial transactions systemically differ from antitrust actions. Life, in short, is complicated, and one of the tasks for the legal system is sorting out that complexity.

The history of both the Federal Rules of Procedure and the Federal Rules of Evidence reflect, at least implicitly, these analytical points; Congress enacted them in part to offset what it believed to be distorting aspects of the systems that they replaced.⁸ Lawmakers believed the previous systems disadvantaged plaintiffs by raising their costs much too high.⁹ The solution to this problem was to simplify pleading requirements and to allow cases to proceed to what pundits believed would be low-cost discovery, followed by low-cost trials.¹⁰ Discovery costs would be low because the assumption was that both parties shared knowledge of the typical cases, and thus, a substantial investment in discovery would not be required. In addition, both parties would have the incentive to keep costs of discovery to their necessary minimum. It is immediately obvious how these conceptions map onto the previous analytical points. “In a world of symmetrical information and low transaction costs . . . the Federal Rules most likely accomplished the goal of facilitating the accurate and efficient resolution of disputes without distorting the underlying substantive law, values that the procedural regime the Federal Rules replaced did not adequately secure.”¹¹ If the original assumptions about litigation are true, procedural wrangling serves no purpose.¹² Moreover, costs were not, and could not be, lowered to zero, so there remained reciprocal incentives to avoid litigation through other means of resolving disputes.

One should note the historical contingency of the era that adopted the Federal Rules of Civil Procedure. It involved substantive assumptions about the relative positions of plaintiffs and defendants that were empirically true

8. See Ronald J. Allen & Alan E. Guy, Conley as a *Special Case of Twombly and Iqbal: Exploring the Intersection of Evidence and Procedure and the Nature of Rules*, 115 PENN. ST. L. REV. 1, 7–8 (2010).

9. *Id.*

10. *Id.*

11. *Id.* at 11–12.

12. *Id.* at 12.

but not logically entailed.¹³ Thus, changes in the relative positions of plaintiff and defendant from the pre-Rules situation may justify changes in the procedural context, which could entail, among other things, a reallocation of costs.¹⁴ Perhaps originally, the procedural regime favored defendants and thus subsidized socially wasteful activity; however, now, in some set of cases, perhaps this regime favors plaintiffs with the opposite effect.¹⁵

In such cases, defendants will be deterred from productive activities, not by the law, but by litigation costs that increase the *in terrorum* [sic] value of even meritless suits that put pressure on a defendant to settle and burden otherwise lawful conduct. Potential defendants will engage in litigation avoidance tactics that are likely to be socially wasteful, and they will settle to avoid litigation costs rather than risk liability on the merits.¹⁶

The high price of litigation increases the cost of socially useful activity that is indistinguishable from socially costly activity at an acceptable price through litigation.¹⁷ The alternative is to buy peace through settlements even though the underlying primary behavior is perfectly acceptable. The effect is a tax on useful behavior.

To generalize, the legal system must take into account the interactive effects of primary behavior and litigation behavior. The effects or consequences of primary behavior on litigation behavior are often noted, but litigation behavior affects primary behavior as well.¹⁸ This means that the regulatory problem is unlikely to be solved by simple slogans such as those concerning access to court. Before addressing how to approach regulating such a complex problem, another issue involving the inadequacy of the conventional understandings of the litigation matrix needs to be addressed. In addition to inadequately considering the relationship between primary behavior and litigation behavior, the conventional conception of an error is inadequate.

The conventional conception of an error is composed of two parts: denying a petitioner access to an adjudication on the merits (through

13. *Id.* at 11–13.

14. *Id.* at 13–14.

15. *Id.* at 14.

16. *Id.*

17. *Id.*

18. See, e.g., Gideon Parchomovsky & Alex Stein, *The Distortionary Effect of Evidence on Primary Behavior*, 124 HARV. L. REV. 518, 547–48 (2010) (arguing that evidentiary concerns about prospective litigation have pervasive effects on primary behavior).

narrowing the courthouse door)¹⁹ and a belief that Type I (false positive) and Type II (false negative) errors²⁰ are roughly equivalent. These parts add up to suggest that the procedural goals should be to treat parties roughly equally and to minimize total errors.²¹ Although these ways of thinking have been around for a considerable period, it is plain that they suffer from serious defects.

First, an error is made each time an undeserving litigant imposes costs on an adversary, a point that seems, rather remarkably, to have been neglected by those who complain of the Supreme Court's recent forays into procedural matters. The image of the federal court system (or any other court system, for that matter) being constantly open and easily accessible for all neglects that walking through the courthouse door by a plaintiff imposes costs on a defendant. If the defendant has behaved inappropriately by reference to the substantive law, the defendant should bear these costs. But as elaborated above, if the defendant has not behaved inappropriately by reference to that same substantive law—if a plaintiff's claim is unjustified—the costs imposed on defendants are errors that create taxes on productive behavior and thus are likely socially perverse. This point is so obvious that it needs little further elaboration. An undeserving plaintiff deprives a deserving defendant of its assets, and the best-case scenario is that the deserving defendant passes those costs on to a hapless public. The best-case scenario, in short, is decidedly unappealing. The point, of course, is that the conventional view seems dominated by the belief that there are no wrongful complaints filed, which is ludicrous. More importantly, in an era of asymmetrical costs, where filing a complaint can generate enormous costs for the defendant, the defendant will be consistently in the position discussed above, having to minimize extra costs attached to socially useful behavior, and having to pass whatever costs it cannot avoid on to someone else, if possible.

There is a second fundamental error in the conventional thinking about errors. It focuses on just two of the decisions that a court can reach at trial—an error for one side or the other—but there are four possible outcomes at trial, and all have social benefits or costs. In addition to errors, correct decisions are possible. Neglecting correct decisions is peculiar. For example, in civil cases, the error equalization

19. See Miller, *supra* note 1, at 9–10 (discussing how changes in litigation strategy and judicial interpretation of the Federal Rules have influenced a “retreat from the principles of citizen access” to courts).

20. For more on Type I and Type II errors in connection with pleading standards, see Allen & Guy, *supra* note 8, at 6–7.

21. See, e.g., David Kaye, *The Limits of the Preponderance of the Evidence Standard: Justifiably Naked Statistical Evidence and Multiple Causation*, 7 AM. B. FOUND. RES. J. 487, 488, 496–500 (1982).

policy is satisfied by making errors in every single case, so long as the base rates of cases that go to trial include roughly the same number of deserving plaintiffs and defendants.

The relationship between the four possible outcomes at trial and procedural regulation is itself more complicated than it appears on the surface. In general, without knowledge of the base rates of deserving parties that go to trial and the relationship between the assessments of fact-finders and true states of affairs, there is literally no way to predict the effect of procedural regulation on correct or incorrect decisions. For example, implicit in the conventional discourse is that a finding that the probability of liability is 0.8 means that in eight out of ten similar cases, the true facts are consistent with liability. However, there could be any relationship between fact-finders' findings of probability and true states of affairs. In the set of all cases where fact-finders find a 0.8 probability of liability, it could be true that all cases in that subset are cases fact-finders should find no liability. Similarly, if everyone who goes to trial is guilty or liable, there can be no convictions of the innocent or mistakes against deserving plaintiffs, no matter how low the standard of proof, and vice versa.

The conventional discourse on procedural regulation also assumes a static system, whereas in fact it is dynamic. One aspect of this dynamism is that parties decide which cases to take further into the procedural system and can adjust their decision in light of changes in the rules. Thus, the simple assumption that changing the burden of pleading or persuasion, or any other part of procedure, causes more errors of one kind than another, or any other suggested cause-and-effect relationship between regulations and outcomes, is obviously not analytically true; it depends on how the system responds to the change.

The combined effect of the neglect of the interactive relationship between primary and litigation behavior and the curious conception of an error is obvious. The result is to obscure the fact that trial decisions are only one part of the output of the legal system. Parties negotiate outcomes in both civil and criminal cases. They do so in the shadow of trials (among other things), but the outcomes in those cases are obviously part of the total social welfare effects of a legal system. In addition, parties make those decisions in a dynamic, not a static, environment, which leads to the question how to regulate such complex processes most effectively.

In the abstract, the answer is clear. How to translate the abstractions into feasible regulation is another matter. First, the abstract answer is addressed in the quote below from my recent Meador Lecture, which is followed by my further reflections on social optimization of the procedural system:

[T]he reality of the legal system is that it is not a nice, tidy, simple, and static context but instead is a bubbling cauldron of messy, complicated, organic, evolutionary processes. The standard tool used to regulate this bubbling mess is rules, and it is the friction between that tool and many of the uses to which it is put that explains in general why fact-finding and legal regulation are viewed as so often problematic. This same relationship is explanatory of many legal puzzles, such as, in ascending order of importance, the curious implications of standard legal error analysis, the rules versus standards debate, and the meaning of “law.”

The simple concept of a rule as setting necessary or sufficient conditions from which outcomes may be deduced is an example of monotonic logic in which the addition of postulates or assumptions simply adds to what may be deduced from the previous assumptions. Monotonic logics are powerful tools, as the rise of modern mathematics and the success of many scientific fields demonstrate. They work best when their operant assumptions accurately capture their domains, which means they work quite well, in Hayek’s famous dichotomy, in made systems such as games, and less well in grown or organic systems, which typify much of the human condition. A large part of debate over rules and their limits is often implicitly about the complexity of the relevant domain and one’s tolerance for mistakes of different kinds. As the number of pertinent variables increases or when some of them are continuous rather than discrete, the deductive problem quickly becomes computationally intractable, even for computers let alone humans. And of course if a new variable pops up that was not previously anticipated, all deductive bets are off, as it were. In either case (computational intractability or failure of imagination), algorithmic approaches that rely on extant rules generate the standard critiques of the indeterminate nature of rules. In reality, it is not that rules are indeterminate but that they are being put to a task for which they are not optimal.²²

That lecture suggested that the central problem of the legal system is similar to the central problem of rationality, which is the taming of complexity. In both cases, simple deductive tools were being put to uses that were suboptimal. That raises the important question of what other

22. Ronald J. Allen, *Rationality and the Taming of Complexity*, 62 ALA. L. REV. 1047, 1060 (2012) (footnote omitted).

approaches may be more fruitful. Inspired by a brilliant article by an artificial intelligence researcher, Tim van Gelder, I explored one possible answer. The struggle of rationality to tame complexity may be less like digital computation and more akin to a dynamic regulator, such as the Watt Centrifugal Governor that was a critical part of the Industrial Revolution.²³ Analogously, legal analysis may need to evolve to deal with the complexities of systems. Van Gelder's example is a metaphor rather than an argument for my purposes, for it provides just the suggestion of possibilities rather than a defined research program. Nonetheless, it is interesting.

The dynamic regulator solved a very interesting problem. The growth of the textile industry in England depended upon a consistent energy source with very limited variability.²⁴ The steam engine provided the energy but its pistons provided episodic bursts of energy rather than a smooth, continuous stream.²⁵ Flywheels were helpful, but still not adequate. As van Gelder made clear, one potential solution to this problem is computational:

1. Measure the speed of the flywheel.
2. Compare the actual speed against the desired speed.
3. If there is no discrepancy, return to step 1.
Otherwise,
 - a. measure the current steam pressure;
 - b. calculate the desired alteration in steam pressure;
 - c. calculate the necessary throttle valve adjustment.
4. Make the throttle valve adjustment.
5. Return to step 1.²⁶

Unfortunately, this computational solution requires a costly person doing it, and it will rarely produce a smooth enough source of energy.²⁷ Scottish engineer James Watt solved this problem by placing movable arms on a spindle at the center of the flywheel. The flywheel's motion was instantaneously transmitted to the valve regulating the flow of steam.²⁸ As the rotation of the flywheel speeds up, the arms extend, which transmits energy to the valve and closes it until the proper equilibrium is reached, and vice versa.

Regardless whether the centrifugal regulator captures something important about rationality, viewing the legal system with this metaphor

23. Tim van Gelder, *What Might Cognition Be, if Not Computation?*, 92 J. PHIL. 345 (1995).

24. *Id.* at 347.

25. *Id.*

26. *Id.* at 348.

27. *Id.* at 347.

28. *Id.* at 349.

in mind may be fruitful. The most dramatic point is that some problems can be solved other than through deductive arguments or simple rules; the contrary belief is a consistent constraint on legal scholarship generally. It is undoubtedly useful to break problems down into smaller parts, and so on, but at the same time that process can be counterproductive, disguising rather than highlighting the nature of the entity under examination. The alternative is to think of the legal system more, perhaps, like fluid dynamics treats the flow of liquids and gases, to embrace, in other words, the messiness of real life rather than abstract it away.

How does this apply in the procedural context? Telling trial judges to behave as centrifugal regulators in order to optimize social productivity is probably not likely to yield satisfactory results. The second-best solution would be to assign the true costs of parties' actions to them. However, it is impossible to determine—practically and maybe theoretically—the “true” costs of litigation behavior. For example, when I ask for discovery, I may be trying to build my case or respond to the opponent's case. I should be responsible for building my case, but responding to my opponent's case perhaps is a cost that he should bear. When a lawyer cross-examines, whose costs are those? If it is pointing out the limits of the adversary's case, he should bear those costs; but if through cross-examination I am building my case, I should bear those costs. How could procedural law sort these different effects into the categories that are useful for one side or the other? A crude rule—opposing party pays for my costs of cross-examination—leads to obvious potential manipulation. Nor is adopting a British-style loser-pays system an obvious solution. Recent empirical work shows both that simple predictions about the effect of a loser-pays system are likely false (in fact, it can increase transaction costs), and that people do not opt for the English rule in contract negotiations.²⁹

Alternatively, the objective could be to structure the process so that the parties have the incentive to properly allocate costs, with, when necessary, the involvement of the trial judge. That objective would involve categorical cost allocation, with the possibility of relief from the trial judge. One category probably ripe for such treatment is discovery costs. Discovery costs generally benefit the party asking for the discovery, and discovery has been a cause of considerable injustice because of increasingly

29. Kong-Pin Chen & Jue-Shyan Wang, *Fee-Shifting Rules in Litigation with Contingency Fees*, 23 J.L. ECON. & ORG. 519, 522–23 (2007); Keith N. Hylton, *Rule 68, the Modified British Rule, and Civil Litigation Reform*, 1 MICH. L. & POL. REV. 73, 74 (1996); see also Theodore Eisenberg & Geoffrey P. Miller, *The English vs. the American Rule on Attorneys Fees: An Empirical Study of Attorney Fee Clauses in Publicly-Held Companies' Contracts* 1 (NYU Law & Econ. Research Paper Series, Working Paper No. 10-52, 2010), available at <http://ssrn.com/abstract=1706054>.

asymmetric allocation.³⁰ Plaintiffs, simply by filing, can impose enormous costs on defendants while bearing virtually none themselves.³¹ One should note how far from the original conceptions giving rise to the Federal Rules of Civil Procedure the modern condition may be. If each side will have about the same amount of discovery costs, it makes perfect sense to let each side bear their own costs. That is identical to cost shifting, and any resources spent in shifting costs are simply wasted. Asymmetric costs, by contrast, cause skewed cost allocation and provide the opportunity for strategic exploitation. By contrast, placing the costs of discovery provisionally on the person asking for it, but allowing for judicial involvement to make adjustments, may both generally give incentives for the optimal production of information and permit a safety valve in the unusual case.

Although the possibilities are diverse, an example of an “unusual” case would be where there is good reason to believe that an adversary is acting strategically, primarily to impose costs. In such a case, the “benefit” is to the adversary, and that is who should bear the costs. The optimal cost shifting would be accomplished by petitioning the trial judge for relief. In making such determinations, the adversarial process will construct the judge’s decision, and the parties will have the correct incentives to educate the trial judge. This is not a guarantee of perfection, but it provides some hope for reasonable outcomes. It exploits the advantages of both an initial “bear your own costs” scheme with the apparent inertia of trial courts that do not want to get involved with cost allocation or discovery regulation unless forced to do so. Courts would be forced to be involved only when the situation was egregious enough to justify a well-grounded petition for relief.

30. See Frank Easterbrook, *Discovery as Abuse*, 69 B.U. L. REV. 635, 636–37 (1989)

31. *Id.*