William & Mary Law Review

Volume 59 (2017-2018)
Issue 1

Article 3

10-1-2017

The Information-Forcing Dilemma in Damages Law

Tun-Jen Chiang

THE INFORMATION-FORCING DILEMMA IN DAMAGES LAW

TUN-JEN CHIANG*

ABSTRACT

Courts assessing compensatory damages awards often lack adequate information to determine the value of a victim's loss. A central reason for this problem, which the literature has thus far overlooked, is that courts face a dilemma when applying their standard information-forcing tools to the context of damages. Specifically, the standard method by which courts obtain information is through a burden of proof. In the context of damages, this means a rule requiring plaintiffs to prove the value of a loss. But courts will often face a situation where a plaintiff can clearly prove the existence of a loss, yet cannot prove the value of the loss with any precision. A court that strictly enforces the burden of proof would award zero damages in such a case, producing a harsh result. But a court that avoids this result by instead awarding its best guess at the correct amount—effectively forgiving the inadequacy of plaintiff's proof—then undermines future incentives for plaintiffs to produce rigorous evidence.

The result of this dilemma is that courts oscillate between strict and forgiving approaches, causing much confusion. Explaining the dilemma helps alleviate the confusion and points to a solution. In principle, courts should require a party to produce more rigorous damages evidence if, and only if, the party is the lower-cost provider of that evidence, and the benefit of having the evidence (in facilitating a more accurate damages award) outweighs the cost of collecting it. The messy legal standards for calculating damages in various

* Associate Professor, George Mason University Antonin Scalia Law School. Thanks to Michael B. Abramowicz, Adam Badawi, John M. Golden, Herbert Hovenkamp, Keith N. Hylton, Douglas Laycock, Jonathan S. Masur, Lisa Larrimore Ouellette, Hillary A. Sale, Brian Z. Tamanaha, and participants at the IP Scholars Conference, the Works-in-Progress IP Colloquium, and a workshop at Washington University School of Law for comments.

fields can be understood as clumsy attempts by courts to arrive at this unifying principle. Interpreting the messy doctrine in light of this "cheaper cost-effective producer" principle thus helps make damages law more coherent.

TABLE OF CONTENTS

INTRODUCTION

Compensatory damages awards are a matter of controversy in many areas of law, ranging from torts to contracts to patents.[1] Courts and commentators devote much attention to the problem, yet much of the debate is under-theorized and misdirected. Specifically, the literature often treats the problem as a matter of insufficiently clear damages standards,[2] and efforts at reform generally focus on devising better legal standards for calculating damages.[3] Such a diagnosis, however, misses the deeper problem, which is that courts lack the information to apply whatever legal standards might be devised. The literature has sometimes noted the paucity of information that courts possess in regard to damages;[4] but it has not

1. *See, e.g.*, Winsness v. M.J. Conoco Distribs., Inc., 593 P.2d 1303, 1305 (Utah 1979) ("The subject of certainty of proof as to [contract] damages has frequently concerned this Court and most others."); P.S. ATIYAH, THE DAMAGES LOTTERY 143-45 (1997) (arguing that tort damages awards are akin to a lottery); FED. TRADE COMM'N, THE EVOLVING IP MARKETPLACE: ALIGNING PATENT NOTICE AND REMEDIES WITH COMPETITION 162 (2011) (arguing that there is a "lottery-ticket mentality" around patent damages); Janet Cooper Alexander, *Rethinking Damages in Securities Class Actions*, 48 STAN. L. REV. 1487, 1488 (1996) ("It is difficult for the parties to estimate the size of the potential jury award with any degree of confidence.").

2. *See, e.g.*, S. REP. NO. 110-259, at 11-12 (2008) (arguing that patent "[j]uries are given little useful guidance" in determining reasonable royalty damages); Paul DeCamp, *Beyond* State Farm*: Due Process Constraints on Noneconomic Compensatory Damages*, 27 HARV. J.L. & PUB. POL'Y 231, 258-68 (2003) (arguing that "[j]urors [r]eceive [i]nadequate [g]uidance" and appellate review "[l]acks [p]rinciple or [p]redictability"); Mark Geistfeld, *Placing a Price on Pain and Suffering: A Method for Helping Juries Determine Tort Damages for Nonmonetary Injuries*, 83 CALIF. L. REV. 773, 776 (1995) ("The absence of well-defined standards for determining tort damages for nonmonetary injuries largely explains why such damages have been and continue to be a focal point in the debate over tort reform.").

3. *See, e.g.*, S. 1145, 110th Cong. § 4 (2008) (proposing a new standard for patent damages); Randall R. Bovbjerg et al., *Valuing Life and Limb in Tort: Scheduling "Pain and Suffering,"* 83 NW. U. L. REV. 908, 909 (1989) (proposing damages schedules to "constrain the operation of vague and open-ended legal rules and the wide latitude of discretion heretofore given legal decisionmakers"); Geistfeld, *supra* note 2, at 776 ("The absence of well-defined standards for determining tort damages for nonmonetary injuries largely explains why such damages have been and continue to be a focal point in the debate over tort reform.").

4. *See, e.g.*, John M. Golden, *Principles for Patent Remedies*, 88 TEX. L. REV. 505, 508 (2010) ("We really have little specific sense of what the value of remedies for patent infringement generally is or should be."); Robert P. Merges, *Of Property Rules, Coase, and Intellectual Property*, 94 COLUM. L. REV. 2655, 2665-67 (1994) (arguing that the lack of accurate information about damages for the typical IP right justifies property rule treatment).

considered in much depth why this problem persists, or, in other words, why courts cannot obtain better damages evidence.

One immediate answer to this question might be that courts cannot collect better damages information because better information does not exist—in other words, that human knowledge is inherently limited and courts are doing the best they can. A belief that the low quality of damages information is due to inherent epistemic limits, and is therefore uninteresting and not amenable to policy solutions, may explain why there is little discussion about the topic. As this Article will explain, however, inherent epistemic limits are not the only cause of low quality damages evidence. Another reason—one that better policy can improve—is that courts face an unappreciated dilemma when attempting to obtain better information about the value of many types of losses.

Specifically, the most obvious way for a court to obtain better damages information is to impose a rigorous burden of proof on the plaintiff—that is, to hold that the plaintiff must prove the true value of the loss with precision using rigorous evidence. But courts will often face a situation in which it is clear that the plaintiff has suffered *some* loss, but the plaintiff cannot prove the *amount* of the loss with any precision.[5] A court that strictly enforces the burden of proof would give zero damages in such a case, producing a harsh result.[6] A court that relaxes the rule—awarding the court's best guess at the correct amount—reaches a better result in the individual case, but undermines incentives for future plaintiffs to produce rigorous evidence.

One can vary the basic setup, but doing so does not escape from the tradeoff described. For example, a court could award an intermediate amount—more than zero, but less than the court's best guess at the correct amount.[7] This does not escape the dilemma but simply results in a mix of both negative consequences. Awarding an

5. *See, e.g.*, Gilbert v. Kennedy, 22 Mich. 117, 129 (1871) ("[C]ases will often occur in which ... no element of certainty can be found by which to measure with accuracy the actual amount of the damages, though it is evident ... that large damages have resulted from the injury.").

6. *See, e.g.*, Dobson v. Dornan, 118 U.S. 10, 18 (1886) (directing an award of six cents because plaintiffs could not prove the amount of profit attributable to infringement); Dobson v. Hartford Carpet Co., 114 U.S. 439, 447 (1885) (directing an award of six cents).

7. *See infra* Part III.C.

intermediate amount produces a *somewhat* unfair result that undercompensates the plaintiff, though not as severely as awarding zero, and it *somewhat* undermines incentives for future plaintiffs to produce rigorous evidence, though not as severely as fully awarding the court's best guess at the correct amount.

Another variation is to place the burden of proof on the defendant rather than the plaintiff—holding that the defendant must prove the correct amount of damages with precision, or be faced with a massive award (say, $1 billion).[8] But this merely changes who gets saddled with the potential harsh result without diminishing the dilemma: there will be many cases in which it is clear that the value of the loss is less than $1 billion, but the defendant cannot prove the exact number. A court that awards $1 billion will create an unjust result. A court that blinks—awarding its best guess at the correct amount—will reveal the threat to be empty, undermining incentives for evidence production in future cases.

Understanding the dilemma leads to three payoffs. First, it provides a framework to bring together doctrinal debates in many different areas—ranging from torts to contracts to patents—as manifestations of a common problem. Second, it explains why the doctrine surrounding damages throughout the law is messy and confused. The reason is that courts reacting to the dilemma will naturally oscillate between strict and loose approaches, depending on which side of the dilemma is more salient at a particular moment and in a particular case. A court facing complaints about speculative evidence and runaway juries will favor strictly enforcing the burden of proof to force plaintiffs to provide better evidence to support their claims.[9] A court facing outcries about deserving plaintiffs being denied compensation will favor loosening the burden of proof to avoid harsh outcomes.[10] The result is inconsistent and messy

8. *See infra* Part III.B.

9. *See, e.g.*, Garretson v. Clark, 111 U.S. 120, 121 (1884) (holding that patent plaintiffs must prove the amount of damages attributable to the patented feature using "reliable and tangible" evidence); Allegheny Pepsi-Cola Bottling Co. v. Mid-Atl. Coca-Cola Bottling Co., 690 F.2d 411, 415 (4th Cir. 1982) ("An antitrust plaintiff is entitled to recover only for that portion of a price reduction caused by the defendants' unlawful conduct.").

10. *See, e.g.*, Story Parchment Co. v. Paterson Parchment Paper Co., 282 U.S. 555, 563 (1931) (holding in an antitrust case that "it will be enough if the evidence show the extent of the damages as a matter of just and reasonable inference, although the result be only approximate," because "it would be a perversion of fundamental principles of justice to deny all relief

doctrine. Contrary to common portrayal, the difficulty of damages calculation is not *caused by* vague doctrinal standards;[11] rather, the inherent dilemma that courts face in collecting damages information *causes* messy and vague doctrine.

Third, once the dilemma is understood, a solution emerges. The solution comprises two principles. First, courts should only require a party to produce information when the social benefit of the information (in enhancing the accuracy and precision of a damages calculation) exceeds the costs of producing the information.[12] Second, courts should impose the burden of proof on the party that can produce the required evidence at lower cost.[13]

This "cheaper cost-effective producer" principle is obviously hard to implement in practice. My point is not to say that implementing this solution is easy or even feasible, but to articulate it as an organizing principle to understand what courts should be, and to some extent are, *aiming* to do.[14] Viewed from this perspective, the vagueness and inconsistency that characterizes damages doctrine across the law is not just aimless flailing around by courts, but represents clumsy attempts to arrive at a coherent principle. Articulating the cheaper cost-effective producer principle therefore provides a unifying lens with which to understand, interpret, and bring coherence to the messy doctrine that characterizes damages law.[15]

This Article is organized into four Parts. Part I first provides a description of the dilemma. Part II then provides several examples

---

to the injured person"); Westinghouse Elec. & Mfg. Co. v. Wagner Elec. & Mfg. Co., 225 U.S. 604, 616 (1912) (stating that the burden of proof "principle must not be pressed so far as to override others equally important in the administration of justice").

11. *See supra* notes 2-3 and accompanying text.

12. *Cf.* United States v. Carroll Towing Co., 159 F.2d 169, 173 (2d Cir. 1947) (defining reasonable care in cost-benefit balancing terms).

13. *Cf.* GUIDO CALABRESI, THE COSTS OF ACCIDENTS: A LEGAL AND ECONOMIC ANALYSIS 139-40 (1970) (arguing for allocating the burden of avoiding accidents to the cheaper avoider).

14. *Cf.* WILLIAM M. LANDES & RICHARD A. POSNER, THE ECONOMIC STRUCTURE OF TORT LAW 85 (1987) ("Hand was purporting only to make explicit what had long been the implicit meaning of negligence.").

15. The analogy to tort law illustrates the contribution. The importance of concepts like the Hand formula and the lower-cost-avoider principle is not that they are susceptible to direct implementation. Rather, the point is that they provide an analytical framework to understand what the vague and messy doctrine of negligence is about. *See* RICHARD A. POSNER, TORT LAW: CASES AND ECONOMIC ANALYSIS 2 (1982) ("This casebook is premised on the belief that the Hand formula ... provides a unifying perspective in which to view all of tort law.").

of how the dilemma underlies problems in damages calculation across many different fields, such as contracts, torts, patents, and antitrust. Part III takes the analysis further and explains that the dilemma not only directly manifests itself in damages doctrine, but can also be seen in various doctrinal reforms that attempt to solve the dilemma—all of which have severe shortcomings because courts and legislatures do not fully understand the underlying problem they are trying to solve. Part IV then presents a solution.

I. THE INFORMATION-FORCING DILEMMA

This Part will provide a description of the information-forcing dilemma and the dynamics by which it operates. Section A will first set the background by characterizing compensatory damages as an information problem. Section B will then describe how courts generally overcome information problems through the mechanism of information-forcing rules, specifically through imposing burdens of proof. Section C will then explain why, unlike other issues in litigation, obtaining information through burdens of proof leads to problems when it comes to assessing compensatory damages. Specifically, courts face a dilemma when a plaintiff who has clearly suffered some loss cannot prove the exact amount. A court in these circumstances must choose between strictly enforcing the burden of proof and giving the plaintiff nothing (a harsh result), or relaxing the burden of proof and undermining its information-forcing effect. Section D then more concretely illustrates how this dilemma operates by examining a pair of patent cases.

A. Damages as an Information Problem

Before explaining the difficulties courts face in collecting information, we must first frame the nature of the problem. It may seem obvious to state that courts need information to calculate damages, and that uncertainty in damages law today fundamentally traces to a lack of adequate information. But I will begin by defending this premise because much of the literature focuses on a somewhat different theory. Specifically, much of the literature treats the problem of unpredictable compensatory damages awards as arising from

the vagueness or inadequacy of substantive legal standards.[16] The literature accordingly focuses less on improving judicial information, such as by helping courts collect better evidence, and more on devising better and clearer substantive legal standards.[17]

To see why vague or defective substantive legal standards are not the fundamental problem, it is useful to remember that, although damages awards are often controversial in many areas, there are also many circumstances when they are *not* controversial. For many run-of-the-mill tort and contract cases involving goods that are transacted on a readily observed market, such as if *X* negligently damages *Y*'s car while it is parked on the street without physically injuring *Y*, the correct amount of damages is not difficult to determine and is not commonly subject to much dispute.[18] This is not because the substantive legal standard is different in such cases—the standard for compensatory damages is the same in virtually all areas of law where they are awarded: restore the victim to the same position as he or she would have been in but for the violation.[19] The

16. *See, e.g.*, Jutzi-Johnson v. United States, 263 F.3d 753, 759 (7th Cir. 2001) (Posner, J.) (arguing that the problem lies in "a standardless, unguided exercise of discretion"); S. REP. NO. 110-259, at 12 (2008) ("The current damage statute is vague and provides little guidance to judges or juries determining the proper damage award."); AM. TORT REFORM ASS'N, TORT REFORM RECORD 34 (2015), http://www.atra.org/wp-content/uploads/2016/11/Record-6-25-15.pdf [https://perma.cc/G6P9-ZXBU] (blaming "the broad and basically unguided discretion given juries in awarding damages"); Marcus L. Plant, *Damages for Pain and Suffering*, 19 OHIO ST. L.J. 200, 203 (1958) (diagnosing the problem as "the absence of any standard for measurement of the damages to be awarded").

17. *See, e.g.*, *Jutzi-Johnson*, 263 F.3d at 758 ("Various solutions, none wholly satisfactory, have been suggested [for calculating pain and suffering damages]."); S. REP. NO. 110-259, at 13-14 (proposing new standards for patent damages); William F. Lee & A. Douglas Melamed, *Breaking the Vicious Cycle of Patent Damages*, 101 CORNELL L. REV. 385, 389 (2016) (suggesting enhanced damages to deter patent infringement); Plant, *supra* note 16, at 210-11 (suggesting a cap); Ted Sichelman, *Purging Patent Law of "Private Law" Remedies*, 92 TEX. L. REV. 517, 554-60 (2014) (rejecting tort law remedies in patent law).

18. Robert N. Strassfeld, *If ...: Counterfactuals in the Law*, 60 GEO. WASH. L. REV. 339, 346 (1992) ("At bottom, all determinations of tort damages imply a comparison between the actual world and a counterfactual one in which the defendant had not injured the plaintiff. Often that comparison gives the trier of fact little pause.").

19. *See, e.g.*, Seymour v. McCormick, 57 U.S. (16 How.) 480, 490 (1853) (describing inquiry as "[w]hat a patentee 'would have made, if the infringer had not interfered with his rights'"); McDougald v. Garber, 536 N.E.2d 372, 374 (N.Y. 1989) ("The goal [of tort damages] is to restore the injured party, to the extent possible, to the position that would have been occupied had the wrong not occurred."); RESTATEMENT (FIRST) OF CONTRACTS § 329 cmt. a (AM. LAW INST. 1932) ("In awarding compensatory damages, the effort is made to put the injured party in as good a position as that in which he would have been put by full performance of the

difference is that, in the run-of-the-mill case involving a car, the information necessary to apply this standard is easy to collect. Courts can quite easily determine *what would have happened* in the counterfactual world (*Y* would have had an undamaged car), and they can also easily determine the *monetary value of the difference* because there is a readily observable market (both a repair market to fix the car and a rental market for a replacement in the interim). When courts have reliable information to determine the state of the counterfactual world and to monetize the difference, the legal standard for compensatory damages is not vague and is not difficult to apply.[20]

My point here is not that the critics are wrong about the legal standard being vague—it *is* vague and difficult to apply in some situations—but that the vagueness does not arise from the standard itself. Rather, the vagueness of the "but-for" standard traces to a deeper information problem. When there is no deficiency of information about the value of the loss, there is little vagueness in the legal standard. When courts lack information, either because they have difficulty determining what would have happened in a counterfactual world,[21] or because there is no observable market,[22] *then* the legal standard becomes vague and hard to apply.

Once we understand the problem of damages as fundamentally a problem of information, the next question becomes: How can we improve the quality of judicial information about damages? It is at this point that the literature on information-forcing mechanisms becomes relevant.

---

contract."); ABA SECTION OF ANTITRUST LAW, PROVING ANTITRUST DAMAGES: LEGAL AND ECONOMIC ISSUES 53 (2d ed. 2010) ("The essential starting point for any damages quantification is the but-for premise."); *see also* Hetzel v. Balt. & Ohio R.R. Co., 169 U.S. 26, 38 (1898) ("[T]he rule of damages should not depend upon the form of the action.").

20. *See* Strassfeld, *supra* note 18, at 346.

21. *See id.* ("Yet often we require the factfinder to engage in much more uncertain inquiries to establish the damage award.").

22. *See* RESTATEMENT (SECOND) OF TORTS § 912 cmt. b (AM. LAW INST. 1979) ("There is no market price for a scar or for loss of hearing.").

B. The Burden of Proof as an Information-Forcing Rule

If we were contemplating the question of how courts could collect information without any preexisting constraints, there would be many theoretical possibilities. For example, we could have judges make direct inquiries themselves, as in fact occurs in European inquisitorial systems.[23] Or we could select jurors who already have personal knowledge of the relevant facts, as occurred under early versions of the jury system.[24] Within our modern adversarial system, however, judges and juries are expected to eschew personal inquiry and instead obtain information from the parties that appear before them.[25] Specifically, courts require the party seeking redress—that is, the plaintiff—to plead and prove facts that support recovery.[26] By making a plaintiff prove his case, the law forces the plaintiff to disclose information to the court.[27]

This initial insight—that courts must force information from parties, and they accomplish this through the mechanism of a burden of proof—is general; it applies to all issues in litigation, including but not limited to damages. If a court wishes to discern whether accused defendant *X* is guilty of shooting victim *Y*, the way that the court obtains this information is not to have the judge personally visit the alleged crime scene and collect forensic evidence, but instead to force a plaintiff—or, in the criminal context, a prosecutor—to produce evidence to prove the proposition.[28] Somewhat

---

23. *See* John H. Langbein, *The German Advantage in Civil Procedure*, 52 U. CHI. L. REV. 823, 826 (1985) (explaining that, in the German inquisitorial system, "the court rather than the parties' lawyers takes the main responsibility for gathering and sifting evidence").

24. John Marshall Mitnick, *From Neighbor-Witness to Judge of Proofs: The Transformation of the English Civil Juror*, 32 AM. J. LEGAL HIST. 201, 201 (1988).

25. *See* 9 JOHN HENRY WIGMORE, EVIDENCE IN TRIALS AT COMMON LAW § 2483 (James H. Chadbourn ed., rev. ed. 1981) ("The apportionment of the task of adducing evidence is one of the most characteristic features of the Anglo-American system. It is placed wholly upon the *parties*."); 21B CHARLES ALAN WRIGHT & ARTHUR R. MILLER, FEDERAL PRACTICE AND PROCEDURE EVID. § 5122 (2d ed. 2017) (discussing "the American system of party initiation and party presentation").

26. *See* Bruce L. Hay, *Allocating the Burden of Proof*, 72 IND. L.J. 651, 652 (1997) ("[T]he plaintiff is almost always assigned the burden of proof."). As Part III.B explains, placing the burden of proof on the defendant does not fundamentally change the dilemma.

27. *See* WRIGHT & MILLER, *supra* note 25, § 5122 ("[T]he court can compel the parties to produce evidence by the threat that their failure to do so will lead to a decision against them on the merits.").

28. I am merging two senses of "burden of proof" that are sometimes distinguished: the

curiously, although there is a vast literature on information-forcing mechanisms in the law,[29] this literature only rarely considers the burden of proof as an instance of such a mechanism.[30] Properly understood, however, a burden of proof is not only an apt example of an information-forcing mechanism but also arguably the most important example: in the absence of a burden of proof, courts would have virtually no evidence and no information with which to make decisions.[31]

Because the burden of proof is the primary doctrinal mechanism through which courts obtain information, it follows that when courts seek better information on some issue, the way in which they attempt to accomplish this objective is by raising the burden of

---

burden of proof encompasses a burden of production and a standard of persuasion. *See* James B. Thayer, *The Burden of Proof*, 4 HARV. L. REV. 45, 48 (1890). The information-forcing effect I describe requires both to work in tandem. A party has an incentive to produce evidence only if it bears the burden of production. A party has an incentive to produce *quality* evidence only up to what is necessary to meet the standard of persuasion.

29. The literature on information-forcing is most developed in the context of contract default rules. *See, e.g.*, Barry E. Adler, *The Questionable Ascent of* Hadley v. Baxendale, 51 STAN. L. REV. 1547, 1553-54 (1999) (identifying limits of Ayres-Gertner model); Ian Ayres & Robert Gertner, *Filling Gaps in Incomplete Contracts: An Economic Theory of Default Rules*, 99 YALE L.J. 87, 91 (1989) (discussing "penalty default[]" rules in contract law that incentivize disclosure); Eric Maskin, *On the Rationale for Penalty Default Rules*, 33 FLA. ST. U. L. REV. 557, 557 (2006) (challenging penalty disclosure rationale of Ayres-Gertner framework); J.H. Verkerke, *Legal Ignorance and Information-Forcing Rules*, 56 WM. & MARY L. REV. 899, 906-08 (2015) (asserting "clause-forcing" rules because most people do not read boilerplate language).

30. The closest articles on this topic that I am aware of are Richard D. Friedman, *Dealing with Evidentiary Deficiency*, 18 CARDOZO L. REV. 1961, 1966 (1997) ("[A]llocation of the burden of production is a significant tool by which the courts can make it likely that the party with better access to the evidence will present it."), Bruce L. Hay & Kathryn E. Spier, *Burdens of Proof in Civil Litigation: An Economic Perspective*, 26 J. LEGAL STUD. 413, 413-14 (1997) ("Optimally used, the burden of proof may minimize the expenditures devoted to gathering, presenting and processing information in litigation."), and Dale A. Nance, *Evidential Completeness and the Burden of Proof*, 49 HASTINGS L.J. 621, 621 (1998) (arguing "that the burden of proof should reflect the need to avoid rendering a judgment on the basis of an evidentiary package that is unreasonably incomplete"). *See also* DALE A. NANCE, THE BURDENS OF PROOF: DISCRIMINATORY POWER, WEIGHT OF EVIDENCE, AND TENACITY OF BELIEF 156-58 (2016); John Leubsdorf, *Evidence Law as a System of Incentives*, 95 IOWA L. REV. 1621, 1642 (2010).

31. Rules of pleading serve a similar function, and some authors have applied the information-forcing insight to that context. *See, e.g.*, Samuel Issacharoff & Geoffrey Miller, *An Information-Forcing Approach to the Motion to Dismiss*, 5 J. LEGAL ANALYSIS 437, 439 (2013); Alex Reinert, *Pleading as Information-Forcing*, 75 LAW & CONTEMP. PROBS. 1, 3 (2012).

proof.[32] In the context of assessing compensatory damages, this translates into a rule requiring the plaintiff to prove the value of the loss with high certainty and precision (the two work out to be the same in this context, because uncertainty about the value of a loss translates into an imprecise estimate).[33] There are in fact cases that suggest a very high burden of proof when it comes to damages.[34] Understanding the information-forcing effect of a burden of proof helps flesh out the rationale underlying such decisions.

One point to clarify here is that, when courts require losses to be *proven* with precision, they are not merely requiring plaintiffs to get an expert to testify to some precise number. An expert who asserts that the plaintiff is entitled to $1 million may phrase his testimony in a very precise manner, but the naked assertion does not prove anything. What matters is not the style and phrasing of the assertion, but whether the expert can provide substantive evidence and analysis to back the assertion up—to convince the fact-finder that the true state of the facts is that the plaintiff suffered exactly $1 million in losses (and not $0.9 million or some other number).

---

32. In speaking of "raising" the burden of proof, I do not mean that a court will change the verbal formulation from "preponderance of evidence" to "clear and convincing" or "beyond reasonable doubt." I mean that a court will require more *certainty* before considering a factual proposition to be established. Formulations such as "preponderance of evidence" and "beyond reasonable doubt" are often useful to express increasing certainty in binary determinations such as guilt or liability (if one has no doubt about a defendant's guilt, then one is very certain of it), but they map awkwardly to increasing certainty when it comes to scalar determinations like valuing losses. A rule that "plaintiff must prove the value of a loss to the nearest dollar by a preponderance of evidence" requires more certainty and is a higher standard of proof than a rule that "plaintiff must prove the value of the loss to the nearest thousand dollars by a preponderance of evidence," even though both use the "preponderance of evidence" formulation.

33. *See* WILLIAM B. HALE, HANDBOOK ON THE LAW OF DAMAGES § 31 (2d ed. 1912) ("Losses must be certain in amount, ... or damages therefor cannot be recovered. The burden ... is on the plaintiff."); *see also* Dobson v. Hartford Carpet Co., 114 U.S. 439, 444 (1885) ("There is but one safe rule—to require the actual damages or profits to be established by trustworthy legal proof.").

34. *See, e.g.*, *Dobson*, 114 U.S. at 444; Garretson v. Clark, 111 U.S. 120, 121 (1884) ("The patentee ... must in every case give evidence tending to separate ... the patentee's damages between the patented feature and the unpatented features, and such evidence must be reliable and tangible, and not conjectural or speculative." (quoting Garretson v. Clark, 10 F. Cas. 45, 45 (C.C.N.D.N.Y. 1879) (No. 5,250))); French v. Ramge, 2 Neb. 254, 260-61 (1872); Griffen v. Colver, 16 N.Y. 489, 491 (1858) ("It is a well established rule of the common law that the damages to be recovered for a breach of contract must be shown with certainty."); Winston Cigarette Mach. Co. v. Wells-Whitehead Tobacco Co., 53 S.E. 885, 887-88 (N.C. 1906).

Thus clarified, it should be obvious that establishing a damages amount with precision is often very difficult. Even market-traded goods (such as a car) have some variance in their value—the price of a particular model of car will fluctuate by hundreds of dollars between different dealers; nonmarket items such as physical injuries are even harder to value with precision. The difficulty of precisely valuing losses means that courts need a great deal of information to do their jobs properly. It also means that plaintiffs have much difficulty supplying the information required.

C. The Dilemma in Forcing Damages Information

The prior Section explained why courts would be tempted to impose a high burden of proof on plaintiffs in order to force information. What this Section will explain is that such an approach has an important limitation when it comes to damages assessment. As a predicate matter, it is important to understand that a burden of proof is only effective as an information-forcing mechanism if it is enforced through a penalty for noncompliance. Rational plaintiffs will not produce costly evidence simply because courts say that plaintiffs have a duty to produce such evidence; they will only produce evidence if there is a private cost to the failure to produce evidence.[35] As a matter of standard common law principles, the way that courts enforce a burden of proof is to say that a party that fails the burden of proof will lose the issue. In the context of damages, this principle logically translates into a rule that a plaintiff who fails to prove the value of a loss with the required degree of precision should receive zero damages.

The problem is that a plaintiff will often be able to clearly prove that he has suffered *some* loss, but at the same time will be unable to prove the precise *amount* of the loss using rigorous evidence.[36] For example, if a defendant negligently destroys the *Mona Lisa*, then the Louvre Museum clearly has suffered some loss. But it

35. *See* Robert Cooter, *Prices and Sanctions*, 84 COLUM. L. REV. 1523, 1524 (1984) ("The mere existence of an obligation or issuance of a legal command may provide insufficient motivation for obedience. In addition, a sanction or threat may be needed to induce conformity." (footnote omitted)).

36. *See* Gilbert v. Kennedy, 22 Mich. 117, 129 (1871).

would be difficult for the Louvre Museum to prove the precise amount of this loss because there is no easily accessible market to establish the exact value of the *Mona Lisa*.[37] A court that strictly enforced a rule that plaintiffs must prove the value of the loss would then have to award zero damages, even though it is obvious that the correct amount of damages is not zero.[38] Such a result is problematic not only because it feels intuitively unfair, but also because denying compensation to victims who have suffered real losses frustrates the policy goals that compensatory damages are designed to serve.[39]

A court faced with the prospect of *actually* having to give zero damages to a plaintiff who has clearly suffered some (unquantifiable) harm, therefore, will feel a strong temptation to avoid the result.[40] That is, instead of giving an injured victim no redress, a court will be tempted to give an award by taking its best guess at the correct amount, based on whatever evidence is available, accepting that the result will necessarily involve some speculation.[41] But, as soon as the court does this, it undermines the information-forcing effect of the burden of proof.

The reason is that, although a court's guess will have a high *variance*, as long as the court is unbiased its *average* will center on the true value.[42] That is, if the true value of a loss is $50 million, a court taking speculative and uninformed guesses will often guess too high (for example, $90 million) or too low (for example, $10 million), but

37. One can imagine many possible ways to approximate the value, such as by looking to recent sale prices of other famous artworks. But no methodology will yield a precise value.

38. A plaintiff may still be entitled to other remedies, such as an injunction against future violations. *See, e.g.*, Blue Cross & Blue Shield United of Wis. v. Marshfield Clinic, 152 F.3d 588, 591 (7th Cir. 1998) (finding that the inability to prove future damages helps satisfy irreparable injury requirement for an injunction). The availability of other remedies does not alleviate the problem because the plaintiff is still being denied full compensation. *See id.* at 594.

39. For example, if the owner of a valuable but difficult-to-value painting knows that he is unlikely to receive damages for harm to the painting, then the owner will have an incentive to take excessive precautions, such as by not publicly displaying the painting.

40. *See* U.S. Frumentum Co. v. Lauhoff, 216 F. 610, 617-18 (6th Cir. 1914) ("To send the successful plaintiff away after years of litigation and with only nominal damages is repellent to the sense of justice.").

41. *See id.* at 626 ("The books are full of cases where damages or profits have been allowed upon reasoning that poorly satisfies the rules of proof.").

42. HOWARD G. TUCKER, MATHEMATICAL METHODS IN SAMPLE SURVEYS 99 (1998) ("If z is an observable random variable, and if λ is some real number, then z is called an unbiased estimate of λ if $E(z) = \lambda$.").

it is just as likely to guess too high as it is to guess too low. The errors will balance each other out, and the *expected* result is that the court will guess the true value. A plaintiff deciding whether to submit evidence—in a regime where courts always awarded their best guess—will therefore reason thus: If I spend money to collect rigorous evidence and thereby prove the true value of the loss, then a court will award me the true value of the loss as compensatory damages. If I spend no money and submit no evidence, a court will take its best guess, and on average it will still award me the true value of the loss as compensatory damages. Since the expected recovery is the same either way, but the latter course saves the cost of collecting evidence, I will not collect any evidence.[43]

Courts that face a situation where it is clear that there is some loss, but where there is little certainty regarding the precise amount of the loss, are thus subject to contradictory impulses: they have the temptation to impose a strong information-forcing rule to incentivize rigorous evidence, but also the temptation to not actually enforce the rule with harsh zero-damages penalties.[44] Of course, one cannot have the information without the enforcement,[45] and therefore courts face a dilemma. The result of the dilemma is that courts balance the competing considerations in an ad hoc and seemingly unprincipled manner,[46] making the case law messy and the doctrine self-contradictory.[47]

---

43. At least if the plaintiff is risk neutral. For risk-averse plaintiffs, see *infra* text accompanying note 51.

44. *See U.S. Frumentum*, 216 F. at 626 ("[T]he cases present a constant and uncertain contest between the desire to do what seems to be justice and the necessity of observing legal rules.").

45. *See generally* Meir Dan-Cohen, *Decision Rules and Conduct Rules: On Acoustic Separation in Criminal Law*, 97 HARV. L. REV. 625, 630 (1984) (exploring the concept of "acoustic separation" between rules of conduct and enforcement).

46. *See* DOUGLAS LAYCOCK, MODERN AMERICAN REMEDIES 104 (concise 4th ed. 2012) (noting the inconsistency in the case law and generalizing the state of play as the "plaintiff must prove damages with as much certainty as is reasonably possible under the circumstances," while noting that "'[c]ircumstances' vary" and "occasionally, [even meeting the above-stated standard] might not be good enough"). I say "seemingly" unprincipled because, as I will explain in Part IV, one can understand judges to be mostly acting in accordance with some implicit principles that they have not managed to articulate because they have not fully grasped the nature of the underlying problem.

47. *Compare, e.g.*, S.C. Johnson & Son, Inc. v. Louisville & Nashville R.R. Co., 695 F.2d 253, 261 (7th Cir. 1982) (affirming judgment of no damages because "the admitted fact of damage is insufficient to prove the amount of damage" (quoting Locklin v. Day-Glo Color

This dilemma is unique to the context of damages; the dilemma does not arise during the liability phase of litigation because liability determinations are binary.[48] That is, a defendant is either liable on a claim or it is not, so if a court is unsatisfied with the plaintiff's proof regarding liability, it is necessarily unconvinced that there is *any* liability. Finding no liability in such circumstances does not seem unjust because the result is not clearly wrong. In contrast, damages determinations have the characteristic that a court can be unsatisfied with the plaintiff's proof regarding the amount of a loss, while still being convinced that there is *a* loss. Awarding zero damages in such circumstances *does* seem unjust because the court knows the correct value of the loss is not zero. The unique existence of the dilemma in damages law means that the burden of proof works substantially less well as an information-forcing mechanism when it comes to damages assessments than in other areas.

The more inherently difficult it is to collect information about the value of a loss, the more acute the dilemma becomes. This creates a depressing irony in which the primary information-forcing mechanism that courts use (that is, the burden of proof) functions least well precisely when courts have the greatest need for it. When the value of a particular loss is inherently very easy to determine—such as in the prior example of a damaged car—then the dilemma poses little difficulty.[49] A court can impose a burden of proof on the plaintiff, and then the plaintiff will comply (because the cost to comply is very low), and so the unappealing prospect of actually having to impose a zero-damages penalty will not arise. It follows that doctrine is stable and damages awards are generally uncontroversial in those situations.[50] In contrast, when information about the correct value of a loss is inherently scarce and therefore very costly to collect (such as trying to value the *Mona Lisa*), then courts

---

Corp., 429 F.2d 873, 879 (7th Cir. 1970))), *with, e.g.*, Fontana Aviation, Inc. v. Beech Aircraft Corp., 432 F.2d 1080, 1085 (7th Cir. 1970) (reversing judgment of no damages because "[o]nce the fact is established, uncertainty concerning the exact amount of loss will not preclude recovery").

48. A. Leah Vickers, Daubert*, Critique, and Interpretation: What Empirical Studies Tell Us About the Application of* Daubert, 40 U.S.F. L. REV. 109, 124 (2005) ("The law ... generally forces its participants to render decisions and verdicts in binary pairs, such as causation/no causation and liability/no liability.").

49. *See supra* text accompanying note 18.

50. *See supra* text accompanying note 20.

will have a greater need to incentivize plaintiffs to produce better information by imposing a rigorous burden of proof. At the same time, the inherent difficulty of the task means that plaintiffs will more often fail to come up with the necessary evidence, even with strong incentives to do so. At this point the prospect of actually imposing a zero-damages penalty will present itself, creating the dilemma and—to the extent courts react by relaxing the burden of proof—undermining the effectiveness of the information-forcing mechanism.

Two qualifications should be added at this point. The first is that plaintiffs will lose *all* incentive to produce evidence only if plaintiffs are risk neutral, which is not a realistic assumption.[51] I begin with the risk neutral plaintiff merely as a way to clearly illustrate the competing impulses that underlie the dilemma. Incorporating risk aversion does not change the analysis significantly[52]: if the plaintiff is risk averse, then the plaintiff's incentive to collect evidence will not quite fall to zero because a court taking a speculative guess has higher variance than if the plaintiff provided affirmative evidence of the correct amount. The important point, however, is that a plaintiff's incentive to produce evidence is *weakened* when courts adopt a forgiving stance, even if the incentive does not always fall to zero. The dilemma thus remains: a court can force better damages information with a strict approach, or it can achieve more just outcomes in individual cases with a forgiving approach, but it cannot do both at the same time.

The second objection is that my analysis has thus far ignored the defendant's ability to present counter-evidence: I have posited a model in which plaintiffs either prove the correct amount of damages, or receive zero; or receive the court's best guess—there is no option for the court to award what the defendant argues. In the real world, of course, the adversarial system features a plaintiff producing evidence and arguing for damages of $\$X$ (where $\$X$ is usually higher than the true number), followed by a defendant producing

51. *See* W. Kip Viscusi, *Product Liability Litigation with Risk Aversion*, 17 J. LEGAL STUD. 101, 120-21 (1988) (finding that "[r]isk aversion plays a statistically significant role" in influencing plaintiff behavior in product liability cases).

52. *See id.* at 121 (finding that expected payoffs play a much greater role in influencing plaintiff behavior than risk aversion).

counter-evidence and arguing for damages of $\$Y$ (where $\$Y$ is usually lower than the true number), followed by the plaintiff producing counter-counter-evidence, and so on. Focusing only on the plaintiff seems to miss an important part of the story.

The reason that I focus only on plaintiffs in this Part is that this adversarial process must *start* with an initial burden of proof, and my analysis thus far is about the initial burden of proof. At this stage of the analysis, a defendant's counter-evidence is not yet relevant. If courts set the initial burden of proof very high—at the extreme, if they always enforced a rule that the plaintiff must prove the true value of the loss with complete certainty or receive zero—then we will never get to the stage where a defendant produces counter-evidence (because the plaintiff will either irrefutably establish the true amount and thus render counter-evidence moot, or the award will be zero by default). The fact that real world litigation regularly proceeds beyond the plaintiffs' initial showing and features defendants producing counterevidence shows that the initial burden of proof is not in fact set at such a high level, but demonstrating this proposition is a *payoff* of the analysis (one that enables subsequent payoffs, such as asking and answering *why* courts do not set the burden that high). My failure to discuss defendants is not an oversight; it is because defendants are not yet in play. In Part IV.B, I will expand my analysis to incorporate the defendant, but one cannot get to that question without first analyzing the initial burden of proof and the judicial dilemma in setting it.

D. An Example from Patent Law

This Section will illustrate how the dilemma operates by examining a pair of cases on patent damages.[53] The dilemma is by no means limited to patent cases—Part II will provide examples from other areas—but I begin with these two cases because they provide a vivid illustration of how courts struggle with the dilemma without realizing it.

---

53. *See, e.g.*, Apple Inc. v. Motorola, Inc., 757 F.3d 1286 (Fed. Cir. 2014), *overruled on other grounds by* Williamson v. Citrix Online, LLC, 792 F.3d 1339 (Fed. Cir. 2015); Lucent Techs., Inc. v. Gateway, Inc., 580 F.3d 1301 (Fed. Cir. 2009).

As background, patent damages are commonly measured according to the so-called "reasonable royalty," where courts attempt to reconstruct what license royalty the plaintiff and defendant would have agreed on, if they had negotiated ahead of time.[54] This inquiry is difficult because it involves a counterfactual.[55] The most important evidence for the reconstruction are so-called "comparable licenses"—licenses that the plaintiff or defendant had entered into on similar patents covering similar technology—on the theory that a hypothetical license between the parties would resemble similar licenses that they had actually agreed to elsewhere.[56] It turns out, however, that instead of submitting licenses that are actually comparable to the patent-in-suit and which would provide a plausible basis to make inferences about the true royalty, both plaintiffs and defendants often choose licenses that are tactically advantageous but analytically useless.[57] That is, plaintiffs would submit licenses that bore little resemblance to the actual transaction at issue, but which had high royalty rates, while defendants would submit equally dissimilar licenses that had low royalty rates.[58] The result of juries being left with such loose evidence and contrasting numbers was tremendous unpredictability in damages awards.[59] A

54. 7 DONALD S. CHISUM, CHISUM ON PATENTS § 20.07, at 20-1188 to -1190 (Matthew Bender ed., 2016) ("A reasonable royalty ... is judicially defined as the amount that would have been set in a hypothetical negotiation between a willing patent owner and a willing potential user as of the date when the infringement began."). The underlying theory for the hypothetical license inquiry being that, had the defendant taken a license, he would have thereby avoided infringement. *See* Seymour v. McCormick, 57 U.S. (16 How.) 480, 488-90 (1853).

55. *See* Panduit Corp. v. Stahlin Bros. Fibre Works, 575 F.2d 1152, 1159 (6th Cir. 1978) (discussing the counterfactual nature of the inquiry).

56. *See* CHISUM, *supra* note 54, § 20.07, at 20-1190. ("In determining this supposititious license rate, the courts consider a variety of factors or categories of evidence. The most influential factor is that of prior and existing licenses negotiated under the patent in suit.").

57. *See* William Choi & Roy Weinstein, *An Analytical Solution to Reasonable Royalty Rate Calculations*, 41 IDEA 49, 51 (2001) (describing parties who submit damages figures according to "whichever seems to benefit them most"); *see also* Jonathan S. Masur, *The Use and Misuse of Patent Licenses*, 110 NW. U. L. REV. 115, 156 (2015) (arguing that "existing licenses cannot provide a useful guide to the value of a patent").

58. *See* Choi & Weinstein, *supra* note 57, at 59.

59. *See* Daralyn J. Durie & Mark A. Lemley, *A Structured Approach to Calculating Reasonable Royalties*, 14 LEWIS & CLARK L. REV. 627, 631-32 (2010) ("The breadth of the available factors also means that it is difficult to exclude evidence or expert testimony espousing virtually any theory of reasonable royalty damages, no matter how outlandish."); Roy J. Epstein & Alan J. Marcus, *Economic Analysis of the Reasonable Royalty: Simplification*

string of large verdicts in the first decade of this century made this problem politically salient and led to much criticism of the Federal Circuit,[60] the court with exclusive jurisdiction over patent appeals.[61]

In response to these criticisms, in 2009 the Federal Circuit decided *Lucent Technologies, Inc. v. Gateway, Inc.*[62] In simple terms, *Lucent* raised the burden of proof on plaintiffs to require more rigor in damages evidence,[63] specifically by requiring the "comparable" licenses submitted to be really comparable.[64] In stark contrast to prior decisions that had cursorily allowed virtually anything to serve as the basis for affirming a jury award,[65] the *Lucent* court undertook a detailed examination (spanning eight pages of the Federal Reporter) of each of the eight licenses that the plaintiff had relied on, and explained why each license was not truly comparable to the transaction at issue and therefore could not justify the $358 million jury verdict.[66] The court then vacated the award and remanded for a new determination.[67]

The *Lucent* court did not specify what the district court should do on remand if the plaintiff could not produce better licenses that met

---

and Extension of the Georgia-Pacific *Factors*, 85 J. PAT. & TRADEMARK OFF. SOC'Y 555, 555 (2003) ("The overriding problem is that the [*Georgia-Pacific*] factors typically can be used to justify a very wide range of outcomes.").

60. *See, e.g.*, Golden, *supra* note 4, at 507 ("Eye-catching nine-figure damage awards stoke calls for changes to the law of patent damages." (footnote omitted)); Christopher B. Seaman, *Reconsidering the* Georgia-Pacific *Standard for Reasonable Royalty Patent Damages*, 2010 BYU L. REV. 1661, 1663-65 (providing a list of damages awards over $1 million as evidence of a problem).

61. 28 U.S.C. § 1295(a)(1) (2012).

62. 580 F.3d 1301 (Fed. Cir. 2009).

63. *See* ROBERT PATRICK MERGES & JOHN FITZGERALD DUFFY, PATENT LAW AND POLICY 931 (6th ed. 2013) (stating *Lucent* "signifies a tightening of attitudes regarding the admissibility and persuasiveness of evidence pertaining to 'comparable' licensing agreement[s]").

64. *See Lucent*, 580 F.3d at 1325-26 ("For similar license agreements to be used as a proxy for derivation of a fair market royalty, the form of license compensation should be on a like-kind basis." (quoting RUSSELL L. PARR, ROYALTY RATES FOR LICENSING INTELLECTUAL PROPERTY 64 (2007))).

65. *See, e.g.*, Micro Chem., Inc. v. Lextron, Inc., 317 F.3d 1387, 1392 (Fed. Cir. 2003) ("[I]t is not the role of the trial court to evaluate the correctness of facts underlying one expert's testimony."); Hunt Bros. Fruit-Packing Co. v. Cassiday, 64 F. 585, 587 (9th Cir. 1894) ("[T]he only measure of damages was such sum as, under all circumstances, would have been a reasonable royalty for the defendant to have paid. This amount it was the province of the jury to determine.").

66. *Lucent*, 580 F.3d at 1325-32.

67. *Id.* at 1340.

the new standards of true comparability.[68] The issue came to a head, however, in *Apple v. Motorola*, where Judge Richard Posner, sitting as a district judge by designation, excluded all of the plaintiff's damages evidence because it was insufficiently rigorous.[69] Judge Posner then awarded summary judgment of zero damages.[70] On appeal, the Federal Circuit reversed, holding that a district court may not award zero damages for lack of evidence because the patent statute states that a plaintiff is entitled to "no ... less than a reasonable royalty."[71] Rather, a judgment of zero damages is permissible only if the evidence affirmatively proves zero is the true value on the merits.[72] To the extent the record does not have enough evidence to support any particular award, a court or jury must simply take its best guess.[73]

What the analysis of this Article contributes to the discussion of this pair of cases is two things. First, it provides a conceptual tool to more clearly articulate the underlying rationale in each case, allowing a reader to see past irrelevant details and misleading rhetoric in the formal opinions. *Lucent* is not just about the individual fact-specific defects of the licenses at issue; the case is more generally about raising the burden of proof in order to incentivize better damages evidence.[74] *Apple* is not just about a legalistic pars-

68. *See id.* at 1332, 1340.

69. Apple Inc. v. Motorola, Inc., No. 1:11-cv-08540, 2012 WL 1959560, at *6, *9-12 (N.D. Ill. May 22, 2012), *rev'd*, 757 F.3d 1286 (Fed. Cir. 2014).

70. Apple Inc. v. Motorola, Inc., 869 F. Supp. 2d 901, 904-13, 924 (N.D. Ill. 2012), *rev'd*, 757 F.3d 1286 (Fed. Cir. 2014).

71. 35 U.S.C. § 284 (2012); Apple Inc. v. Motorola, Inc., 757 F.3d 1286, 1327 (Fed. Cir. 2014) ("'[T]he jury's finding of no damages cannot be supported' because 'the statute requires that damages to a successful claimant in a patent infringement suit shall not be less than a reasonable royalty.'" (quoting Norian Corp. v. Stryker Corp., 363 F.3d 1321, 1333 (Fed. Cir. 2004))), *overruled on other grounds by* Williamson v. Citrix Online, LLC, 792 F.3d 1339 (Fed. Cir. 2015).

72. *Apple*, 757 F.3d at 1328 ("[A] fact finder may award no damages only when the record supports a zero royalty award.").

73. *Id.* at 1327-28 ("If a patentee's evidence fails to support its specific royalty estimate, the fact finder is still required to determine what royalty is supported by the record.").

74. *See* Uniloc USA, Inc. v. Microsoft Corp., 632 F.3d 1292, 1316-17 (Fed. Cir. 2011) (invoking *Lucent* to reject a "25 percent rule of thumb" because "[t]he rule does not say anything about a particular hypothetical negotiation"); ResQNet.com, Inc. v. Lansa, Inc., 594 F.3d 860, 868-69 (Fed. Cir. 2010) (citing *Lucent* to reverse award based on "speculative and unreliable evidence divorced from proof of economic harm").

ing of the patent damages statute[75]—which does not define how much is "reasonable" and thus does not textually foreclose a zero award[76]—but is more fundamentally about avoiding harsh zero-damages awards on moral and policy grounds.[77] Viewed through these lenses, the cases are more significant and stand for deeper principles than what appears at a cursory glance.

Second, besides allowing us to better understand each case on its own, the analysis of this Article reveals how the cases irreconcilably conflict. The *Apple* court does not acknowledge—it may not even realize—that its holding is inconsistent with *Lucent*, but the two cases are in fact completely at odds with each other. *Lucent* attempts to force plaintiffs to provide better damages evidence by setting a high burden of proof.[78] *Apple* effectively guts this burden of proof by removing any penalty for noncompliance.[79] The information-forcing effect of *Lucent* is thereby undermined.

This is not to say that *Apple* was wrongly decided. There are real downsides to denying compensation to plaintiffs who have suffered losses, just as there are real benefits to having accurate and precise damages information. My goal in describing the conflict between the two cases is not to take sides on which decision is better on the merits. Rather, my point is antecedent: it is to say that there is a tradeoff, and courts should be more aware of the competing considerations at stake. What *Lucent* and *Apple* illustrate is what happens when courts lack this awareness: the dilemma does not go away;

---

75. *Cf. Apple*, 757 F.3d at 1327 ("'[T]he jury's finding of no damages cannot be supported' because 'the statute requires that damages to a successful claimant in a patent infringement suit shall not be less than a reasonable royalty.'" (quoting *Norian Corp.*, 363 F.3d at 1333)).

76. *See* Oskar Liivak, *When Nominal Is Reasonable: Damages for the Unpracticed Patent*, 56 B.C. L. REV. 1031, 1033-34 (2015) (arguing that the damages statute can be interpreted to permit nominal damages awards).

77. Understood in this way, *Apple* accords with many nineteenth-century cases, which embraced an anti-zero-award principle well before the patent statute had any language about reasonable royalties. *See infra* text accompanying notes 154-60.

78. *See* Lucent Techs., Inc. v. Gateway, Inc., 580 F.3d 1301, 1332 (Fed. Cir. 2009) ("[I]t was [the plaintiff's] burden to prove that the licenses relied on were sufficiently comparable to sustain a lump-sum damages award.").

79. *See Apple*, 757 F.3d at 1327-28 (allowing a zero royalty award only when the evidence in the record supports a zero royalty; however, a complete lack of evidence does not require this result).

instead, courts unwittingly oscillate between competing approaches, causing the law to become messy and self-contradictory.[80]

One response here is that the doctrinal oscillation might be considered harmless because courts generally reach sensible results anyway.[81] My reply is that the observation is true but the conclusion does not follow. That is, it is true that courts mostly reach sensible results, but this is because judges apply the open-ended doctrine in a way that the individual judge regards as intuitively sensible.[82] The fact that courts mostly reach sensible results through unreasoned intuition, however, is not responsive to the problem that existing doctrine does not provide coherent guidance for those results. Our legal system is not supposed to have judges determine, Solomon-like, what constitutes a sensible result based on subjective intuition; our legal system is built around the idea that judges are supposed to reach reasoned decisions based on objective rules and principles. The fact that damages doctrine is messy and provides little guidance is thus problematic even if bottom-line results are mostly sensible. And even from a purely consequentialist perspective, the fact that results depend on each judge's subjective intuition at least means that outcomes are rather unpredictable and haphazard, which is harmful because parties are generally risk-averse.[83]

II. MORE EXAMPLES OF THE DILEMMA IN ACTION

A. Contract Law

A common problem in contract law is the calculation of lost profits.[84] For example, if General Motors (GM) has a contract to buy

80. *Cf.* Karl N. Llewellyn, *Remarks on the Theory of Appellate Decision and the Rules or Canons About How Statutes Are to Be Construed*, 3 VAND. L. REV. 395, 401-06 (1950) (showing how canons and counter-canons of statutory construction conflict).

81. *Cf.* O.W. HOLMES, JR., THE COMMON LAW 1 (1881) ("The life of the law has not been logic; it has been experience.").

82. *See id.* ("The felt necessities of the time, the prevalent moral and political theories, intuitions of public policy, avowed or unconscious, even the prejudices which judges share with their fellow-men, have had a good deal more to do than the syllogism in determining the rules by which men should be governed.").

83. *See supra* text accompanying note 51.

84. *See* Mark Glick & Cory Sinclair, *Damages Resulting from a Lost Opportunity: The Proper Damage Date in Utah Contract and Tort Cases*, 23 UTAH B.J. 30, 30 (2010) ("The issue

engine parts from Delphi and Delphi fails to supply the parts, then GM loses not only the engine parts, but also sales of cars that it could have made with the parts. Contract law then allows GM to recover the lost profits on car sales, provided that the amount of these lost profits can be proven.[85]

For a company like GM, with a well-established history of sales and detailed financial records, it is generally not too difficult to prove the lost profits caused by a supply disruption to a high degree of precision.[86] GM usually has a good idea of how many cars it will sell in a particular quarter based on historical sales patterns and studies of consumer behavior, and it has detailed financial records on the amount of profit it makes per car. If a supply disruption causes fewer cars to be manufactured and sold, the effect can thus be readily quantified. In these circumstances, proving the lost profits from a breach of contract poses little difficulty.

What courts quickly faced, however, were cases in which breaches of contract led to supply disruptions for startup companies without a record of prior sales. For example, in *Evergreen Amusement Corp. v. Milstead*, a contractor's delay forced a movie theatre to postpone its opening by three months, and the question was whether the movie theatre could recover its lost profits from the postponement.[87] Historically, many courts imposed a demanding burden of proof, requiring plaintiffs to prove the amount of lost profits with "certainty."[88] Strict enforcement of this burden of proof then led to the result that startup businesses could not recover lost profit damages for breaches of contract because a startup business would never have a sufficient record of prior sales to establish the amount of lost profits with certainty.[89]

---

of how to calculate the damages from a lost opportunity often arises in contract ... cases.").

85. *See* U.C.C. §§ 2-713, -715 (AM. LAW INST. & NAT'L CONFERENCE OF COMM'RS OF UNIF. STATE LAWS 2014) (allowing recovery of consequential damages).

86. *See* RESTATEMENT (SECOND) OF CONTRACTS § 347 cmt. b (AM. LAW INST. 1981) ("Where the injured party's expected advantage consists largely or exclusively of the realization of profit, it may be possible to express this loss in value in terms of money with some assurance.").

87. 112 A.2d 901, 902 (Md. 1955).

88. *See, e.g.*, Griffen v. Colver, 16 N.Y. 489, 491 (1858) ("It is a well established rule of the common law that the damages to be recovered for a breach of contract must be shown with certainty, and not left to speculation or conjecture."); HALE, *supra* note 33, § 31.

89. *See, e.g.*, *Evergreen*, 112 A.2d at 904 ("[D]amages for profits anticipated from a business which has not started may not be recovered."); *see also* Cent. Coal & Coke Co. v.

This so-called "new business rule" attracted much criticism.[90] In response, courts lowered the burden of proof. Instead of requiring "certainty," courts relaxed the requirement to one of "reasonable certainty,"[91] and they gradually moved away from the new business rule.[92] The "reasonable certainty" formulation is now used by every American jurisdiction,[93] and the rule that startup businesses can never recover lost profits is regarded as defunct.[94]

The transition from a "certainty" standard to a "reasonable certainty" standard and the attendant demise of the new business

---

Hartman, 111 F. 96, 98 (8th Cir. 1901); The Lively, 15 F. Cas. 631, 634-35 (C.C.D. Mass. 1812) (No. 8403); Cramer v. Grand Rapids Show Case Co., 119 N.E. 227, 228-29 (N.Y. 1918); Sinclair Ref. Co. v. Hamilton & Dotson, 178 S.E. 777, 780 (Va. 1935).

90. *See, e.g.*, Vickers v. Wichita State Univ., 518 P.2d 512, 517 (Kan. 1974) ("To hold recovery is precluded as a matter of law merely because a business is newly established would encourage those contracting with such a business to breach their contracts."); Bernadette J. Bollas, Note, *The New Business Rule and the Denial of Lost Profits*, 48 OHIO ST. L.J. 855, 855 (1987); Frank Lane Williamson, Comment, *Remedies—Lost Profits as Contract Damages for an Unestablished Business: The New Business Rule Becomes Outdated*, 56 N.C. L. REV. 693, 693-96 (1978); Note, *The Requirement of Certainty in the Proof of Lost Profits*, 64 HARV. L. REV. 317, 318-19 (1950).

91. *See* TAS Distrib. Co. v. Cummins Engine Co., 491 F.3d 625, 631-32 (7th Cir. 2007) ("The absolute certainty requirement was reformed some time ago to the less demanding requirement of 'reasonable certainty.'"); M & R Contractors & Builders, Inc. v. Michael, 138 A.2d 350, 355 (Md. 1958) ("Courts have modified the 'certainty' rule into a more flexible one of 'reasonable certainty.'"); HARVEY MCGREGOR, MAYNE AND MCGREGOR ON DAMAGES § 174, at 163 (12th ed. 1961).

92. *See* Bollas, *supra* note 90, at 859 ("The clear and growing majority of courts, however, now apply the new business rule as a rule which delimits the sufficiency of evidence [rather than as a categorical bar]."); *see also M & R Contractors*, 138 A.2d at 355 ("[T]he last hundred years have witnessed continual modification of the once rigid rule that anticipated profits, because inherently uncertain, were *per se* not a proper element of damages for breach of contract." (quoting Note, *Speculative Profits as Damages for Breach of Contract*, 46 HARV. L. REV. 696, 696 (1933))).

93. *TAS Distrib.*, 491 F.3d at 632 ("Today, all United States jurisdictions enforce the requirement of 'certainty' in damage award amount, but limit this requirement to 'reasonable certainty.'"); RESTATEMENT (SECOND) OF CONTRACTS § 352 (AM. LAW INST. 1981) ("Damages are not recoverable for loss beyond an amount that the evidence permits to be established with reasonable certainty.").

94. *See, e.g.*, 22 AM. JUR. 2D *Damages* § 58 (2013) ("The new business rule is not a per se rule forbidding an award of lost-profits damages to new businesses but rather an evidentiary rule that creates a higher level of proof needed to achieve reasonable certainty as to the amount of damages."); ROBERT L. DUNN, RECOVERY OF DAMAGES FOR LOST PROFITS § 4.3, at 378 (6th ed. 2005) ("Most recent cases reject the once generally accepted rule that lost profits damages for a new business are not recoverable.... The earlier cases are either ignored or rationalized.").

rule have been documented elsewhere.[95] Understanding the dilemma adds to prior accounts in two ways. The first is that it provides a fuller account of the underlying reasons and concerns that led to courts first creating the "certainty" standard and then relaxing it.[96] The "certainty" standard functions as a high burden of proof.[97] Strict enforcement of this standard has an information-forcing effect,[98] which explains why some courts once thought it to be a good idea. At the same time, strict enforcement of a high burden of proof has the downside that it generates harsh outcomes—requiring startup businesses to prove lost profits with certainty effectively means they cannot recover lost profits from breaches of contract, which is both intuitively unfair and also discourages business formation.[99] These downsides eventually led to courts relaxing the rule and undermining its information-forcing effect.

The second payoff is that understanding the dilemma gives more meaning to the "reasonable certainty" standard. It should be obvious that the "reasonable certainty" formulation itself is very open-ended; no one knows what it means, and the case law applying the standard is not consistent.[100] What the information-forcing dilemma explains is *why* courts are not consistent, and it fleshes out the competing considerations at play. To the extent that courts apply the standard very stringently, so that "reasonable certainty" essentially requires absolute certainty, then courts will create long-term incentives for better information disclosure. But they will also

95. *See TAS Distrib.*, 491 F.3d at 631-32; DUNN, *supra* note 94, §§ 4.1, -.3, -.4, at 374-92; John M. Golden, *Reasonable Certainty in Contract and Patent Damages*, 30 HARV. J.L. & TECH. (SPECIAL SYMPOSIUM) 257, 266-72 (2017).

96. *See* Comment, *Lost Profits as Contract Damages: Problems of Proof and Limitations on Recovery*, 65 YALE L.J. 992, 1000 & n.43 (1956) (observing that many courts "have broken from precedent" without "stating the reasons for their departure" and "in fact often seem unaware of, or ignore, their departures from precedent").

97. E. Allan Farnsworth, *Legal Remedies for Breach of Contract*, 70 COLUM. L. REV. 1145, 1210-11 (1970) (characterizing the standard as a raised burden of proof).

98. Despite the extensive attention paid to information-forcing rules in the contracts literature, I can find no one who has made this point about the certainty standard.

99. *See* Vickers v. Wichita State Univ., 518 P.2d 512, 517 (Kan. 1974) ("Strict application of the certainty doctrine would place a new business at a substantial disadvantage.").

100. *See* DUNN, *supra* note 94, § 1.7, at 21 ("Relatively few cases analyze what is meant by reasonable certainty."); Robert M. Lloyd, *The Reasonable Certainty Requirement in Lost Profits Litigation: What It Really Means*, 12 TRANSACTIONS 11, 12 (2010) ("In spite of this universal adoption of the language, however, courts have never really explained what they mean by the term 'reasonable certainty.'").

create harsh outcomes that undercompensate deserving victims, such as denying recovery to startup businesses who have clearly suffered injury from a defaulting contractor. To the extent that courts apply the standard very loosely, so that "reasonable certainty" requires no certainty at all, then courts can avoid those harsh and unjust outcomes, but at the price of undermining incentives for rigorous damages evidence and creating greater uncertainty in damages awards. Courts that seek to balance these competing considerations naturally oscillate between the two approaches.

In this sense, the universal adoption of the "reasonable certainty" formulation does not indicate consensus or consistency in the law but instead merely obscures a debate that was once more transparent. Instead of different courts explicitly adopting or rejecting a rule that startup businesses cannot recover lost profits, courts today all purport to agree on a "reasonable certainty" standard.[101] But they apply the standard with different degrees of stringency,[102] so there is really no agreement at all.[103] Understanding the dilemma provides a conceptual tool to see that the same debate is still going on under the surface and to bring out the competing considerations to explain what judges are disagreeing about.

B. Tort Law

An obvious context where the dilemma arises in tort law is the issue of pain and suffering damages. The question for pain and suffering damages is not whether there is *a* loss—even the most diehard skeptic concedes that pain and suffering are real losses[104]—but how to *quantify* the loss.[105] Pain and suffering are inherently difficult to

101. *See supra* notes 98-100 and accompanying text.

102. *See supra* note 100 and accompanying text.

103. *See* Cass R. Sunstein, Commentary, *Incompletely Theorized Agreements*, 108 HARV. L. REV. 1733, 1739 (1995) (noting that different sides can appear to agree on a verbal formulation while glossing over substantive disagreements).

104. *E.g.*, Paul V. Niemeyer, *Awards for Pain and Suffering: The Irrational Centerpiece of Our Tort System*, 90 VA. L. REV. 1401, 1402 (2004) ("Pain is real, and, of course, the suffering it causes is real."); *see also* Kwasny v. United States, 823 F.2d 194, 197 (7th Cir. 1987) (Posner, J.) ("We disagree with those students of tort law who believe that pain and suffering are not real costs .... No one likes pain and suffering and most people would pay a good deal of money to be free from them.").

105. *See* Niemeyer, *supra* note 104, at 1402 (diagnosing as "damages that are irrationally

value, because no readily accessible market exists by which physical pain and suffering can be objectively translated into a monetary amount.[106] The result is that courts face a conundrum when it comes to devising doctrinal rules for pain and suffering damages: if courts strictly enforce a very stringent burden of proof, requiring plaintiffs to objectively establish the value of their pain through rigorous evidence, the result will be many zero-damages awards that deny compensation to victims of real pain.[107] But if courts allow a very lax burden of proof, the result will be speculative awards that invite political backlash,[108] such as legislative proposals to cap awards.[109]

Although pain and suffering damages in one sense provide a good example of the dilemma, I will not use them as my main example, because any discussion about pain and suffering damages requires entering into a predicate debate about whether pain can be valued in money in the first place. This "incommensurability" argument is

---

quantified").

106. *See* Flory v. N.Y. Cent. R.R. Co., 163 N.E.2d 902, 905 (Ohio 1959) (per curiam) ("Of all the items of compensatory damages ... human pain and suffering is perhaps the most difficult to determine."). An important assumption of my analysis here is that it is *theoretically possible* to determine a "correct" value for pain and suffering damages, even if doing so is practically very difficult. This assumption is often contested—the counterargument being that pain and suffering is inherently susceptible of no monetary quantification. *See* Margaret Jane Radin, *Compensation and Commensurability*, 43 DUKE L.J. 56, 75 (1993) (discussing the general problem of incommensurability); *see also* Cass R. Sunstein, *Incommensurability and Valuation in Law*, 92 MICH. L. REV. 779 (1994) (same). To the extent one takes such a position, however, it then becomes impossible to intelligently discuss pain and suffering damages—no monetary amount, from zero to infinity, is either "right" or "wrong" if pain and suffering simply cannot be measured in monetary terms. No court truly subscribes to the most extreme version of the incommensurability argument, since every court reviews pain and suffering damages for excessiveness, and an excessiveness standard logically presupposes that there is *some* benchmark against which an award is measured, so that an award can be so "wrong" as to be excessive. *See, e.g.*, Linn v. United Plant Guard Workers of Am., Local 114, 383 U.S. 53, 65-66 (1965) ("If the amount of damages awarded is excessive, it is the duty of the trial judge to require a remitter or a new trial.").

107. *Cf. supra* notes 35-39 and accompanying text.

108. *See* Seffert v. L.A. Transit Lines, 364 P.2d 337, 346-47 (Cal. 1961) (Traynor, J., dissenting) (criticizing pain and suffering award because "no evidence was introduced" on the value of the suffering); Plant, *supra* note 16, at 203 (arguing that the requirements are so lax that "unless the instruction baldly suggests to the jury that they engage in free speculation, it is likely to be upheld"); William Zelermyer, *Damages for Pain and Suffering*, 6 SYRACUSE L. REV. 27, 28 (1954) (arguing that pain and suffering damages are "conceived at the end of a speculative and uncertain journey").

109. *See* AM. TORT REFORM ASS'N, *supra* note 16, at 34-42 (listing laws of twenty-nine states capping nonpecuniary damages).

a never-ending philosophical debate beyond the scope of this Article to resolve.[110] For now, another good illustration of the dilemma from the field of tort law is the situation in which a wrongful act does not cause an entire injury by itself, but instead *aggravates* some preexisting injury—for example, when a plaintiff who was already suffering a back injury gets worse due to an accident caused by defendant's negligence.[111] Under standard principles of negligence law, one might reason that the plaintiff should only be able to recover damages *to the extent* of the aggravation.[112] That is, if the original back injury was worth $100, while the postaccident back injury is worth $150, then the defendant *caused* only $50 of harm and should only be liable for that amount. And indeed, many cases stand for this proposition.[113] Moreover, several cases explicitly hold that the plaintiff has the burden of proving the amount of the aggravation.[114]

As an initial matter, it is important to appreciate that this rule makes a certain amount of sense. If the goal of tort law is to compensate plaintiffs only for losses caused by a defendant's wrongdoing,[115] then courts need to figure out how much of a plaintiff's injury is attributable to the defendant's wrongdoing. And the only way they can do that is by using an information-forcing mechanism to obtain the necessary information, or, in other words, by imposing a

110. *See generally* Radin, *supra* note 106; Sunstein, *supra* note 106. I will add that, to the extent that one truly accepts the incommensurability argument, it would imply that there is no such thing as an excessive pain and suffering award, whereas every court reviews pain and suffering damages for excessiveness.

111. *See* St. Louis Sw. Ry. Co. of Tex. v. Johnson, 97 S.W. 1039, 1040 (Tex. 1906).

112. *See id.* ("[T]he plaintiff would be entitled to recover, if at all, ... only for the increase or aggravation of the troubles.").

113. *See, e.g.*, Louisville & Nashville R.R. Co. v. Mattingly, 318 S.W.2d 844, 849 (Ky. 1958) ("L. & N. cannot be held responsible to [the plaintiff] Mattingly for his prior injury and diseased condition except to the extent that such disablement was aggravated or extended by the one involved here."); *St. Louis Sw. Ry. Co.*, 97 S.W. at 1040.

114. *See, e.g.*, Rowe v. Munye, 702 N.W.2d 729, 740 (Minn. 2005) ("[T]o require the defendant to separate the new injury from the preexisting injury improperly places the burden on the party with the lesser amount of information and again might have the tendency to overcompensate the plaintiff."); Gulf, Colo. & Santa Fe Ry. Co. v. Mannewitz, 8 S.W. 66, 67 (Tex. 1888) ("The burden of proof is on the plaintiff, Mannewitz, to show the extent of his injuries which were caused by the turning over of the car.").

115. *See* John C.P. Goldberg & Benjamin C. Zipursky, *Torts as Wrongs*, 88 TEX. L. REV. 917, 918 (2010) (arguing that "tort law is about *wrongs*").

burden of proof.[116] A rule requiring the plaintiff to prove the amount of the aggravation quite logically follows from these principles.

However, apportioning the value of an injury between two causes is generally fraught with difficulty. It is already difficult, if not impossible, to place a monetary value on physical injuries;[117] it is even more difficult to apportion this monetary value between the *preexisting old* injury and the *incremental new* injury.[118] Requiring plaintiffs to prove the extent of the aggravation—if one takes the requirement seriously and demands anything even remotely resembling rigorous precision—then becomes a de facto rule holding that plaintiffs cannot recover for aggravations of preexisting injuries.

It should therefore come as little surprise that there are lines of countercases that either relax the burden of proof and allow juries to take quite speculative guesses at the correct apportionment,[119] or eliminate the apportionment question altogether by holding that the plaintiff can recover the entire amount of the injury.[120] But both of these options have significant drawbacks: relaxing the burden of proof leads to speculative and uncertain damages awards, while allowing plaintiffs to recover the entire amount of the injury leads to excessive awards that are unfair to the defendant (and induce potential victims to take too few precautions and potential injurers to take too many).[121] Because every possible rule has its own downside, courts oscillate and the doctrine becomes incoherent and self-contradictory.

---

116. *See supra* text accompanying notes 23-31.

117. *See supra* note 106.

118. LaMoureaux v. Totem Ocean Trailer Exp., Inc., 632 P.2d 539, 544 (Alaska 1981) ("[I]t is often difficult to determine how much of a plaintiff's injury is due to the preexisting condition and how much to the aggravation caused by the defendant.").

119. *Id.* ("The requirements of proof usually have been somewhat relaxed in such cases." (quoting WILLIAM L. PROSSER, THE LAW OF TORTS 319 (4th ed. 1978))).

120. *E.g.*, Lovely v. Allstate Ins. Co., 658 A.2d 1091, 1093-94 (Me. 1995); Tingey v. Christensen, 987 P.2d 588, 592 (Utah 1999) (holding that the plaintiff can recover for the entire injury because "a defendant should not escape liability because the amount of damage cannot be proved with precision" (quoting Winsness v. M.J. Conoco Distribs., Inc., 593 P.2d 1303, 1306 (Utah 1979))).

121. *See* Rowe v. Munye, 702 N.W.2d 729, 740 (Minn. 2005) (declining to shift the burden of proof to the defendant because this "could force the defendant to pay for damages he did not cause"); *see also infra* Part III.B.

C. Antitrust Law

Of all areas, antitrust law has adopted perhaps the most consistent approach to the dilemma. Ever since the Supreme Court decision in *Story Parchment Co. v. Paterson Parchment Paper Co.* held that "[w]here the tort itself is of such a nature as to preclude the ascertainment of the amount of damages with certainty, it would be a perversion of fundamental principles of justice to deny all relief to the injured person,"[122] courts have emphasized that a low standard of proof applies to antitrust damages.[123] While damages still may not be quantified "by mere speculation or guess," a jury is given broad leeway to fix the amount of damages "as a matter of just and reasonable inference, although the result be only approximate."[124]

Two points are worth making here. The first is that this represents making a choice *within* the dilemma, not an exception to it. Antitrust courts accept loose evidence and uncertain awards as the price for avoiding harsh zero-damages outcomes.[125] Consciously picking one poison over another has allowed antitrust law to (mostly) avoid the doctrinal oscillation that is seen in other areas, but my predicate point that courts face a choice between two poisons still stands.

The second point is that, even in the antitrust context, the underlying policy tension inherent in the dilemma ultimately reasserts itself, and various doctrinal manifestations of this tension can still be seen if one looks carefully. Most obviously, antitrust law still holds that awards may not be speculative.[126] This means that *at some point* courts must draw the line and hold that a plaintiff

122. 282 U.S. 555, 563 (1931).

123. *See* Bigelow v. RKO Radio Pictures, Inc., 327 U.S. 251, 263-65 (1946); *see also* Zenith Radio Corp. v. Hazeltine Research, Inc., 395 U.S. 100, 123 (1969) ("Trial and appellate courts alike must also observe the practical limits of the burden of proof which may be demanded.").

124. *Story Parchment*, 282 U.S. at 563.

125. *See supra* notes 122-24 and accompanying text.

126. *See* J. Truett Payne Co. v. Chrysler Motors Corp., 451 U.S. 557, 567 (1981) ("[I]t is a close question whether petitioner's evidence would be sufficient to support a jury award even under our relaxed damages rules."); 2A PHILLIP E. AREEDA ET AL., ANTITRUST LAW: AN ANALYSIS OF ANTITRUST PRINCIPLES AND THEIR APPLICATION § 392a, at 378 (4th ed. 2014) ("'[S]peculative' is an epithet that is used to characterize insufficient damage proof."); Roger D. Blair & William H. Page, *"Speculative" Antitrust Damages*, 70 WASH. L. REV. 423, 426 (1995) (conceptualizing the "speculative" standard as an ex post label that courts attach to a conclusion that the plaintiff failed the burden of proof).

who has suffered a loss has nevertheless failed to prove the amount of that loss with sufficient precision and therefore receives zero damages.[127] Antitrust law's acceptance of loose evidence and aversion to harsh zero-damages awards is not absolute, despite the strong language ("perversion of fundamental principles of justice")[128] that might suggest otherwise. Courts are always balancing the desire to do justice in an individual case against the need to force litigants to produce reliable information.[129] The fact that antitrust courts somewhat tip the scales in favor of doing justice in individual cases does not obviate the fact that there is still *a* balance, nor does it remove the difficulty of the balancing act.[130]

Less obviously, another doctrinal manifestation is that lower courts sometimes surreptitiously move the damages quantification inquiry into other doctrinal silos in order to avoid the *Story Parchment* precedent and impose a higher standard of proof. For example, a court might reclassify what is really an inquiry into the *amount* of damages as an inquiry about *causation* of antitrust injury. A good example of this move is *MCI Communications Corp. v. American Telephone & Telegraph Co.*, a Seventh Circuit case in which the jury found $600 million in compensatory damages (that 15 U.S.C. § 15(a) automatically trebled to $1.8 billion) for alleged anticompetitive practices by AT&T in excluding MCI from the telecommunications market.[131] The $600 million number was calculated by comparing the difference between MCI's original business plans— showing what a supposedly "undamaged" MCI would have earned if AT&T had allowed MCI completely free rein—and MCI's actual operating numbers.[132]

---

127. *See* ABA SECTION ON ANTITRUST LAW, *supra* note 19, at 56 ("Setting a low threshold is one thing; determining whether the threshold has been reached is another.").

128. *See Story Parchment*, 282 U.S. at 563.

129. *See* ABA SECTION OF ANTITRUST LAW, *supra* note 19, at 56 ("Proving as a fact something that never occurred ('what the plaintiff's situation would have been in the absence of the defendant's antitrust violation') is impossible. Resort to assumptions and inferences, supported by real-world data, evidence, and economic theory, is inevitable.").

130. *See id.* ("Many of the battles over antitrust damages revolve around whether the plaintiff has provided sufficient support for its quantification."); AREEDA ET AL., *supra* note 126, § 392a, at 378 ("[I]t is difficult to identify a bright line that separates a 'just and reasonable inference' from 'speculation.'").

131. 708 F.2d 1081, 1092-93 (7th Cir. 1983).

132. *Id.* at 1160.

In reversing this verdict, the Seventh Circuit began by acknowledging that "the Supreme Court has been willing to accept a degree of uncertainty in the calculation of damages" and "the *amount* of damages may be determined by a just and reasonable estimate as long as the jury verdict is not the product of speculation or guess work."[133] After this initial perfunctory acknowledgement, it went on to say,

> It is essential, however, that damages reflect only the losses directly attributable to *unlawful* competition....
>
> ...
>
> The courts have always distinguished between proof of *causation* of damages and proof of the *amount* of damages. Thus, the courts have been consistent in requiring plaintiffs to prove in a reasonable manner the link between the injury suffered and the *illegal* practices of the defendant.[134]

The court held that MCI failed to prove how much of the $600 million in lost profits were specifically attributable to AT&T's unlawful activities, as opposed to lawful competition, and that this was a failure of proof of causation as opposed to one pertaining to the amount of damages.[135] It then vacated the award and remanded for a new trial.[136]

It is important to see that, although the *MCI* court characterized its holding as a matter of proving a causal "link between the injury suffered and the *illegal* practices of the defendant,"[137] its decision is not about causation in the conventional sense, which asks whether the defendant's actions caused an injury that supports finding liability.[138] There is no doubt that AT&T's unlawful practices caused MCI *some* antitrust injury, or that there is antitrust liability.[139] The

133. *Id.* at 1161.

134. *Id.*

135. *See id.* at 1162-64.

136. *Id.* at 1168, 1174.

137. *Id.* at 1161.

138. *See* RESTATEMENT (SECOND) OF TORTS § 9 (AM. LAW INST. 1965) (defining legal cause as "the fact that the causal sequence by which the actor's tortious conduct has resulted in an invasion of some legally protected interest of another is such that the law *holds the actor responsible* for such harm unless there is some defense to *liability*" (emphasis added)).

139. *MCI*, 708 F.2d at 1166-67 ("A new trial on liability is unwarranted since this court has affirmed liability for monopolization on the basis of most of AT&T's actions involving intercon-

only question is the amount of damages.[140] What the *MCI* decision effectively requires is for a plaintiff to provide detailed and rigorous evidence regarding the amount of damages attributable to unlawful acts.[141] This is also precisely what *Story Parchment* holds is *not* required.[142] Logically, the two cases are irreconcilable; recharacterizing the issue as one of "causation" rather than as about the "amount of damages" merely obscures the conflict.[143]

The contradiction between *MCI* and *Story Parchment* has not gone unnoticed.[144] Nor is *MCI* the only case to use some fancy footwork to surreptitiously impose a heightened burden of proof for antitrust damages.[145] My point is not to criticize the Seventh Circuit on legal hierarchy grounds for defying Supreme Court precedent;[146]

---

nection with MCI.").

140. *Id.* at 1168 ("The issue to be re-tried on remand is solely the amount of damages that MCI is entitled to receive for those acts by AT&T which have been found to be unlawful."); *cf.* Story Parchment Co. v. Paterson Parchment Paper Co., 282 U.S. 555, 567 (1931) ("That there was actual damage due to depreciation in value was not a matter of speculation, but a fact which could not be gainsaid. The amount alone was in doubt.").

141. *See MCI*, 708 F.2d at 1168 ("MCI must be able to rationally separate these damages from those losses (or reductions in profit) which are caused by the purely lawful competitive actions of AT&T.").

142. *See Story Parchment*, 282 U.S. at 566-67 ("Whether the unlawful acts of respondents or conditions apart from them constituted the proximate cause of the depreciation in value, was a question, upon the evidence in this record, for the jury."); *see also* Nat'l Farmers' Org., Inc. v. Associated Milk Producers, Inc., 850 F.2d 1286, 1306-07 (8th Cir. 1988).

143. To be clear, I am not saying that the "causation" and "amount of damages" inquiries are conceptually indistinguishable when properly conceived. I am saying that *MCI* distorts the concept of causation.

144. *See* James R. McCall, *The Disaggregation of Damages Requirement in Private Monopolization Actions*, 62 NOTRE DAME L. REV. 643, 670 (1987) ("The proof standard mandated by the disaggregation requirement ... has been rejected in antitrust treble damage actions under the decisions of the Supreme Court."); M. Sean Royall, *Disaggregation of Antitrust Damages*, 65 ANTITRUST L.J. 311, 311 (1997) ("Within the sea of case law upholding broadly permissive standards for proof of antitrust damages, there is a competing undercurrent known as the disaggregation rule.").

145. *See, e.g.*, Concord Boat Corp. v. Brunswick Corp., 207 F.3d 1039, 1057 (8th Cir. 2000) (excluding plaintiff's damages theory "because it did not separate lawful from unlawful conduct"); Image Tech. Servs., Inc. v. Eastman Kodak Co., 125 F.3d 1195, 1224 (9th Cir. 1997) ("The ISOs must segregate damages attributable to lawful competition from damages attributable to Kodak's monopolizing conduct.").

146. Indeed, the *MCI* court's reasoning quite plausibly follows from another line of Supreme Court decisions starting with *Brunswick Corp. v. Pueblo Bowl-O-Mat, Inc.*, 429 U.S. 477, 489 (1977), which held that an antitrust violation that causes competitive injury to a plaintiff might not support any damages if the technically unlawful conduct ultimately benefits consumers. *Brunswick*'s rationale can be extended to implicitly require plaintiffs to prove the amount of consumer injury.

rather, it is to illustrate how a deep-seated policy tension ultimately reasserts itself notwithstanding a Supreme Court decision. Doctrinal technicalities about the difference between causation and amount of damages are mere excuses; what is really driving the conflict is a policy tension between the benefits of a rigorous burden of proof in forcing information, and the desire for just outcomes and aversion to harsh zero-damages awards.[147] One cannot understand the interaction between *MCI* and *Story Parchment* without understanding the underlying policy dispute at stake, and one cannot understand the policy dispute without understanding the information-forcing dilemma.

D. Nineteenth-Century Patent Law

The dilemma is not a new problem: courts have been facing the dilemma for as long as courts have been awarding compensatory damages. The timelessness of the problem is partially demonstrated by the fact that many of the cases in my contract law discussion are quite old.[148] But to really bring this point home, it is useful to compare the modern patent cases that I discussed in Part I.D with some older cases.

As background, although the standard measure of damages in patent law today is the "reasonable royalty,"[149] early patent law did not use this formulation. Instead, patentees who wished to recover lost royalty revenue had to prove an "established royalty," that is, a royalty rate established by a long record of prior licenses.[150] As the Supreme Court held in *Rude v. Westcott*:

> In order that a royalty may be accepted as a measure of damages against an infringer, who is a stranger to the license

147. *See* McCall, *supra* note 144, at 652 ("The burden of persuasion and the standard of proof are important concepts to keep in mind when considering the disaggregation requirement."); Charles N. Charnas, Comment, *Segregation of Antitrust Damages: An Excessive Burden on Private Plaintiffs*, 72 CALIF. L. REV. 403, 423 (1984) ("A literal application of the segregation doctrine may reduce the possibility of speculative treble damage awards, but it may also raise the plaintiff's burden of building a prima facie case.").

148. *See supra* Part II.A.

149. 35 U.S.C. § 284 (2012).

150. *See* Birdsall v. Coolidge, 93 U.S. 64, 68 (1876) ("That when a person, without license, appropriates the patented invention of another, the measure of damages, if a royalty has been established, is the regular royalty paid by purchasers and licensees.").

> establishing it, it must be paid or secured before the infringement complained of; it must be paid by such a number of persons as to indicate a general acquiescence in its reasonableness by those who have occasion to use the invention; and it must be uniform at the places where the licenses are issued.[151]

Viewed from a law and economics perspective, the rationale underlying this rule should be obvious: the rule forces the production of more rigorous evidence of comparable licenses.[152] Determining the true amount of lost royalties is generally very difficult because the question ultimately comes down to what royalty the patentee and defendant would have agreed on if they had (counterfactually) negotiated ahead of time, when in fact no negotiation occurred.[153] A court can construct the counterfactual with more confidence if it has an extensive record of prior licenses with a single, uniform royalty rate—if everyone else had agreed to the same royalty rate for the same patent, then it is likely that the defendant would also have agreed to this royalty rate as well.[154] A rule that patentees seeking to recover royalty damages must produce an extensive record of prior licenses thus has an information-forcing effect.

The problem that courts quickly confronted were situations where it was not possible for a plaintiff to meet the burden of proof because the evidence demanded did not exist—either there were no (or only a few) prior licenses, or the licenses did not have a single, uniform royalty rate.[155] Under the logic of *Rude*, a court in such a situation would then have to deny the patentee all recovery for lost royalties.[156] This outcome would be quite unjust, as the patent clearly has

---

151. 130 U.S. 152, 165 (1889).

152. *See* Diamond Stone-Sawing Mach. Co. v. Brown, 166 F. 306, 309 (2d Cir. 1908) ("The rule which we sanction requires proof instead of conjecture and facts instead of theories.").

153. *See supra* text accompanying notes 54-55.

154. As Jonathan Masur has argued, a gap between the evidence and the hypothetical inquiry remains because the hypothetical inquiry is supposed to assume that the patent is fully valid and infringed, whereas real-life licenses are almost never negotiated under such conditions. Masur, *supra* note 57, at 130-31. But having robust evidence of comparable licenses is at least better than not having such evidence.

155. *See, e.g.*, Coupe v. Royer, 155 U.S. 565, 583-84 (1895); Hunt Bros. Fruit-Packing Co. v. Cassiday, 64 F. 585, 587 (9th Cir. 1894).

156. *See Rude*, 130 U.S. at 167 (ordering nominal damages); *see also Coupe*, 155 U.S. at 583 (same). A patentee could still pursue other theories, such as for lost profits from displaced

at least some value (the defendant did, after all, use the patented technology, and presumably would not have done so if it has absolutely no value) and a rational defendant would thus have been willing to pay *some* royalty for it.[157] The analytical question is not whether there has been *any* lost royalty, but *how much* the lost royalties are.[158] We thus have another example of the (now familiar) setup of a doctrinal rule logically calling for giving plaintiffs zero damages because they cannot prove exact amount of the loss, when no one thinks that zero is actually the correct amount as a first-principles matter.

What courts did in response was to develop the "reasonable royalty" doctrine, which largely replicated the established royalty analysis but with a lower burden of proof.[159] Instead of requiring patentees to rigorously prove an established royalty amount through an extensive record of prior licenses (on pain of receiving no royalty damages at all), juries would now be permitted to fix the royalty damages at an amount they found to be "reasonable."[160] As Judge Learned Hand observed, this methodology was hardly ideal as a matter of informed decision-making or precise awards, but it had a very big upside in avoiding harsh zero-damages awards: "The whole notion of a reasonable royalty is a device in aid of justice, by which that which is really incalculable shall be approximated,

---

sales. But not every patentee has displaced sales—nonpracticing patentees, for example, would not have any sales to lose. *See Rude*, 130 U.S. at 167 (stating that there was no issue of lost profits because "no machines had been manufactured and put on the market by the patentee"). Moreover, even when it is clear that there are some lost sales, the amount of lost profits is *also* very hard to prove, so the dilemma repeats itself. *See* Mark A. Lemley, *Distinguishing Lost Profits from Reasonable Royalties*, 51 WM. & MARY L. REV. 655, 656 (2009) (arguing that "courts have distorted the reasonable royalty measure in various ways" in order to allow recovery for "patent owners who have in fact probably lost sales to infringement [but] cannot prove lost profits damages" due to "strict standards of proof").

157. *See Hunt Bros.*, 64 F. at 587 ("The plaintiff was clearly entitled to damages for the infringement.").

158. *See* U.S. Frumentum Co. v. Lauhoff, 216 F. 610, 616-17 (6th Cir. 1914) ("It is a travesty to allow property rights to be seized and enjoyed without remedy simply because of the supposed difficulty in establishing their value.").

159. Dowagiac Mfg. Co. v. Minn. Moline Plow Co., 235 U.S. 641, 648 (1915) (stating that where there is no established royalty, "it was permissible to show the value by proving what would have been a reasonable royalty").

160. *Id.* at 648-49 ("This amount it was the province of the jury to determine." (quoting *Hunt Bros.*, 64 F. at 587)).

rather than that the patentee, who has suffered an indubitable wrong, shall be dismissed with empty hands."[161]

At this point we catch up to the story told in Part I.C. The lax standards that applied to reasonable royalty calculations eventually led to complaints about unpredictable awards and runaway juries.[162] This situation then prompted the Federal Circuit to attempt to tighten standards of proof in *Lucent*.[163] Although *Rude* and *Lucent* are separated by a vast distance in time and are ostensibly about two different doctrines[164]—*Rude* is about the "established royalty" doctrine while *Lucent* is about the "reasonable royalty" doctrine—the two decisions share many similarities. At a fundamental level, both opinions are trying to make the calculation of lost royalties more rigorous and precise by demanding plaintiffs produce *truly comparable* licenses.[165] Although *Lucent* does not go nearly as far as *Rude* did—*Rude* spelled out a list of onerous bright-line requirements for what plaintiffs must provide—the underlying logic of both decisions is the same.

The similarities do not end there. *Rude* was eventually undermined by later decisions and today is defunct.[166] Similarly, as discussed in Part I.C, *Lucent* has been significantly undermined by the Federal Circuit's later decision in *Apple v. Motorola*.[167] The reason for this eventual undermining of a strict rule requiring plaintiffs to produce rigorous evidence of truly comparable licenses in both cases is the same: the strong desire to award *something* to a patentee who has "suffered an indubitable wrong," rather than dismissing the patentee with "empty hands."[168] To be sure, the specific mechanism for this undermining differs slightly between

161. Cincinnati Car Co. v. N.Y. Rapid Transit Corp., 66 F.2d 592, 595 (2d Cir. 1933) (emphasis added).

162. *See supra* text accompanying notes 51-59.

163. *See supra* text accompanying notes 60-67.

164. Roger D. Blair & Thomas F. Cotter, *Rethinking Patent Damages*, 10 TEX. INTELL. PROP. L.J. 1, 7 n.21 (2001).

165. *See* MERGES & DUFFY, *supra* note 63, at 931.

166. JOHN SKENYON ET AL., PATENT DAMAGES LAW AND PRACTICE § 3:15 (2016) ("[A] patent owner [is] no longer required to prove each of the *Rude* requirements to obtain damages.").

167. *See supra* text accompanying notes 69-73.

168. Cincinnati Car Co. v. N.Y. Rapid Transit Corp., 66 F.2d 592, 595 (2d Cir. 1933); *see also* U.S. Frumentum Co. v. Lauhoff, 216 F. 610, 617-18 (6th Cir. 1914) ("To send the successful plaintiff away after years of litigation and with only nominal damages is repellent to the sense of justice.").

the two cases. In the context of *Rude*, courts moved claims for royalty damages to a new doctrinal silo—from "established royalty" to "reasonable royalty"—which allowed them to shed disfavored prior precedent without any formal inconsistency between decisions (*Rude* is technically still good law with respect to claims for established royalties;[169] it is just that hardly anybody ever makes such a claim). In the context of *Lucent*, the inconsistency with the later decision in *Apple* is more direct.[170] But, despite this superficial difference, the underlying dynamics are the same. Patent law is essentially having the same debate about damages in the twenty-first century as it had at the end of the nineteenth; we have made no progress in over one hundred years. Explicating the information-forcing dilemma helps bring this (depressing) point out clearly.

III. SECOND-LEVEL REACTIONS TO THE DILEMMA

Thus far, I have described what might be considered the "first-level" version of the dilemma. That is, low-quality damages information leads to courts imposing a high burden of proof (to force better information). A case then arises where a plaintiff has clearly suffered loss but cannot prove the exact amount, and the court is unwilling to impose a harsh zero-damages award. The aversion to harsh zero-damages awards, in turn, undermines the information-forcing effect, landing us back to square one. As Part II demonstrated, even this basic version of the dilemma has much explanatory power in providing a way to understand the development of damages doctrine in many areas of law.

This Part goes a step further. Although courts have not fully understood the dilemma, they have always been facing the underlying problem, and courts and legislatures have reacted to it in various ways beyond simply oscillating between strict and loose burdens of proof. This Part will discuss additional strategies that courts and legislatures have tried. The payoff is twofold. First, the information-forcing dilemma provides a unifying lens that allows us to see all of these additional strategies—which often bear little superficial resemblance to each other—as reactions to the same

169. Blair & Cotter, *supra* note 164, at 7 n.24.

170. *See supra* text accompanying note 76.

underlying problem. Second, fully appreciating the nature of the underlying dilemma allows us to understand why these attempted solutions have not worked. Namely, because lawmakers have not fully understood the problem they were trying to solve, the solutions they attempt have always been incomplete, and whichever side of the dilemma is not addressed by a solution ends up reasserting itself over time.

A. The Fact/Amount Rule

Courts have not been entirely unaware of the dilemma described in this Article. Indeed, many courts have recognized one facet or another of the problem. For example, in the early case of *Gilbert v. Kennedy*, the Supreme Court of Michigan described the downside of a strict approach in these terms:

> [C]ases will often occur in which, *first*, no reliable data, no element of certainty can be found by which to measure with accuracy the actual amount of the damages, though it is evident to the court and jury that large damages have resulted from the injury .... [A]re the jury to give merely nominal, or what is the same thing, *no* damages, and is the injured party to obtain no redress, because the case happens to be one which does not furnish a rule for their accurate measurement? ... To deny the injured party the right to recover any actual damages in such cases, because they are of a nature which cannot be thus certainly measured, would be to enable parties to profit by, and speculate upon, their own wrongs, encourage violence and invite depredation.[171]

The most common doctrinal result to emerge from this line of reasoning is the fact/amount rule: "[O]nce the plaintiff has met its burden of proving the fact of damage, some uncertainty with respect to the amount of damages will not preclude recovery."[172] This

171. 22 Mich. 117, 129-30 (1871).

172. J.F. Feeser, Inc. v. Serv-A-Portion, Inc., 909 F.2d 1524, 1539 (3d Cir. 1990) (citing Story Parchment Co. v. Paterson Parchment Paper Co., 282 U.S. 555, 562 (1931)); *see also* Adams v. United States, 358 F.2d 986, 993 (Ct. Cl. 1966) (per curiam) ("[U]ncertainty as to the amount of the damage does not preclude recovery where the fact of damage is clearly established." (quoting Addison Miller, Inc. v. United States, 70 F. Supp. 893, 900 (Ct. Cl.

statement, or something substantially similar, appears throughout the case law and in leading treatises.[173]

Once analyzed, it becomes apparent that the fact/amount rule does not escape the dilemma but simply represents a statement favoring the "award the court's best guess" approach. If uncertainty as to the value of a loss does not preclude recovery—so long as the court is certain of the fact of injury and so the amount is more than zero—what is a court supposed to award except its best guess at the correct amount? The fact/amount rule's stated principle of "do not award zero in situations of uncertainty" thus logically[174] translates into a corollary principle of "award the best guess in situations of uncertainty."[175] This lands us right back at the problem that, if courts always awarded their best guess at the correct amount in situations of inadequate information, then plaintiffs would have no incentive to produce rigorous evidence and better information.

It should therefore be unsurprising that there is a line of counter-cases that state that plaintiffs must prove both the fact *and amount* of damages with reasonable certainty,[176] with the implication that failure to prove the value of the loss with sufficient certainty will lead to zero damages.[177] This counter-rule also appears in leading

---

1947))).

173. *See, e.g.*, 22 AM. JUR. 2D *Damages* § 719 (2013) ("Damages are considered 'speculative' if there is an uncertainty concerning the fact of damages, not the amount."); 25A C.J.S. *Damages* § 363 (2017) ("Where it clearly appears that a party has suffered damage, uncertainty as to the exact amount thereof will not bar recovery."); DUNN, *supra* note 94, § 1.8 (collecting cases).

174. A less-principled third option is to award some intermediate amount. I address that possibility in Part III.C.

175. *See* Story Parchment Co. v. Paterson Parchment Paper Co., 282 U.S. 555, 565 (1931) ("[S]uch reasonable and probable estimate, as in the exercise of good sense and sound judgment they shall think will produce adequate compensation." (quoting *Gilbert*, 22 Mich. at 131)).

176. *See, e.g.*, Cole v. Homier Distrib. Co., 599 F.3d 856, 864 (8th Cir. 2010) ("[A] plaintiff in a contract action generally has the burden of proving 'the existence and amount of ... damages with reasonable certainty.' ... The failure to meet this burden properly results in summary judgment." (omission in original) (quoting Ulrich v. CADCO, Inc., 244 S.W.3d 772, 779 (Mo. Ct. App. 2008))); Aldon Indus., Inc. v. Don Myers & Assocs., 517 F.2d 188, 191 (5th Cir. 1975).

177. S.C. Johnson & Son, Inc. v. Louisville & Nashville R.R. Co., 695 F.2d 253, 261 (7th Cir. 1982) (affirming judgment of no damages because "the admitted fact of damage is insufficient to prove the amount of damage" (quoting Locklin v. Day-Glo Color Corp., 429 F.2d 873, 879 (7th Cir. 1970))).

treatises.[178] The two logically incompatible lines of case law then sit side-by-side, their incompatibility unacknowledged but very real, creating messy doctrine.[179]

B. Reversing the Burden of Proof

To this point, I have presented the dilemma as a complication that arises when courts attempt to apply the burden of proof to—and force information from—the *plaintiff*. Yet there is no intrinsic reason that courts must confine their information-forcing attempts to the plaintiff side of the ledger. A court can place the burden of proof on either plaintiffs or defendants.[180] And if a court places the burden of proving damages on a defendant, then the harsh result of denying recovery to an injured plaintiff merely because plaintiff cannot prove the precise amount of the loss will not arise. Superficially, therefore, reversing the burden of proof seems a viable solution to the dilemma described in Part I.

The copyright statute can be understood as incorporating the logic described in the prior paragraph. As background, copyright law often faces a situation where less than all of the revenue from an infringing product is properly attributable to copyright infringement.[181] For example, if a book by Taney makes one hundred dollars in sales and contains one chapter that is plagiarized from Marshall, then Taney infringes Marshall's copyright. But it would be difficult to argue that Marshall should receive the entire one hundred dollars in revenue from Taney's book—some of that revenue is properly attributable to Taney's own contributions and to the expenses of printing the book.[182] A court cannot easily determine the

178. *See, e.g.*, 22 AM. JUR. 2D *Damages* § 719 (2013) ("Generally, the burden of proof is upon the plaintiff ... to show the fact and extent of an injury and to show the amount and value of the plaintiff's damages."); 25A C.J.S. *Damages* § 305 (2017) ("The plaintiff also has the burden of proving the amount of damages with reasonable certainty." (footnotes omitted)).

179. *See* LAYCOCK, *supra* note 46, at 104 (noting the tension between the fact/amount rule and cases that reject plaintiff's proof as speculative "even when it seems clear that plaintiff suffered some damages").

180. Hay, *supra* note 26, at 652 (providing some considerations in allocating the burden of proof). In principle, a court can even force information from third parties, FED. R. CIV. P. 45(b)(2), (g), but I will limit my consideration to only parties in litigation.

181. *See, e.g.*, Sheldon v. Metro-Goldwyn Pictures Corp., 309 U.S. 390, 396 (1940).

182. *See id.* at 405 ("[W]e perceive no ground for saying that ... the court may make an award of profits which have been shown not to be due to the infringement.").

relative proportion by some mechanical formula,[183] such as using the number of plagiarized pages to the number of nonplagiarized pages—after all, the pages written by Marshall might be really important, even if small in number. Because courts do not inherently have sufficient information to make the apportionment decision, they need to impose a burden of proof to obtain better information.

By its terms, § 504(b) of the copyright statute allocates the burden of proof as follows: "In establishing the infringer's profits, the copyright owner is required to present proof only of the infringer's gross revenue, and the infringer is required to prove his or her deductible expenses and the elements of profit attributable to factors other than the copyrighted work."[184]

In effect, this statute allocates the burden of proving the value of the copyrighted portion to the defendant.[185] That is, the value of the copyrighted portion of a larger work must logically be the total value of the larger work, minus everything that is *not* attributable to the copyrighted portion. By requiring the defendant to prove how much of the profit from the larger work is *not* attributable to copyright infringement, therefore, the copyright statute implicitly makes the defendant also prove how much of the profit *is* so attributable. And allocating the burden of proof to the defendant avoids the problem of plaintiffs getting stuck with zero awards due to failure of proof.

However, the statutory methodology produces a problem on the other side: literally applied, it tends to produce absurdly large awards. Consider, for example, the case of *Polar Bear Productions, Inc. v. Timex Corp.*, in which the watchmaker Timex used some copyrighted content owned by Polar Bear in its advertising for

183. *See id.* at 404 (discussing the "the difficulty of making an exact apportionment").

184. 17 U.S.C. § 504(b) (2012). Strictly speaking, this provision relates to disgorging profits and not to compensatory damages. The two measures are not always the same because if an infringer is a more efficient producer than the copyright holder, then he will earn more profits than the copyright holder otherwise could have, and the extra profits do not represent a "loss" as such. *See* Robert L. Birmingham, *Breach of Contract, Damage Measures, and Economic Efficiency*, 24 RUTGERS L. REV. 273, 284 (1970) (describing the concept of an efficient breach of a legal obligation). However, the problem of inadequate information applies equally when calculating compensatory damages and calculating profits, so I will not distinguish between them in this Article.

185. Frank Music Corp. v. Metro-Golden-Mayer, Inc., 772 F.2d 505, 518 (9th Cir. 1985) ("The burden of proving apportionment ... is the defendant's.").

"Expedition" brand watches.[186] Polar Bear sued and won, and the primary issue was the amount of money to be awarded.[187] Following the plain text of the statute, Polar Bear argued that it was entitled to claim profits based on the total sales of Expedition watches.[188] The court rejected this claim, reasoning that such a methodology would result in an excessive award:

> Although the statute only references the broad term "gross revenue," to conclude that a copyright plaintiff need only provide the company's overall gross revenue, without regard to the infringement, would make little practical or legal sense. Otherwise, the plaintiff in a copyright action against a multidivision, multi-product company such as General Mills, would need to do nothing more than offer an overall gross revenue number—like $11.5 billion—and sit back.[189]

The court instead held that "a copyright plaintiff must present a modicum of proof linking the infringement to the profits sought" and vacated the jury award because Polar Bear had failed to present sufficient evidence.[190]

The point that § 504(b), when literally applied, produces absurdly large awards is not a new one. Courts have repeatedly noted this problem.[191] The contribution of my analysis is not to simply repeat this well-known point but to add more depth to the analysis: to explain what § 504(b) is trying to accomplish, why rational lawmakers would ever have enacted it, and how attempts to overcome the downsides of § 504(b) lead to their own problems.

From the perspective of the information-forcing dilemma, § 504(b) is a second-level reaction to the problems that arise from a first-level attempt to solve the dilemma. That is, because courts do not inherently possess information about the correct apportionment of value between the infringing aspects of a work and the nonin-

186. 384 F.3d 700, 703-04 (9th Cir. 2004).

187. *Id.* at 703.

188. *Id.* at 715 ("Polar Bear claimed a portion of Timex's profits from all Expedition watch sales.").

189. *Id.* at 711.

190. *Id.* at 715-16.

191. *See, e.g.*, Bonner v. Dawson, 404 F.3d 290, 294 (4th Cir. 2005); Davis v. The Gap, Inc., 246 F.3d 152, 160 (2d Cir. 2001); Taylor v. Meirick, 712 F.2d 1112, 1122 (7th Cir. 1983).

fringing aspects, courts must obtain the needed information from parties using an information-forcing mechanism.[192] If we followed traditional legal principles and required the plaintiff to prove the correct apportionment,[193] then what would happen is that we would get the first-level dilemma described in Part I, where plaintiffs would often be unable to prove the correct apportionment with any precision and would therefore get nothing. In order to combat the problem of zero awards, Congress reversed the burden of proof and placed the burden of proof on defendants.[194]

This second-level reaction to the first-level dilemma, however, does not really solve the dilemma but simply changes the direction in which the problem manifests. We no longer have a problem of plaintiffs facing unjust, undercompensatory, zero-damages awards because they cannot comply with an impossibly onerous burden of proof. We get instead a problem of defendants facing unjust, overcompensatory, total-revenue awards because they cannot comply with an impossibly onerous burden of proof.[195] And relaxing the burden of proof in both instances will undermine the information-forcing effect and leave courts struggling to come up with the correct apportionment by themselves. The shoe is on the other foot, but the fundamental dynamics of the dilemma remain the same.

Courts react to this second-level problem by creating a third-level solution: they require copyright holders like Polar Bear to prove a causal link between the infringement and the claimed profits.[196] But this is just re-reversing the burden of proof. Now Polar Bear must prove how much profit is attributable to infringement, in default of

192. *See supra* Part I.

193. HALE, *supra* note 33, § 31 ("Losses must be certain in amount ... or damages therefor cannot be recovered. The burden ... is on the plaintiff.").

194. *See* Lottie Joplin Thomas Tr. v. Crown Publishers, Inc., 592 F.2d 651, 657 (2d Cir. 1978) ("[P]lacing the apportionment burden on defendants best comports with our admonition that in apportioning profits 'every indulgence should be granted plaintiff in an attempt to arrive at a sum which is assuredly adequate.'" (quoting Orgel v. Clark Boardman Co., 301 F.2d 119, 121-22 (2d Cir. 1962))).

195. Sheldon v. Metro-Goldwyn Pictures Corp., 106 F.2d 45, 49 (2d Cir. 1939) ("[I]t is nearly as unfair to cast the infringer for all the profits, as it would be to deny the patentee or author any recovery whatever."), *aff'd*, 309 U.S. 390 (1940).

196. Polar Bear Prods., Inc. v. Timex Corp., 384 F.3d 700, 715 (9th Cir. 2004); *see, e.g.*, Bouchat v. Balt. Ravens Football Club, Inc., 346 F.3d 514, 520-23 (4th Cir. 2003); Univ. of Colo. Found., Inc. v. Am. Cyanamid Co., 196 F.3d 1366, 1375 (Fed. Cir. 1999).

which it cannot recover any profit at all.[197] Of course, plaintiffs like Polar Bear will have tremendous difficulty precisely proving how much profit is attributable to the infringement, and strictly enforcing the burden of proof means that Polar Bear recovers nothing—even though nobody thinks that Timex truly made no profit from its infringement.[198] In short, this third-level reaction takes us full circle back to the dilemma in its original form.

C. Intermediate Award Amounts

If the reaction that triggers the dilemma is that a zero award feels particularly harsh when a plaintiff fails the burden of proof, then an obvious solution is that a court should not award zero, but should award an intermediate amount—that is, more than zero, but less than the court's best estimate of the fully compensatory amount. Awarding an intentionally low-ball amount when the court is unsatisfied with the plaintiff's damages evidence would at first glance seem to achieve the best of both worlds: it avoids the extreme harshness of a zero award, yet it still provides plaintiffs with some incentive to produce better evidence in future cases because successfully proving the true value of the loss would lead to a higher expected award than the intentionally low-ball guess.

This strategy is rarely explicitly suggested—what court is going to admit to intentionally giving an incorrect amount?—but one might suspect that it occurs quietly in practice. And hints of this strategy can occasionally be found if one looks carefully. For example, in *Apple v. Motorola*, after the Federal Circuit rejected the trial judge's awarding zero damages for insufficient evidence,[199] the court went on to say that "if the patentee's proof is weak, the court is free to award a low, perhaps nominal, royalty."[200] This suggests that district courts may somewhat reduce the damages award as a

197. *Polar Bear*, 384 F.3d at 716 (vacating entire award).

198. *See id.* at 715 (acknowledging this with the statement "even if we suspect that Timex derived some quantum of profits from the infringement because its infringement was part of promotional efforts").

199. *See supra* text accompanying notes 69-73.

200. 757 F.3d 1286, 1328 (Fed. Cir. 2014), *overruled on other grounds by* Williamson v. Citrix Online, LLC, 792 F.3d 1339 (Fed. Cir. 2015).

penalty for insufficient evidence, so long as the reduction does not go all the way to zero.

For purposes of my argument here, the important point is that an intermediate approach is not a *refutation* of the dilemma but a *manifestation* of it.[201] That is, the dilemma is about a tradeoff between promoting incentives for rigorous information in the long run versus the desire to reach just outcomes in an immediate case. This tradeoff is not a dichotomy but a spectrum. I have devoted attention thus far to describing the extreme ends of the spectrum—either awarding no damages at all (strongest information-forcing incentives), or awarding full compensation (best outcome for the individual case)—simply to illustrate the consequences that are at stake in the tradeoff.[202] An intermediate result within this spectrum mixes those consequences but does not escape the tradeoff. A court that awards a nonzero but low-ball amount creates a *somewhat* undercompensatory result that *somewhat* frustrates the policy goals of compensatory damages (for example, if patentees know that courts give low-ball awards, this will negatively affect the incentives to invent that the patent system is designed to foster), while it also *somewhat* undermines future incentives for rigorous evidence.[203] The essential dilemma remains unchanged. A court could provide stronger incentives for better evidence by going closer to a zero award, but only at the cost of a harsher outcome in the individual case before it. Conversely, a court could create a better outcome in the case before it by going closer to its estimate of the fully compensatory amount, but only at the cost of more severely undermining future incentives for rigorous evidence. Courts that desire to achieve both objectives are thus still trapped in a dilemma about what to do.

The fact that intermediate award amounts represent a pragmatic compromise between two principles—following neither one to its logical conclusion—also means that as a solution it fits awkwardly within our legal system, which is one where courts are supposed to make decisions according to logic and principle.[204] Courts do not

201. In this sense, an intermediate award strategy is not even a second-level reaction to the dilemma but is a first-level manifestation of the dilemma itself.

202. *See supra* Part II.

203. *See supra* text accompanying notes 41-43.

204. *See* Herbert Wechsler, *Toward Neutral Principles of Constitutional Law*, 73 HARV. L. REV. 1, 15 (1959) ("[T]he main constituent of the judicial process is precisely that it must be

openly say that they award intermediate amounts because they have no principle to explain and justify such decisions;[205] an intermediate award is neither the correct amount the victim deserves as compensation, nor is it the obvious amount to award as a penalty for failing the burden of proof. Without a principle, intermediate awards are a solution in practice, but not in theory, and the lack of a theory means that such damages awards are (and will continue to be) unpredictable.

D. Statutory Benchmarks and Caps

If the problem of damages fundamentally arises because courts lack the information to determine the true value of a loss, then one way to circumvent all of this is to have a legislature prescribe some alternative benchmark upon which to base damages calculations. To the extent the new benchmark is easy to apply, there would be no information difficulty and no dilemma.

Section 289 of the design patent statute provides an example of this kind of statutory benchmark. This section states that an infringer of a design patent is liable to a prevailing patent owner "to the extent of his total profit."[206] Setting the benchmark for design patent damages at the defendant's total profit rather than at the value of the patented design means that design patent damages are relatively easy to compute. For example, if Samsung infringes an Apple design patent by using the iTunes icon in a Galaxy smartphone, then it is difficult to determine how much the iTunes icon is worth as a proportion of the entire smartphone; but it is easy to

---

genuinely principled, resting with respect to every step that is involved in reaching judgment on analysis and reasons quite transcending the immediate result that is achieved.").

205. In some sense, my proposal in Part IV can be understood as providing a principle for how to calculate intermediate awards, though as a matter of semantic clarity I would prefer to style deviations from the court's best estimate of the correct amount as an explicit "penalty."

206. 35 U.S.C. § 289 (2012). A reader might notice some similarity between this statute and 17 U.S.C. § 504(b) (2012). The two are different at a theoretical level. Section 504(b) is ultimately trying to determine the value of the copyrighted work; it merely places the burden on the defendant to prove this value by deducting everything not attributable to the copyrighted portion. *See supra* note 184. Section 289 is not trying to determine the value of the patented feature and thus has no provision for deducting nonpatented features.

determine Samsung's total profit.[207] A statutory benchmark thus solves the dilemma in some sense.

The information-forcing dilemma nonetheless provides many insights about a statutory benchmark like § 289. First, statutory benchmarks are often created as a second-level reaction to a first-level incarnation of the dilemma. Section 289 itself was enacted in the aftermath of the Supreme Court's decision in *Dobson v. Hartford Carpet Co.*, in which the Court held that "[t]he burden [of proving damages] is upon the plaintiff, and, if he fails to give the necessary evidence, but resorts, instead, to inference and conjecture and speculation, he must fail for want of proof,"[208] and accordingly directed an award of six cents to the patentee.[209] By setting the damages benchmark at total profit rather than at the true value of the patented feature, § 289 makes it easy for a patentee to prove the relevant benchmark and solves the problem of harsh zero (or six-cent) awards.

The second insight is that, as a reaction to the first-level incarnation of the dilemma, a statutory benchmark solves one side of the problem, but at the price of neglecting the other side. Specifically, a statutory benchmark makes damages computations easy, but, because damages are no longer measured according to a benchmark of compensating the plaintiff for the value of his loss, awards become unjust. To return to the example of Samsung using the iTunes icon, few people would find it sensible to award Apple *all* of Samsung's profit for infringement by a minor feature.[210]

The unjust outcomes generate a third-level reaction, which is that courts attempt to reform the benchmark to achieve better correspondence between damages awards and actual losses. This is exactly what happened with § 289. In the recent case of *Samsung Electronics Co. v. Apple Inc.*, the Supreme Court held that § 289 should be understood not as granting the patentee a right to the total profit of

207. *See generally* SAMSUNG, 2014 SAMSUNG ELECTRONICS ANNUAL REPORT 3 (2014), http://www.samsung.com/us/aboutsamsung/investor_relations/financial_information/downloads/2015/SECAR2014_Eng_Final.pdf [https://perma.cc/ZM4M-RSRW].

208. 114 U.S. 439, 444 (1885).

209. *Id.* at 447; *see also* Dobson v. Dornan, 118 U.S. 10, 18 (1886) (directing six-cent award for failure of damages proof).

210. *See* Mark A. Lemley, *A Rational System of Design Patent Remedies*, 17 STAN. TECH. L. REV. 219, 221 (2013) (arguing the total profit rule "makes no sense" and "drastically overcompensates the owners of design patents").

the infringer, but only to the profit based on the particular "article of manufacture" to which the patent pertained.[211] More importantly, the Court held that the relevant "article of manufacture" is not necessarily the finished product that is sold to consumers, but could be a smaller component such as a screen or case.[212] The Court did not, however, further specify what constituted the relevant "article of manufacture"—when can it be a component instead of the finished product? And how small a component can count?—and instead simply remanded the case to the lower court.[213]

The payoff of understanding the dilemma is that it allows us to see both the impetus for the Court taking the case to begin with, as well as the reasons for its minimalist and uninformative eventual disposition. The impetus for taking the case quite clearly is that the most obvious interpretation of the statute—patentees receive a defendant's entire profit from the infringing sale of a finished product—leads to overcompensatory awards when the product includes both patented and unpatented components. By potentially limiting the recovery to profits attributable to a smaller component, the Court can avoid grossly overcompensatory outcomes.

But doing so creates a new problem—the court now has an information-forcing dilemma in trying to ascertain the amount of profit attributable to a component, which, unlike a finished product, does not have easily observed prices and profits. If the court requires the plaintiff to *prove* the profit attributable to the component, then the plaintiff will fail to come up with a precise number, and the court will be faced with the choice of either awarding zero—an undercompensatory outcome that the statute clearly is designed to avoid[214]—or awarding its own guess—which will be arbitrary, as well as reduce incentives for evidence production in the future. If the court requires a defendant to prove the amount, the same dilemma emerges, in that defendants will also have difficulty proving a precise profit number. The court can again react to this failure of proof either by awarding its own guess (same problems as above), or it can impose some kind of penalty on the defendant, such

211. 137 S. Ct. 429, 434 (2016) (quoting 35 U.S.C. § 289 (2012)).

212. *See id.* at 435 (quoting 35 U.S.C. § 289).

213. *See id.* at 436.

214. More precisely, "$250," since the statute provides this minimum amount. 35 U.S.C. § 289 (2012). Of course, given the cost of litigation, a $250 recovery is effectively a zero award.

as by giving the total profit for entire product to the plaintiff, which brings back the overcompensation problem. By issuing a minimalist opinion and remanding, the Supreme Court avoided definitively choosing any of these unsavory options,[215] but it does not solve the information difficulty and instead simply passes the problem to the lower court.

IV. PAYOFFS AND A SOLUTION

Thus far, the payoffs of my analysis have mostly been a matter of conceptual clarity. Understanding the information-forcing dilemma allows a reader to see past vague and unhelpful doctrinal formulations and better understand what damages doctrine in various fields is about. It also reveals that the ostensibly different doctrinal inquiries in various fields share many commonalities.[216] Once we look beyond the surface, it turns out that damages debates in patent, contract, tort, and antitrust law share many similarities, even though the literature in these areas rarely interact.[217]

Better conceptual clarity also helps explain why debates keep going in circles and making no progress. Stated simply, the dilemma is a multisided problem, but courts, legislators, and academic commentators have tended to focus on one side or another at any one time, depending on what is salient at the particular time or in a particular case. When courts notice a problem of low quality evidence, they react—as the court in *Lucent* did—by raising the burden of proof,[218] without realizing that this can solve the problem of low quality evidence only at the price of creating a problem of unjust zero-damages awards. When courts notice the problem of unjust zero-damages awards, they react—as the court in *Apple* did—by directing juries to take their best guess,[219] without realizing that this destroys the incentive for rigorous evidence. Without understanding the full nature of the problem, any attempted

215. *See Samsung*, 137 S. Ct. at 436.

216. *See supra* Part II.

217. *See* John M. Golden & Karen E. Sandrik, *A Restitution Perspective on Reasonable Royalties*, 36 REV. LITIG. (forthcoming 2017) (arguing that patent law should draw on general restitution principles to better theorize how damages should be calculated).

218. *See supra* text accompanying notes 60-67.

219. *See supra* text accompanying notes 69-73.

solution will necessarily be incomplete, and whatever facet of the problem gets neglected will resurface later to undermine the solution. The law therefore becomes either full of self-contradictory pronouncements (for example, plaintiffs must prove the amount of their loss,[220] except when they do not have to[221]), or is consistent only at the level of vague generalities that provide no guidance (for example, that damages must be proven with "reasonable" certainty).

This kind of conceptual payoff may not satisfy the more practical-minded readers in my audience, who ask what courts should *do*. This Part will provide an answer to that question. Courts should require a party to produce damages information—which is the same as saying that a court should penalize a party for failing to produce information—if and only if two conditions are met. First, the social benefit of having additional information on some issue must outweigh the social cost of collecting the information and presenting it in court (the cost-effectiveness criterion). Second, courts should impose the burden of proof on the party that can more cheaply produce the information required (the cheaper-producer criterion). Put together, this "cheaper cost-effective producer" principle provides a coherent principle to guide courts on how to approach the dilemma.

My claim here is both normative and descriptive. My argument is not only that courts ought to craft damages doctrine according to the cheaper cost-effective producer principle, but also that the existing doctrine can be understood to have *already* done this to some extent. Viewed through this lens, the vague and messy existing doctrine is not just an incoherent morass; it is a clumsy and imperfect attempt by courts to arrive at this unifying principle, which courts have somewhat intuited but have not been able to articulate. Clearly articulating the principle therefore allows the doctrine to be coherently interpreted and better implemented in the future.

A. The Cost-Effectiveness Criterion

The cost-effectiveness criterion states that courts should require the production of a particular piece of information only when the

220. *See supra* notes 176-78 and accompanying text.
221. *See supra* text accompanying notes 178-81.

benefits of having the information exceed the costs of collecting and producing the evidence to prove it. The normative justification for this criterion should be self-evident: the law should not do something (or require something to be done) unless the benefits exceed the costs. The same justification underlies cost-benefit tests throughout the law,[222] the most famous being the Hand formula for negligence.[223]

It is important to emphasize that the test is whether the *social* benefits exceed the *social* costs. If the private benefits of a party disclosing information to a court exceed the private costs, then a party would disclose the information voluntarily—there would be no need for a court-imposed, information-forcing rule such as a burden of proof to compel the disclosure.[224] The relevant question is whether the social benefits of having a particular piece of damages information—the social benefits being that a court will be able to render a more informed decision, which in turn better accomplishes whatever social policy objective the respective legal regime is designed to accomplish[225]—outweighs the social cost of collecting and producing the evidence required to furnish a court with the information it seeks.[226]

This cost-effectiveness criterion is admittedly not susceptible to any kind of direct mathematical implementation. Just as courts do not have sufficient information to precisely calculate the value of a loss (what we might call "first-order" information deficiency), they are also unlikely to have sufficient information to calculate the costs

---

222. *See, e.g.*, 2 U.S.C. § 1532(a)(2) (2012) (requiring cost-benefit analysis for significant federal regulatory actions); Richard H. Pildes & Cass R. Sunstein, *Reinventing the Regulatory State*, 62 U. CHI. L. REV. 1, 3-6 (1995) (discussing various Executive Orders by Presidents Reagan and Clinton that define "cost-benefit analysis as the basic foundation of decision" for the regulatory state).

223. United States v. Carroll Towing Co., 159 F.2d 169, 173 (2d Cir. 1947) (defining reasonable care in cost-benefit terms).

224. *See* Alex Raskolnikov, *Accepting the Limits of Tax Law and Economics*, 98 CORNELL L. REV. 523, 531 (2013) (explaining that a basic economic justification for government intervention is to "align private and social costs and benefits where they would otherwise diverge").

225. *See* Louis Kaplow & Steven Shavell, *Accuracy in the Assessment of Damages*, 39 J.L. & ECON. 191, 192 (1996) (explaining the social benefit of accurate damages determinations).

226. The social costs are likely to more closely correspond to the private costs, in that the private costs of information production—such as hiring an investigator to find evidence—are also generally social costs, since they consume social resources that could be deployed elsewhere.

and benefits of obtaining better information on the value of a loss (what we might call "second-order" information deficiency). In particular, courts face a catch-22 in trying to assess the value of additional information when, by definition, they do not know the content of the information they are missing.[227] The sum is that, as a practical matter, courts are not likely to have more than a faint intuition about what additional information might be relevant to damages, what the cost of collecting and producing that information would be, how much a judicial damages determination would be enhanced if the information were produced, and the exact value of the social benefit that comes from marginally better judicial decision-making. All of these things would need to be known if courts were to truly calculate the costs and benefits of additional information.

My goal in articulating the cost-effectiveness criterion is not to suggest that courts should attempt to actually calculate the costs and benefits. As an analogy, no court attempts to calculate the variables in the Hand formula.[228] The point of the Hand formula in tort law is not to allow courts to actually quantify B, P, and L,[229] but to give a meaningful definition of what constitutes "reasonable care," so that courts have something to analyze rather than simply throw the entire question to a black-box jury.[230] Similarly, in the absence of a clear articulation of the cost-effectiveness criterion, the standards of proof in damages law are vague, incoherent, and often self-contradictory.[231] Nobody knows what "reasonable certainty" or "just and reasonable inference" actually means. What the

227. *Cf.* Kenneth J. Arrow, *Economic Welfare and the Allocation of Resources for Invention*, *in* THE RATE AND DIRECTION OF INVENTIVE ACTIVITY: ECONOMIC AND SOCIAL FACTORS 609, 615 (Nat'l Bureau of Econ. Research ed., 1962) (describing the informational paradox where a potential purchaser of information cannot assess its value without first knowing the content of the information).

228. Moisan v. Loftus, 178 F.2d 148, 149 (2d Cir. 1949) (Hand, J.) ("[A]ll such attempts [at calculating the variables in the Hand formula] are illusory.").

229. These refer to the burden (B), probability (P), and amount of loss (L). The Hand formula defines negligence as the failure to take a precaution if $B < PL$. *See* United States v. Carroll Towing Co., 159 F.2d 169, 173 (2d Cir. 1947).

230. *See* McCarty v. Pheasant Run, Inc., 826 F.2d 1554, 1556-57 (7th Cir. 1987) (Posner, J.); Gregory C. Keating, *Reasonableness and Rationality in Negligence Theory*, 48 STAN. L. REV. 311, 331 (1996) ("The real significance of the Hand Formula, then, is not *technical*, but *conceptual*: It isolates the elements of due care and the relations among them.").

231. *See supra* Part II.

cost-effectiveness criterion provides is a way to more coherently understand these terms. Viewed through a cost-effectiveness lens, a plaintiff proves the value of a loss with "reasonable certainty" if and only if a court believes that the plaintiff has produced all the information that could be cost-effectively produced.[232] Conversely, a plaintiff's proof is deficient and "speculative" if the court believes that a plaintiff has not produced all the information that could be cost-effectively produced.[233]

The cost-effectiveness criterion thus helps clarify damages doctrine in two ways. First, it makes clear that there *is* a balance to be struck. Courts already understand this in their actions, but not always in their words. That is, the case law is replete with absolutist statements that suggest the standard of proof is alternatively either infinitely onerous, such as requiring a plaintiff to prove the value of a loss with certainty,[234] or nonexistent, as in the most extreme version of the fact/amount rule.[235] Of course, no court ever truly applies such an extreme standard—even antitrust courts will find a plaintiff to have failed the burden of proof at some point[236]—but the cost-effectiveness criterion helps dispel the semantic confusion that often pervades judicial opinions.

Second, the cost-effectiveness criterion helps explicate the considerations that go into the balance. A court determining whether a plaintiff (or defendant) has produced all the cost-effective information possible must analyze what types of additional information it might want, how much the additional information is likely to improve the court's damages estimate, to what extent a better judicial decision would actually affect society at large (rather than simply effect a wealth transfer between the parties),[237] and weigh all

232. *See* Westinghouse Elec. & Mfg. Co. v. Wagner Elec. & Mfg. Co., 225 U.S. 604, 622 (1912) (holding that the burden of proof shifts to the defendant "after the plaintiff has proved the existence of profits attributable to his invention and demonstrated that they are impossible of accurate or approximate apportionment").

233. *See* Mammoth Oil Co. v. United States, 275 U.S. 13, 51 (1927) ("It is certainly a maxim that all evidence is to be weighed according to the proof which it was in the power of one side to have produced.").

234. *See supra* notes 33-34 and accompanying text.

235. *See supra* text accompanying notes 172-77.

236. *See supra* text accompanying notes 123-24.

237. *See* Louis Kaplow & Steven Shavell, *Legal Advice About Information to Present in Litigation: Its Effects and Social Desirability*, 102 HARV. L. REV. 565, 589 (1989) (arguing that more judicial information is socially significant only if individuals comprehend that infor-

this against the likely costs of collecting and producing the additional information. As discussed above, courts are not likely to have anything more than faint intuitions regarding these variables.[238] But the point is that the cost-effectiveness criterion allows us to identify these variables as *the* relevant factors for consideration. It provides a *target* for the analysis. Courts may not have the information necessary to actually *hit* the target, but at least they will know what they should be aiming at, whereas without the cost-effectiveness criterion they would have no idea what direction they should be pointing.

I would additionally argue that the cost-effectiveness criterion is not only something that courts *should* aim for, but is something that they already *do* aim for, at least to some extent.[239] For example, when the Supreme Court states that "[t]rial and appellate courts alike must ... observe the practical limits of the burden of proof which may be demanded,"[240] it is implicitly describing a cost-benefit balancing test, where the benefits of a more demanding burden of proof are to be balanced against the costs (or "practical limits"). To be sure, the Court did not go into much detail about what the benefits of a more demanding burden of proof are, or why there are practical limits or costs to demanding too high a burden of proof. That is the contribution of the cost-effectiveness criterion: to *explicate* on the reasoning that the courts have not fully articulated. But the criterion itself can be found in the case law already, albeit implicitly and imperfectly applied.[241]

---

mation which affects judicial decisions).

238. *See supra* text following note 82.

239. *See* LAYCOCK, *supra* note 46, at 104 (arguing that "a fair generalization" of the case law is a principle that the "plaintiff must prove damages with as much certainty as is reasonably possible under the circumstances"). My cost-benefit criterion gives a more precise definition to what constitutes "*reasonably* possible under the circumstances."

240. Zenith Radio Corp. v. Hazeltine Research, Inc., 395 U.S. 100, 123 (1969); *see also* Westinghouse Elec. & Mfg. Co. v. Wagner Elec. & Mfg. Co., 225 U.S. 604, 616 (1912) (stating that the burden of proof "principle must not be pressed so far as to override others equally important in the administration of justice").

241. *See* LANDES & POSNER, *supra* note 14, at 23 ("People can apply the principles of economics intuitively—and thus 'do' economics without knowing they are doing it.").

B. The Cheaper-Producer Criterion

The second criterion is that a court should allocate the burden of proof to the party that can more easily (that is, cheaply) produce the information it seeks. The normative justification for this criterion is also quite obvious: imposing the burden of proof on the cheaper producer allows information to be produced at lower cost, promoting efficiency.[242] Once we understand a burden of proof as an instrumental, information-forcing mechanism—rather than as a moral principle that the party seeking to disturb the status quo has the unique obligation to fulfill—then it quickly follows that courts can impose the burden of proof on either a plaintiff or defendant,[243] and they should choose based on which side is better able to produce the proof required.[244]

More than with the cost-effectiveness criterion, courts have already articulated the cheaper-producer criterion in many cases. The clearest statement of this criterion was made by the Supreme Court in *Selma, Rome & Dalton Railroad v. United States*, where the Court stated that "it has been established as a general rule of evidence, that the burden of proof lies on the person who wishes to support his case by a particular fact which lies more peculiarly within his knowledge, or of which he is supposed to be cognizant."[245] Other cases contain similar statements.[246]

Although a cheaper-producer criterion is not strictly novel, articulating it here nonetheless provides two contributions. First,

242. *See* 6 JEREMY BENTHAM, *An Introductory View of the Rationale of Evidence*, *in* THE WORKS OF JEREMY BENTHAM 136 (1843) (arguing that the burden of proof should be placed on the party who can fulfill it with the "least inconvenience").

243. *See* R.H. Coase, *The Problem of Social Cost*, 3 J.L. & ECON. 1, 2 (1960) (explaining that the law can allocate a social cost to one side or the other).

244. *Cf.* CALABRESI, *supra* note 13, at 135-40, 135 n.1 (articulating a "cheapest ... cost avoider" criterion for allocating the burden of taking precautions).

245. 139 U.S. 560, 568 (1891).

246. *See, e.g.*, Gomez v. Toledo, 446 U.S. 635, 640-41 (1980) (placing the burden of proof on the defendant for qualified immunity because the issue "depends on facts peculiarly within the knowledge and control of the defendant"); United States v. Denver & Rio Grande R.R. Co., 191 U.S. 84, 92 (1903) ("It is a general rule of evidence ... 'that where the subject matter of a negative averment lies peculiarly within the knowledge of the other party, the averment is taken as true, unless disproved by that party.'"); Lindahl v. Office of Pers. Mgmt., 776 F.2d 276, 280 (Fed. Cir. 1985) (placing the burden of proof on the party with the disability to prove inability to work because "[t]he party with the best knowledge normally sustains the burden"); Fleming v. Harrison, 162 F.2d 789, 792 (8th Cir. 1947).

courts are not consistent about applying the criterion, and much of the time they reflexively allocate the burden of proof to the plaintiff, without considering the parties' relative costs of production,[247] based on a moralistic rationale that the party seeking to change the status quo should bear the burden of proof.[248] Second, the academic literature, led by McCormick in his influential treatise, often dismisses the cheaper-producer criterion as an insignificant consideration based on the reasoning that modern discovery practice gives both sides essentially equal access to the same evidence.[249]

On the first point, there is little to say except that applying the cheaper-producer criterion will facilitate better and more efficient adjudication than an unthinking, reflexive allocation of the burden of proof to the plaintiff.[250] On the second, saying that modern discovery equalizes the cost of producing evidence miscomprehends the "evidence" courts need. Let me elaborate on this second point.

A conceptual distinction should be drawn between two senses of the word "evidence." The word can either mean a *class of things* that provide information about some topic (for example, evidence *about* a defendant's financial condition) or it can mean *a concrete thing* within the class (for example, the defendant's prior year 10K statement). When McCormick and other authors argue that modern discovery equalizes the cost of producing evidence, they are using "evidence" in the sense of a concrete thing. And if one conceptualizes the cheaper-producer inquiry as a question of which party is better able to produce a concrete thing that already exists and is in a

247. *See* WRIGHT & MILLER, *supra* note 25, § 5122 ("Courts ... frequently ignore policy in favor of formalistic methods of assigning burdens of proof.").

248. *See* 2 MCCORMICK ON EVIDENCE § 337, at 428 (John W. Strong ed., 4th ed. 1992) ("[B]urdens of pleading and proof ... have been and should be assigned to the plaintiff who generally seeks to change the present state of affairs."); *see, e.g.*, Schaffer *ex rel.* Schaffer v. Weast, 546 U.S. 49, 56 (2005).

249. *See* FLEMING JAMES, JR. ET AL., CIVIL PROCEDURE § 7.16, at 344 (4th ed. 1992) ("Access to evidence is a much diminished basis for allocating burden of proof in modern liberal discovery."); 2 MCCORMICK ON EVIDENCE, *supra* note 248, § 337, at 429 n.11 ("Expanded pretrial discovery would seem to have diminished greatly whatever importance [the relative cost of production] had in allocating the burdens."); Hay, *supra* note 26, at 674-75; Jody S. Kraus, *Decoupling Sales Law from the Acceptance-Rejection Fulcrum*, 104 YALE L.J. 129, 142-43 (1994).

250. GENE R. SHREVE & PETER RAVEN-HANSEN, UNDERSTANDING CIVIL PROCEDURE 346 (1989) (arguing that disturbing the status quo is not a good basis for allocating the burden of proof). *See generally* FED. R. CIV. P. 1 (identifying the purpose of civil procedure as "the just, speedy, and inexpensive determination of every action").

party's possession, such as, "which party is better able to produce the defendant's prior year 10K statement?," then the objection that modern discovery allows each side to demand essentially all the documents of the other side—and so largely equalizes their ability to produce the document to a court—has validity.

But that is not what the cheaper cost-effective producer analysis is about. It is useless to have a court consider what concrete documents it would like to see and then ask which party is better able to produce the document. The very fact that the court *does not have the document* means that the court is unlikely to be able to specifically identify the missing document as something it would like to see.[251] Rather, the inquiry is only meaningful if it is conceptualized at a higher level—as asking for information *about* a topic. The question is not "which party is better able to produce the defendant's 10K statement?" The question is "which party is better able to produce information about the defendant's financial condition?" Framed in this manner, it is obvious that the defendant has a significant advantage. Although the plaintiff might have all of the defendant's preexisting documents, the defendant still has superior *knowledge*—and this means the defendant can identify the most relevant documents, can explain any inconsistencies between the documents, can generate new documents, and knows whom to ask for these tasks.[252] Forcing a defendant to disclose information on this issue will thus be much more effective than trying to coax the information out of a plaintiff.

---

251. *See supra* text accompanying notes 60-67, 218-19.

252. *See* Golden, *supra* note 4, at 585 (arguing that patent defendants are likely to have better knowledge of the value of a component relative to its part in a larger product); *see also* Hay & Spier, *supra* note 30, at 419 ("[D]iscovery does not render irrelevant the question of relative presentation costs.").

C. Implementing the Cheaper Cost-Effective Producer Principle

The most obvious objection to my solution is that courts will not have the information to implement the principle. As explained in Section A, one response to this objection is that I am not envisioning anything more than rough guesses on the costs and benefits, imperfectly and intuitively applied. A second response, however, is that the cheaper cost-effective producer principle is capable of more simple application than one might think. Specifically, courts can implement the principle by using the following methodology: (1) At the close of evidence, a court determines whether either party has been negligent in producing evidence; if no party is negligent then the court proceeds to the merits determination. (2) If the court believes one party has been negligent, it penalizes the negligent party (by reducing damages if the negligent party is the plaintiff, or increasing them if the defendant is the negligent party). (3) If the court believes both parties have been negligent, it penalizes the more negligent party.

By using the language of "negligence," I am obviously drawing on the analogy to tort law, and the economic analysis of negligence law explains why the methodology above effectively implements the cheaper cost-effective producer principle.[253] What I am proposing is essentially a comparative negligence regime: in ordinary tort law, "negligence" is the failure to take a cost-benefit-justified precaution to prevent a social loss[254]—the social losses typically being things like car accidents and the precautions being things like careful driving. In my proposal, "negligence" is also the failure to take cost-benefit-justified precautions to prevent a social loss; the only difference is that the social loss is an ill-informed judicial decision and the precaution to prevent this loss is the production of evidence. A court determining whether any party has negligently failed to produce evidence, therefore, is effectively determining whether all reasonably available—that is, cost-effective—information has been

253. *See* Mark F. Grady, *Common Law Control of Strategic Behavior: Railroad Sparks and the Farmer*, 17 J. LEGAL STUD. 15, 15 (1988) (explaining how economic analysis "divulges the basic structure of negligence liability").

254. Richard A. Posner, *A Theory of Negligence*, 1 J. LEGAL STUD. 29, 32 (1972).

produced.[255] Rules (1) and (2) thus combine to effectuate the cost-effectiveness criterion.

Rules (2) and (3) combine to effectuate the cheaper-producer criterion. The effect of (2) is obvious enough—if only one party has not produced all the evidence it can cost-effectively produce, then that party is the cheaper producer of additional damages evidence.[256] Rule (3) overcomes an intuitive problem that a court cannot formally allocate a burden of proof on both sides at the same time. But by threatening to penalize whichever side is *more* negligent, the court sets off a virtuous cycle where the threat of a penalty on one side (say the plaintiff) induces the plaintiff to produce more and better evidence, which in turn induces the defendant to produce more and better evidence, and so on—the cycle stops only at the point where both parties have produced all cost-effective evidence and therefore do not expect to be found negligent at all.

One obvious question is how much the penalty should be. In theory, if parties are fully confident that courts will only apply the penalty when a party has been negligent—that is, where a party fails to produce evidence that it *could have* cost-effectively produced—then there is no upper ceiling to the optimal penalty. Even reducing a plaintiff's damages to zero, or increasing a defendant's liability by billions, would not cause overdeterrence because the threat of a massive penalty would only incentivize compliance, and it would incentivize compliance only in cases where such compliance is socially beneficial. This idealized theory tends to support harsh bright-line penalties—such as reducing a plaintiff's award to zero or automatically doubling a defendant's liability—because they have little substantive downside while minimizing administrative cost.[257] Such a rationale may explain why traditional doctrine penalizes a plaintiff's failure to produce evidence with a zero award and has no

255. *See* Nance, *supra* note 30, at 627-28 (providing similar definition of when evidence is "unreasonably incomplete").

256. It is useful to clarify here that the cheaper cost-effective producer analysis is a marginal analysis, just as the negligence inquiry is a marginal inquiry. *See* LANDES & POSNER, *supra* note 14, at 87. That is, the question for a court is not which side is the cheaper producer of damages information *overall*, but which side is the cheaper producer of *additional* damages information beyond whatever the court already knows.

257. *Cf.* Gary S. Becker, *Crime and Punishment: An Economic Approach*, 76 J. POL. ECON. 169, 183-84 (1968) (presenting the economic model in which extremely severe punishments facilitate lower administrative costs in crime prevention).

allowance for intermediate amounts, even if one might suspect intermediate awards are common in practice.[258]

We do not live in an ideal world, and the possibility of judicial error in the determination of negligence means that it is possible to overdeter parties into spending too much on evidence collection. Given this, courts should generally try to set the penalty at a level to prevent both under- and overdeterrence—that is, the expected penalty should be just slightly greater than the private cost savings from failing to produce evidence,[259] such that a party will be better off producing the evidence. To the extent that precisely determining this optimal amount is administratively difficult, however, courts should generally err on the side of higher penalties because, as explained above, the concern about overdeterrence is limited.

Courts have not explicitly articulated the methodology that I describe here, but sometimes cases suggest a burden-shifting framework that works similarly. A good example is *Westinghouse Electric & Manufacturing Co. v. Wagner Electric & Manufacturing Co.*, a case dealing with the burden of proving the correct apportionment of profits for patent infringement.[260] The Supreme Court held that, if a patentee "has exhausted all available means of apportionment" but still cannot prove a reasonably precise number, then the patentee is entitled to the entire profit unless the defendant can prove the correct apportionment.[261] As the Court acknowledged, "[T]his is but another way of saying that the burden of proof is on the defendant."[262] But the Court emphasized that the burden is not shifted until "*after* the plaintiff has proved the existence of profits attributable to his invention and demonstrated that they are impossible of accurate or approximate apportionment."[263] Viewed from the cheaper cost-effective producer principle, this burden-shifting

258. *See* RICHARD A. POSNER, ECONOMIC ANALYSIS OF LAW § 6.4, at 220 (8th ed. 2011) (arguing that the common law adopted contributory negligence because it was administratively cheaper).

259. By "expected," I mean that courts should adjust for the possibility that a negligent party may not be caught. If a party has only a 25 percent chance of being caught, then, in cases where they are caught, the actual penalty must be multiplied fourfold. *See* A. Mitchell Polinsky & Steven Shavell, *Punitive Damages: An Economic Analysis*, 111 HARV. L. REV. 869, 874 (1998).

260. 225 U.S. 604, 614 (1912).

261. *Id.* at 621-22.

262. *Id.* at 622.

263. *Id.* (emphasis added).

framework reflects the two criteria I have described. The plaintiff bears the initial burden of proof, and must satisfy this burden of proof by producing all cost-effective evidence, that is, "exhaust[ ] all available means of apportionment."[264] Once this is done, then the burden shifts, until the court has the best evidence that can be cost-effectively obtained. At that point the work of an information-forcing mechanism (that is, a burden of proof) is done.[265]

CONCLUSION

The information-forcing dilemma pervades the law of damages. The basic way for a court to obtain better information on the value of a loss is by strictly enforcing a rigorous burden of proof. For this to actually work, however, there must be a credible penalty, and the basic penalty is an award of zero damages, which is unjust when the evidence clearly establishes there is *some* loss but not the precise *amount*. The injustice of the zero-damages penalty, in turn, leads courts to lower the burden of proof, which undermines the information-forcing effect.

The result of this dilemma is that damages law becomes a confusing mess. Courts acting without an understanding of the full nature of the problem will naturally oscillate between opposing approaches, depending on which side of the dilemma is more salient at a particular time and in a particular context. When courts face complaints about speculative awards, they tighten the burden of proof to force better information. When courts face the prospect of awarding zero damages to sympathetic injured plaintiffs, they loosen the burden of proof. The result is a self-contradictory doctrine that goes in circles.

Articulating the dilemma provides a way to both understand what ails damages law today and how to solve the problem. At least in principle, courts should allocate the burden of proof to the cheaper

264. *Id.* at 621.

265. One difference between *Westinghouse* and my proposed methodology is that the Court in *Westinghouse* would still penalize the defendant by awarding the plaintiff with the entire profit even if the defendant produces all cost-effective evidence (as long as the evidence does not suffice to determine the correct apportionment), *see id.* at 621-22, whereas my proposal would have a court take its best guess. The difference is largely immaterial from the perspective of economic incentives. *See* DOUGLAS G. BAIRD ET AL., GAME THEORY AND THE LAW 18 (6th prtg. 2003).

producer of information, and they should require the production of information only when it is cost-benefit-justified. This cheaper cost-effective producer principle may not be susceptible to precise implementation, but it provides a guiding lodestar to explain how damages law should be understood. Following this lodestar allows damages law to break out of the repetitive circle in which it has been trapped.