

THE LOGIC OF COMPARABLE-CASE GUIDANCE IN THE DETERMINATION OF AWARDS FOR PAIN AND SUFFERING AND PUNITIVE DAMAGES

Hillel J. Bavli*

Abstract

Little guidance is provided to triers of fact in arriving at awards for pain and suffering and punitive damages. Such awards are therefore highly variable. This article explains why methods involving comparable-case guidance—information regarding awards in comparable cases as guidance for determining damage awards—are effective not only in reducing unpredictability, but also in improving the accuracy of awards for pain and suffering and punitive damages generally. The article addresses major objections to such methods, and provides relevant legal context and direction for implementation.

I. INTRODUCTION

Lawmakers and scholars have struggled to address the unpredictability of awards for pain and suffering and punitive damages. Jurors are provided with very little guidance in determining such awards, and courts lack objective standards to guide jurors and review their awards. Consequently, awards can vary wildly.

For example, in a well-known case involving allegations that the auto manufacturer BMW had fraudulently failed to disclose to a purchaser that his newly purchased car had been repainted after suffering damage, a jury awarded the purchaser compensatory damages of \$4,000 and punitive damages of \$4 million.¹ However, in an identical case brought by another purchaser in the same court and before the same judge, *but with a different jury*, the jury awarded a similar level of compensatory damages, but no punitive damages at all, finding that BMW's behavior did not rise to the level of reprehensibility deserving of punitive damages.²

The awards in the BMW case, like others—such as a medical

* Harvard University. The views expressed in this article are those of the author, and not those of any organization with which he is affiliated. The author thanks Donald Rubin, David Rosenberg, Joseph Blitzstein, Kenneth Shepsle, and Yang Chen for their helpful comments, and John Kenneth Felter for his helpful edits.

1. See *BMW of N. Am., Inc. v. Gore*, 517 U.S. 559 (1996). The Alabama Supreme Court reduced the award to \$2 million, and the United States Supreme Court ultimately reversed, holding that even \$2 million in punitive damages was grossly excessive. *Id.* at 585–86.

2. See George L. Priest, *Introduction to CASS R. SUNSTEIN ET AL., PUNITIVE DAMAGES: HOW JURIES DECIDE* 2-3 (2002).

malpractice case resulting in a pain and suffering award of \$100 million³ or a deceptive-cigarette-marketing case resulting in a punitive damages award of \$28 billion⁴—represent a stark deviation from Justice Holmes’s characterization of the law as a “systematized prediction.”⁵ As Judge Niemeyer commented regarding the \$100 million award in *Evans v. St. Mary’s Hospital of Brooklyn*, “Because the jury was given no rule nor any rational criteria to apply in setting the amount of such an award, but told simply to do what it thought best, the jury responded with a perceived ‘measurement’ of the pain, which essentially amounted to an emotional response.”⁶ The jury award in *Evans*—an award ten times the amount requested by the plaintiff’s attorney, and one that the trial judge ultimately reduced to \$3.5 million—suggests, at least, that such awards are unpredictable.⁷ Regarding the award in *Bullock v. Philip Morris, Inc.*, Judge Niemeyer commented that “only emotion, not a rule of law, could justify imposing an award of \$28 billion. To confirm this, one need only consider the ‘mind’ of a legislative body developing a prospective rule of law to punish similar conduct.”⁸

Indeed, numerous empirical studies have confirmed substantial anecdotal evidence that awards for pain and suffering and punitive damages are highly unpredictable.⁹

3. See *Evans v. St. Mary’s Hosp. of Brooklyn*, 1 A.D.3d 314 (N.Y. App. Div. 2003).

4. See *Bullock v. Philip Morris Inc.*, No. BC249171, 2002 WL 31833905 (Cal. App. Dep’t Super. Ct. Dec. 18, 2002).

5. Oliver Wendell Holmes, *The Path of the Law*, 10 HARV. L. REV. 457, 458 (1897). Justice Holmes observed that “[p]eople want to know under what circumstances and how far they will run the risk of coming against what is so much stronger than themselves, and hence it becomes a business to find out when this danger is to be feared. The object of our study, then, is prediction, the prediction of the incidence of the public force through the instrumentality of the courts.” *Id.* at 457; see also Paul v. Niemeyer, *Awards for Pain and Suffering: The Irrational Centerpiece of our Tort System*, 90 VA. L. REV. 1401, 1402–04 (2004).

6. Niemeyer, *supra* note 5, at 1403.

7. *Id.* at 1403–04.

8. *Id.* at 1409–10 (noting that the “California legislature has fixed the maximum fine for false advertising at \$2500,” and that “Congress has fixed the maximum fine for corporations’ violations of federal offenses at \$500,000 or twice the defendant’s gain or the victim’s loss”).

9. See Shari Seidman Diamond et al., *Juror Judgments About Liability and Damages: Sources of Variability and Ways to Increase Consistency*, 48 DEPAUL L. REV. 301, 317 (1998) (concluding that there is “considerable variation in both juror and jury awards,” and that “[a] substantial portion of that variation is not predictable from measures of either background or attitudinal individual differences across jurors”); Randall R. Bovbjerg et al., *Valuing Life and Limb in Tort: Scheduling “Pain and Suffering,”* 83 NW. U. L. REV. 908, 919–24 (1989) (concluding that “[a]lthough the median, and even mean, awards in a given category may be considered relatively reasonable, the seemingly uncontrolled variability of awards is cause for concern—similar to anxiety about drowning in a pool averaging only two feet in depth”); Oscar G. Chase, *Helping Jurors Determine Pain and Suffering Awards*, 23 HOFSTRA L. REV. 763, 768–69 (1995) (citing studies); David W. Leebron, *Final Moments: Damages For Pain and Suffering Prior to Death*, 64 N.Y.U. L. REV. 256, 259 (1989) (concluding that “tort awards for even this relatively simple area [of damages for pain and suffering prior to death] vary significantly and that neither the specific facts of the case nor differing theoretical views of the functions of the

Commentators have proposed numerous methods to address unpredictability, including various forms of award schedules that bind or guide a jury in its award determination in light of its findings regarding certain facts, such as the severity of a plaintiff's injuries.¹⁰ Thus, "[s]chedules with categories based on injury severity typically provide the method of classification, and prior awards for injuries within each category provide a range of damages amounts."¹¹ That is, "[a]lthough reforms of this type differ in their details, each proceeds from the premise that prior pain-and-suffering awards for similar cases provide the appropriate basis for computing the present award."¹² Ultimately, "[t]he jury or reviewing court determines where the plaintiff's injury falls on the schedule, and the schedule provides a range or specified amount that can be binding or nonbinding on juries or courts."¹³

To the extent that these methods predetermine the award or range of awards, and to the extent that they bind the jury rather than guide it, they have been criticized as "eviscerat[ing] the various contributions that juries make to the civil justice system," and as being "inconsistent with the basic tort principle that each victim is entitled to an award tailored to his or her circumstances, set by a lay jury."¹⁴

A similar recommendation, proposed previously in various forms, involves "comparability analysis." Using this approach, the court (perhaps by way of an adversarial process involving the litigation parties, and even the trier of fact) would first identify a universe of comparable cases. It would then provide the trier of fact with certain information regarding the awards in these cases in the context of a jury instruction or as expert testimony, and it would instruct the trier of fact to arrive at a damages determination in light of the evidence introduced in the case, using the comparable-case information (or "prior-award information") as guidance.¹⁵ Such methods are based, in part, on

awards can explain such variation"); see also Richard Abel, *General Damages Are Incoherent, Incalculable, Incommensurable, and Inegalitarian (but Otherwise A Great Idea)*, 55 DEPAUL L. REV. 253, 291–303 (2006). But see Yun-chien Chang et al., *Pain and Suffering Damages in Personal Injury Cases: An Empirical Study* (Univ. of Chicago Coase-Sandor Inst. for Law & Econ., Research Paper No. 749, 2016), <http://papers.ssrn.com/abstract=2741180> (concluding that "pain and suffering damages in Taiwan are to a large extent statistically and legally predictable").

10. See *infra* Part III.

11. Mark Geistfeld, *Placing A Price on Pain and Suffering: A Method for Helping Juries Determine Tort Damages for Nonmonetary Injuries*, 83 CAL. L. REV. 773, 791 (1995).

12. *Id.*

13. *Id.*

14. David A. Logan, *Juries, Judges, and the Politics of Tort Reform*, 83 U. CIN. L. REV. 903, 942–43 (2015).

15. See *infra* Part III. Note that there is substantial overlap between such methods and those involving scheduling—particularly to the extent that scheduling methodologies involve providing a trier of fact with scenarios and associated award values, based on comparable cases, as guidance in

empirical studies confirming that they are effective in controlling outlying awards and reducing award variability generally, even using prior awards to *guide*, rather than bind, the trier of fact.¹⁶

For purposes of clarity—because the terms “damage schedule” and “comparability analysis” have been used in the literature to mean various things in various contexts—I will use the term “comparable-case guidance” (CCG) to refer to scheduling or comparability-analysis methods that fulfill three fundamental requirements: 1) information used as guidance must be derived from *prior* “comparable” cases (as opposed to, e.g., damage schedules predetermined arbitrarily by a legislative body); 2) comparable-case information must be considered by the *trier of fact* in particular (as opposed to, e.g., a reviewing court); and 3) comparable-case information must be used as *guidance* only (as opposed to, e.g., imposing a range or amount that is binding on the trier of fact).¹⁷

While CCG methods allow for a case-by-case analysis that remains in the discretion of the trier of fact, the use of prior awards itself has been attacked based on three fundamental objections. First, “[i]f the system has been providing overly arbitrary pain-and-suffering awards, and if we have no method for determining the appropriate award in the first instance, why should we make prior awards the cornerstone of future awards,” since “[b]y doing so, we may ensure that like cases are treated alike in that all involve inappropriate damages awards.”¹⁸ Second, such methods “fail to address the fundamental issue of how one should initially assess the value of pain-and-suffering damages” or arrive at an appropriate punitive damages award.¹⁹ Third, the validity of such

determining damages. See also Michael J. Saks et al., *Reducing Variability in Civil Jury Awards*, 21 LAW AND HUMAN BEHAVIOR 243, 246 (1997); Bovbjerg et al., *supra* note 9, at 953.

16. See, e.g., Saks et al., *supra* note 15, at 249–55.

17. Note that the analysis herein may apply to scheduling and comparability analysis more broadly; however, I focus on CCG methods in particular.

18. Geistfeld, *supra* note 11, at 792 (commenting that “[t]his reliance upon past awards . . . represents the most problematic aspect of these reform proposals”); see Peter H. Schuck, *Scheduled Damages and Insurance Contracts for Future Services: A Comment on Blumstein, Bovbjerg, and Sloan*, 8 YALE J. ON REG. 213, 218 (1991) (“by using earlier awards as the foundation for their new system of damages scheduling, they impound and then compound what they themselves characterize as the distortions of the past, thereby projecting those distortions into the future”); see also Robert L. Rabin, *The Pervasive Role of Uncertainty in Tort Law: Rights and Remedies*, 60 DEPAUL L. REV. 431, 448–49 (2011) (“Scheduling proposals of this kind have been criticized by torts scholars such as Mark Geistfeld, who pointed out the seeming paradox in rejecting unstructured jury decision making in favor of a scheduled approach, which from a horizontal equity perspective takes arbitrary prior awards as the cornerstone for future awards, and from a vertical equity perspective takes the ordering of magnitude in past jury awards as an appropriate key for hierarchical sorting in the designated severity-categories for future awards.”).

19. Ronen Avraham, *Putting A Price on Pain-and-Suffering Damages: A Critique of the Current Approaches and A Preliminary Proposal for Change*, 100 NW. U. L. REV. 87, 104 (2006) (discussing pain and suffering awards in particular, and citing W. Kip Viscusi, *Pain and Suffering: Damages in*

methods relies heavily on the presumption that a “correct” set of cases has been identified—that the prior cases identified are indeed materially comparable to the case at hand.²⁰

These objections boil down to a fundamental concern: if our problem is the unpredictability of awards caused by a trier of fact’s inability to assess objectively the appropriate value of awards for pain and suffering or punitive damages, how is it beneficial to provide a trier of fact with information regarding damages awarded in prior cases that are separate and distinct from the present case and that presumably suffer from the same arbitrariness that we wish to address in the present case? Seemingly, this would only compound the problem.

My aim in this article is to address these objections by explaining in simple but formal terms how, notwithstanding the foregoing objections, CCG methods not only reduce unpredictability but improve the accuracy of awards for pain and suffering and punitive damages generally by allowing for the sharing of relevant information across cases.

Part II discusses the importance of reducing the variability of awards for pain and suffering and punitive damages. Part III provides background regarding current methods and proposals for reducing award variability. Part IV explains how CCG methods reduce unpredictability and improve the accuracy of awards for pain and suffering and punitive damages generally. Part V discusses a number of considerations for identifying a universe of prior cases and for distilling information from such cases for consideration by the trier of fact. Part VI provides additional legal context for CCG methods and concludes.

II. THE IMPORTANCE OF REDUCING THE VARIABILITY OF AWARDS FOR PAIN AND SUFFERING AND PUNITIVE DAMAGES

The Supreme Court has repeatedly emphasized the importance of maintaining fair and consistent awards for pain and suffering and punitive damages. In a recent case, for example, the Court stated:

The real problem, it seems, is the stark unpredictability of punitive awards. Courts of law are concerned with fairness as consistency, and evidence that the median ratio of punitive to compensatory

Search of A Sounder Rationale, 1 MICH. L. & POL’Y REV. 141, 168 (1996) (“The key issue to be resolve[d] for pain and suffering schedules and scales is that establishing such arbitrary benchmarks does not resolve the more fundamental issue of how one should initially assess the value of pain and suffering damages.”).

20. See Logan, *supra* note 14, at 943–44 (“While [an approach allowing the fact finder to consider a range of possible awards for guidance] would improve predictability, such an approach would only be as good as the quality of the methodology for selecting which cases were factually similar enough to be included in the range.”).

awards falls within a reasonable zone, or that punitive awards are infrequent, fails to tell us whether the spread between high and low individual awards is acceptable. The available data suggest it is not.²¹

The Supreme Court continued by discussing the “inherent uncertainty of the trial process” and the resulting inconsistency among awards in cases with similar facts.²² In examining unpredictability as a matter of policy, rather than of constitutional significance,²³ the Court emphasized that the unpredictability of high punitive damage awards is “*in tension with the function of the awards as punitive.*”²⁴ It commented:

Thus, a penalty should be reasonably predictable in its severity, so that even Justice Holmes’s “bad man” can look ahead with some ability to know what the stakes are in choosing one course of action or another. See *The Path of the Law*, 10 Harv. L. Rev. 457, 459 (1897). And when the bad man’s counterparts turn up from time to time, the penalty scheme they face ought to threaten them with a fair probability of suffering in like degree when they wreak like damage. Cf. *Koon v. United States*, 518 U.S. 81, 113 (1996) (noting the need “to reduce unjustified disparities” in criminal sentencing “and so reach toward the evenhandedness and neutrality that are the distinguishing marks of any principled system of justice”). The common sense of justice would surely bar penalties that reasonable people would think excessive for the harm caused in the circumstances.²⁵

The Supreme Court thus emphasized the point that predictability,

21. *Exxon Shipping Co. v. Baker*, 554 U.S. 471, 499 (2008) (citing Theodore Eisenberg et al., *Juries, Judges, and Punitive Damages: Empirical Analyses Using the Civil Justice Survey of State Courts 1992, 1996, and 2001 Data*, 3 J. OF EMPIRICAL LEGAL STUD. 263, 267 (2006); THOMAS H. COHEN, U.S. DEPT. OF JUSTICE, PUNITIVE DAMAGE AWARDS IN LARGE COUNTIES, 2001, at 5 (Mar. 2005), <http://www.bjs.gov/content/pub/pdf/pdalc01.pdf>; Brian J. Ostrom et al., *A Step Above Anecdote: A Profile of the Civil Jury in the 1990s*, 79 JUDICATURE 233, 240 (1996); Erik K. Moller et al., *Punitive Damages in Financial Injury Jury Verdicts*, 28 J. LEGAL STUD. 283, 333 (1999)).

22. *Id.* at 500–01 (quoting *BMW of N. Am., Inc. v. Gore*, 646 So.2d 619, 626 (Ala. 1994)).

23. The Supreme Court highlighted that “the Court’s response to outlier punitive-damages awards has thus far been confined by claims at the constitutional level, and our cases have announced due process standards that every award must pass.” *Id.* at 501. In *Baker*, however, the Court examined a jury award for punitive damages “for conformity with maritime law, rather than the outer limit allowed by due process.” *Id.* at 501–02. In acting “in the position of a common law court of last review,” *id.* at 507, it considered punitive damages not with respect to “their intersection with the Constitution,” but rather in relation to the “desirability of regulating them as a common law remedy,” *id.* at 502.

24. *Id.* at 502 (emphasis added).

25. *Id.* at 502–03.

consistency, and fairness are fundamental to the deterrence objectives underlying punitive damages. This sentiment and others associated with the harms of unpredictability have been echoed repeatedly by lower courts and scholars.

For example, in the case *Payne v. Jones*, the Second Circuit explained the purposes underlying the device of remittitur by emphasizing the harmful effects of variability and outlying awards, including those associated with overdeterrence:

Apart from impairing the fairness, predictability and proportionality of the legal system, judgments awarding unreasonable amounts as damages impose harmful, burdensome costs on society. As an initial matter, an excessive verdict that is allowed to stand establishes a precedent for excessive awards in later cases. The publicity that accompanies huge punitive damages awards, *see, e.g.*, Henry Weinstein, *Philip Morris Ordered to Pay \$28 Billion to Smoker*, L.A. Times, Oct. 5, 2002, will encourage future jurors to impose similarly large amounts. Unchecked awards levied against significant industries can cause serious harm to the national economy. Productive companies can be forced into bankruptcy or out of business. Municipalities can be drained of essential public resources. The threat of excessive damages, furthermore, drives up the cost of insurance premiums, deters both individuals and enterprises from undertaking socially desirable activities and risks, and encourages overspending on “socially excessive precautions” that “cost[] more than the reduction of harm produced by [them].” A. Mitchell Polinsky & Steven Shavell, *Punitive Damages: An Economic Analysis*, 111 Harv. L. Rev. 869, 879 (1998). The prices of goods and services will rise, and innovation will be inhibited. *See id.* at 873.²⁶

Courts have similarly emphasized the need to “minimize the arbitrary variance in awards bound to result from [the] throw-up-the-hands approach” that courts regularly use in determining awards for pain and suffering.²⁷ Courts and scholars have recognized the need to address this “standardless, unguided exercise of discretion by the trier of fact, reviewable . . . pursuant to no standard to guide the reviewing court either.”²⁸

In *Geressy v. Digital Equipment Corp.*, Judge Weinstein explained the court’s decision to consider prior-award information. He discussed

26. *Payne v. Jones*, 711 F.3d 85, 94 (2d Cir. 2013).

27. *Jutzi-Johnson v. United States*, 263 F.3d 753, 759 (7th Cir. 2001).

28. *Id.*

concern by the courts and legislature regarding the “virtually unbridled discretion” of juries in awarding damages for pain and suffering, for which there is “currently no meaningful way to measure such non-quantifiable losses monetarily.”²⁹ As Professor Oscar Chase explained:

Variability is a problem primarily because it undermines the legal system’s claim that like cases will be treated alike; the promise of equal justice under law is an important justification for our legal system. Variability is also claimed to create instrumental defects; that is, it makes it harder to settle cases, thus adding unnecessary transaction costs to the tort system, and delaying payment to needful plaintiffs. Unpredictability also leads to inefficiencies because of over- or under-precautions by affected industries and insurers.³⁰

Thus, whether for purposes of fairness, deterrence, or another objective, courts recognize the importance of generating consistent and predictable damage awards. Indeed, reducing variability is fundamental to achieving accurate legal outcomes. In Part IV, I consider variability in the context of error generally, and discuss in more detail what is meant by “accurate legal outcomes.” First, however, I describe current methods and proposals for addressing award variability.

III. CURRENT METHODS AND PROPOSALS FOR ADDRESSING AWARD VARIABILITY

Additur and Remittitur. Courts use the procedural devices of additur and remittitur to increase or decrease an award found to be insufficient or excessive. For example, a defendant may argue for a new trial based on the excessiveness of the jury’s award. If the judge agrees, he may offer the plaintiff to reduce the award (remittitur) by some amount rather than proceeding with a new trial.³¹ Appellate courts may also modify awards and address challenges to additur and remittitur.³²

Although the devices of additur and remittitur can, *in theory*, address

29. *Geressy v. Dig. Equip. Corp.*, 980 F. Supp. 640, 656 (E.D.N.Y. 1997) (quoting Leslie A. Rubin, *Confronting a New Obstacle to Reproductive Choice: Encouraging the Development of RU-486 Through Reform of Products Liability Law*, 18 N.Y.U. REV. L. & SOC. CHANGE 131, 146 (1990–91)).

30. Chase, *supra* note 9, at 769.

31. David Baldus et al., *Improving Judicial Oversight of Jury Damages Assessments: A Proposal for the Comparative Additur/Remittitur Review of Awards for Nonpecuniary Harms and Punitive Damages*, 80 IOWA L. REV. 1109, 1118–20 (1995).

32. *Id.* Note, the Supreme Court has distinguished additur from remittitur and held that the former procedure is violative of the defendant’s Seventh Amendment right to a trial by jury. See *Dimick v. Schiedt*, 293 U.S. 474 (1934); see also Baldus et al., *supra* note 31, at 1119–20.

the high levels of variability associated with awards for pain and suffering and punitive damages, in practice they cannot. First, additur and remittitur are primarily used to address excessive awards, and are infrequently used to adjust inadequate awards.³³ Second, these procedures are generally reserved for only the most extreme cases—they are used to address extreme outliers that are deemed incorrect, rather than address variability in general.³⁴ Third, courts lack consistent and principled procedures for arriving at better award determinations than juries. As Judge Posner opined in *Jutzi-Johnson v. United States*, “[most courts] treat the determination of how much damages for pain and suffering to award as a standardless, unguided exercise of discretion by the trier of fact, reviewable for abuse of discretion pursuant to no standard to guide the reviewing court either.”³⁵ Further, courts employ these devices infrequently;³⁶ and widespread replacement of jury awards with judicial determinations would be problematic with respect to the Seventh Amendment and fundamental principles of tort law.

Damage Caps. Similarly, *damage caps*, which place upper limits on damage awards or certain types of damage awards, are widely recognized as a particularly poor method for addressing variability.³⁷

33. See Baldus et al., *supra* note 31, at 1119–20.

34. *See id.* at 1118–20.

35. *Jutzi-Johnson v. United States*, 263 F.3d 753, 759 (7th Cir. 2001).

36. *See* Baldus et al., *supra* note 31, at 1119–21 (citing MICHAEL G. SHANLEY & MARK A. PETERSON, RAND: THE INST. FOR CIVIL JUSTICE, POSTTRIAL ADJUSTMENTS TO JURY AWARDS vii–viii, 43–47 (1987), <http://www.rand.org/content/dam/rand/pubs/reports/2007/R3511.pdf>, and others). Baldus et al. note that “[o]n the basis of post-trial information in the 880 1982–84 California and Cook County, Illinois cases reported in [the Shanley and Peterson study], we estimate that in cases involving a damages award, there was a remittitur or new trial 6% of the time and an additur or a negotiated increase in the award in 2–3% of the cases.” *Id.* at 1120 n.21. *See also* Joseph Sanders, *Why Do Proposals Designed to Control Variability in General Damages (Generally) Fall on Deaf Ears? (and Why This Is Too Bad)*, 55 DEPAUL L. REV. 489, 503 (2006).

37. Damage caps have served as the primary tool of tort reform aiming to address variability. Legislation for caps boomed during heightened calls for tort reform in the 1970s and 1980s, and by 1987, twenty-three states had instituted caps ranging from \$150,000 to \$1,000,000. Sanders, *supra* note 36, at 510; *see also* Stephen D. Sugarman, *A Comparative Law Look at Pain and Suffering Awards*, 55 DEPAUL L. REV. 399, 399–400 (2006). Since then, numerous other states have adopted caps in various contexts as well. Sanders, *supra* note 36, at 510. Statutory caps may be applied to specific types of damages, such as punitive damages or general damages, or damage awards generally. *See* Baldus et al., *supra* note 31, at 1121–23. They may be applied in particular contexts, such as medical malpractice. *See generally* Kathryn Zeiler, *Turning from Damage Caps to Information Disclosure: An Alternative to Tort Reform*, 5 YALE J. HEALTH POL’Y, L. & ETHICS 385 (2005). And they can appear in a variety of forms. For example, juries can be informed of damage caps explicitly, or a cap can be applied by the court only after a jury has exceeded it. *See generally* Saks et al., *supra* note 15, at 245. In general, juries are not informed of caps; but, regardless, jurors may be aware of such measures. *See id.* Under some proposals, a judge may be permitted to reduce awards below the quantity provided by the legislation; in others, a judge may be permitted only to reduce an award to the established quantity. *See* Colleen P. Murphy, *Statutory Caps and Judicial Review of Damages*, 39 AKRON L. REV. 1001, 1002–08 (2006).

They give rise to a range of problems inherent in capping damages (or certain types of damages) outright, independent of the circumstances of the case or harm suffered by the plaintiff.

First, caps address only the most extreme cases, and only excessive awards. Second, caps are likely to play a role primarily in cases in which the harm is very severe. If, for example, there is a cap of \$1 million, the cap may not affect an individual who suffered a broken nose, but may greatly affect an individual who suffered paralysis. The cap is unlikely to address variability in the former case, while possibly causing an inappropriately low award in the latter.³⁸ Thus, aside from biasing the award, caps distort incentives to litigate tort claims based on severe harm. Further, caps can cause suboptimal risk-taking by diluting the deterrence effect of punitive damages and tort law generally.³⁹ Third, a number of studies have concluded that caps may in fact exacerbate the variability problem rather than address it. In particular, jurors who are aware of a cap may anchor to it, and their awards may gravitate to it. “Thus, those with the most severe injuries and losses will be unfairly deprived of compensation by caps. In addition, by serving as a psychological anchor, caps appear likely to further exaggerate the error of overcompensating those with smaller losses.”⁴⁰

But, notwithstanding the foregoing problems, damage caps are used, due, in part, to their administrative convenience and their ability to control insurance rates.⁴¹

Comparable-Case Information. Finally, a number of courts and commentators have proposed methods involving the use of information regarding awards in prior comparable cases. Such proposals have appeared in various forms. One set of approaches builds on courts’ regular use of comparability *review*, whereby trial and appellate courts consider awards in comparable cases in their review of a trier of fact’s award for excessiveness.⁴² These proposals attempt to develop a more structured framework for the court’s comparative analysis—for example, instituting a method by which a court identifies a universe of comparable claims; examines whether the current award “deviates materially from what would be reasonable compensation” (to use language from a New York statute upon which such proposals seem to

38. See Baldus et al., *supra* note 31, at 1121–23.

39. See Sanders, *supra* note 36, at 509–11. Constitutionality issues may arise as well. Opponents of caps have argued, for example, that caps are violative of a plaintiff’s right to a trial by jury. See Zeiler, *supra* note 37, at 387.

40. Saks et al., *supra* note 15, at 245. See also Verlin B. Hinsz & Kristin E. Indahl, *Assimilation to Anchors for Damage Awards in a Mock Civil Trial*, 25 J. OF APPLIED SOC. PSYCHOL. 991 (1995).

41. See Sanders, *supra* note 36, at 511.

42. See Baldus et al., *supra* note 31, at 1134.

build);⁴³ and, if so, determines a suitable adjustment of the jury award.⁴⁴

These review-based methods have the advantage of involving only modest modifications to current well-accepted methods, but they are not ideal. Specifically, they incorporate comparable-case information on review rather than in the award directly. As such, they only focus on outlying awards rather than on variability in general. Moreover, regular use of such methods arguably leads to constitutionality and tort issues associated with the replacement of a jury's discretion with that of the judge.⁴⁵

As indicated in the Introduction, another set of proposals uses various types of schedules to bind or guide the trier of fact in its determination of certain types of damage awards.⁴⁶ These proposals, to the extent they involve binding a trier of fact to predetermined values or ranges of values, have been criticized as impinging on the trier of fact's discretion, and as "inconsistent with the basic tort principle that each victim is entitled to an award tailored to his or her circumstances, set by a lay jury."⁴⁷ To the extent that such methods bind rather than guide the trier of fact, and to the extent that they involve award values that are predetermined arbitrarily, for example, by a legislative body in advance of the case, such methods are arguably problematic.

Finally, a number of proposals—overlapping with the category above regarding schedules—involve comparability analysis, whereby the court (perhaps by way of an adversarial process involving the litigation parties, and even the trier of fact) identifies a set of materially comparable cases, provides the trier of fact with certain information regarding the awards in these cases in the context of a jury instruction or presented as expert testimony, and instructs the trier of fact to determine an appropriate damages award in light of the evidence introduced in the case and the comparable-case information.⁴⁸

43. N.Y. C.P.L.R. 5501(c) (McKinney 1995).

44. See Baldus et al., *supra* note 31, at 1134. See also Sanders, *supra* note 36, at 503.

45. See Colleen P. Murphy, *Judicial Assessment of Legal Remedies*, 94 NW. U. L. REV. 153, 192 (1999); J. Patrick Elsevier, *Out-of-Line: Federal Courts Using Comparability to Review Damage Awards*, 33 GA. L. REV. 243, 258–59 (1998) ("Comparability analysis [on review] requires a court to engage in several subjective, fact-intensive inquiries Yet, by asserting that comparability provides an objective framework to review compensatory damage awards, courts are able to substitute their own assessment of what the appropriate damage award should be, while skirting the Seventh Amendment's proscription against reexamination of issues of fact.").

46. See, e.g., Bovbjerg et al., *supra* note 9, at 945.

47. Logan, *supra* note 14, at 943.

48. See, e.g., Chase, *supra* note 9, at 777–90 (proposing that courts inform jurors of "the range of awards made by other juries in the same state for such damages during a contemporaneous time period," to be provided as nonbinding guidance in the court's charge to the jury in the form of a "chart constructed to allow comparison with roughly similar cases in which plaintiffs' verdicts were recovered."). See generally Baldus et al., *supra* note 31, at 1123–24; Saks et al., *supra* note 15, at 246; Sanders, *supra* note 36, at 506–07; JACOB A. STEIN, STEIN ON PERSONAL INJURY DAMAGES TREATISE

As suggested in the Introduction, such methods—and CCG methods in particular—avoid many of the problems caused by other methods, such as damage caps and binding or predetermined schedules. These methods allow for a flexible case-by-case analysis in the discretion of the trier of fact, while still providing substantial guidance. Nevertheless, such methods have been attacked on the grounds that, instead of addressing the fundamental issue of assessing the value of pain and suffering damages or an appropriate punitive damages award, they may *compound* the unpredictability of damage awards by providing the trier of fact with information regarding damages awarded in separate and distinct prior cases that presumably suffer from the same arbitrariness that the court aims to address in the present case.⁴⁹ In the following parts, I address these objections.

IV. THE LOGIC OF CCG

In this part, I explain why, notwithstanding the objections delineated above, CCG methods reduce unpredictability *and* improve the accuracy of awards for pain and suffering and punitive damages generally. I begin by discussing what is meant by “accuracy.”

A. Bias and Variance: The Elements of Error

Part II makes clear that controlling the variability of awards for pain and suffering and punitive damages is important. But reducing variability does not necessarily mean improving the award. For example, it is likely unwise to encourage jurors to anchor to an *arbitrary* value, notwithstanding associated variability benefits. In fact, variability could be zero if the court were to dictate an award without regard to the particulars of the case.

Courts and scholars are correct to be concerned about the high variability associated with awards for pain and suffering and punitive damages. But good policy requires steps toward reducing such variability only insofar as they improve awards generally.

Thus, for purposes of considering the concepts of “error” and “accuracy” in the context of damage awards, assume that there is some “correct” award that could, in theory, be determined as a function of perfect information regarding the state of the world, including the facts

§ 8:32 (3d ed. 2015) (discussing AMERICAN BAR ASSOC., REPORT OF THE ACTION COMMISSION TO IMPROVE THE TORT LIABILITY SYSTEM 10–15 (1987)); *Jutzi-Johnson v. United States*, 263 F.3d 753, 758–61 (7th Cir. 2001).

49. See Geistfeld, *supra* note 11, at 792; Avraham, *supra* note 19, at 104; Logan, *supra* note 14, at 943–44.

of the case and applicable law and norms.⁵⁰ Of course, we neither have perfect information nor know the correct award. Instead, the court asks a jury to arrive at an award, which will serve as an *estimate* of the correct outcome.

More concretely, let us consider the correct award in a given case—say, for example, the BMW case described in the Introduction—to be the mean of the population of possible awards that would emerge from adjudicating the case repeatedly under various conditions (e.g., before different judges and juries, by different attorneys, with different permutations of facts, etc.).⁵¹ A single trial thus generates a sample from the population and an *estimate* of the correct award.⁵² Call the actual award an “estimate,” and the process that generates the estimate an “estimator.”⁵³ We can then define “error” in terms of distance, and (equivalently) “accuracy” in terms of proximity, between the estimate and the correct award.⁵⁴

I define the *reliability* of a legal procedure as the accuracy of the outcome (here, the award) that we can expect by following the procedure. If we expect that a certain legal procedure will produce an accurate award, we say that the procedure (or, for simplicity, the award) is “reliable.”⁵⁵

50. See Hillel J. Bavli, *Aggregating for Accuracy: A Closer Look at Sampling and Accuracy in Class Action Litigation*, 14 LAW PROBABILITY & RISK 67, 74–78 (2015). We can similarly consider a *range* or *distribution* of “correct” awards that reflects, for example, uncertainty regarding the law. *Id.* at 74 n.24.

51. *Id.* at 74–78 (citing Michael J. Saks & Peter David Blanck, *Justice Improved: The Unrecognized Benefits of Aggregation and Sampling in the Trial of Mass Torts*, 44 STAN. L. REV. 815, 833–34 (1992)). For simplicity, I ignore the potential for bias. As noted in Bavli, *supra* note 50, at 74 n.29, “[i]t may be more intuitive to consider the concept of a ‘correct’ verdict in the context of a criminal trial. Consider, for example, the O.J. Simpson murder trial. Polls show that over 50% of Americans believe that the jury arrived at an incorrect verdict. Implicit in the public’s disagreement with the verdict is an assumption that there exists a ‘correct’ outcome. The framework described above is intuitive: had the jury known the true facts of the case, and had there been no ambiguity regarding the application of the law to the facts of the case, the jury would have arrived at the correct conclusion. But given ambiguity regarding either the facts of a case or the state of the law, it is unclear whether a criminal defendant in fact satisfied the elements of the crime charged; and the jury must arrive at a verdict—‘guilty’ or ‘not guilty’—which serves as an estimate of the correct outcome, ‘guilty’ or ‘not guilty.’” Similarly, the determination of a civil damages award (or, e.g., a prison sentence) can be understood as an estimate of a correct outcome. Arguably courts implicitly acknowledge this characterization when, for example, a court finds a jury award to be excessive or inadequate. Finally, the formulation of the correct award as the mean of the population of possible awards from repeated adjudications is intended to capture, e.g., the state of the law as understood by various judges and the norms of the time and facts of the case as understood by various combinations of jurors.

52. See *id.* at 74–75; Saks & Blanck, *supra* note 51, at 833–34.

53. The estimator, in the current context, can be understood as the procedure (broadly speaking) that generates the legal outcome, which, in our example, is the jury’s punitive damages award. Mathematically, it is a random variable that, once “realized,” or decided, becomes an estimate.

54. See Bavli, *supra* note 50, at 74–78.

55. See generally *id.*

In statistics there are two sources of error: bias and variance. If an estimator is “unbiased,” then it will generate the correct value *on average*. If it is “biased,” then it will generate the incorrect value on average, and the “bias” reflects the distance between the value the estimator generates on average and the correct value. Bias is therefore a source of error. Note that it is possible that an unbiased estimator will never generate the correct value—to be unbiased is only to generate the correct value *in expectation*, or *on average*. Additionally, although unbiasedness is generally understood as a good characteristic for an estimator to have, it does not indicate lack of error, since the values generated by the estimator can vary wildly around the correct value. For example, if the correct punitive damages value in the BMW case above is \$100,000, then repetitions of an unbiased adjudication may generate estimate values (i.e., damage awards) of \$0, \$50,000, \$150,000, and \$200,000, which are indeed centered at the correct value of \$100,000; however, the awards are highly dispersed around \$100,000. We would, for example, prefer that repeated adjudications generate the values \$90,000, \$95,000, \$105,000, and \$110,000; or even better, \$100,000, \$100,000, \$100,000, and \$100,000.

Thus, the second source of error is “variance,” which is a measure of dispersion. In particular, if an estimator entails a high level of variance, then it will generate estimates that are highly dispersed around its mean, or average, value. If the estimator entails a high level of variance, but is unbiased, then it will generate estimates that are highly dispersed around the *correct value*. In this case, we say that the estimator is “unbiased,” but that it lacks “precision.” If the estimator is “precise” but “biased,” then it generates values that are tightly centered around the wrong value—an undesirable circumstance. If an estimator is “precise” and “unbiased,” then it will generate estimates that are close in proximity to the correct value, and we say that it is “accurate.”

Thus, in the BMW example above, the awards \$90,000, \$95,000, \$105,000, and \$110,000 reflect greater precision than \$0, \$50,000, \$150,000, and \$200,000. And, the awards \$100,000, \$100,000, \$100,000, and \$100,000 reflect even greater precision.

More formally, let α be the “correct” award and $\hat{\alpha}$ a random variable defined by the actual award. Let $E(\hat{\alpha})$ be the expectation of $\hat{\alpha}$. In other words, $\hat{\alpha}$ will equal $E(\hat{\alpha})$ *on average*; and if we repeat the trial many times then the average of the punitive damage awards will be approximately $E(\hat{\alpha})$.

Bias is defined as the difference between the expectation of the estimator $\hat{\alpha}$ and the correct award α (the object we are trying to estimate). More formally, $Bias = E(\hat{\alpha}) - \alpha$, and it is said that the estimator $\hat{\alpha}$ is “unbiased” if $E(\hat{\alpha}) = \alpha$.

Additionally, let $V(\hat{\alpha})$ be the variance of $\hat{\alpha}$, where variance is defined in the standard way. That is,

$$V(\hat{\alpha}) = E(\hat{\alpha} - E(\hat{\alpha}))^2 = \frac{1}{n} \sum_{i=1}^n (\hat{\alpha}_i - E(\hat{\alpha}))^2.$$

The variance of the realized awards is thus the average of the square differences between the awards and the average of the awards. For example, the variance of \$0, \$50,000, \$150,000, and \$200,000 is $\frac{1}{4}((\$0 - \$100,000)^2 + (\$50,000 - \$100,000)^2 + (\$150,000 - \$100,000)^2 + (\$200,000 - \$100,000)^2)$. The *standard deviation* of $\hat{\alpha}$ is the square root of the variance: $SD(\hat{\alpha}) = \sqrt{V(\hat{\alpha})}$.

Now let us formally define error in terms of bias and variance. There are many ways to measure error (just as there are many ways to measure, for example, dispersion). For instance, we might define error by the raw differences between the estimates and the correct value. But just as it is inconvenient for scientists to define dispersion by calculating the raw differences between the estimates and the mean of the estimates,⁵⁶ similar inconveniences arise from defining error in terms of raw differences. It turns out that a convenient and often suitable measure of error is the well-accepted *mean squared error (MSE)*, defined as $MSE(\hat{\alpha}) = E[(\hat{\alpha} - \alpha)^2]$. In words, we find the expectation of the square difference between the estimator and the correct value.

Now, using *MSE* as our measure of error, it is easy to show that:

$$MSE(\hat{\alpha}) = E[(\hat{\alpha} - \alpha)^2] = \text{Variance} + \text{Bias}^2$$

That is, error (defined here as MSE) can be separated into two elements: bias and variance.⁵⁷

B. Reducing Variability and Error with Prior-Award Information

Let us view a civil case as an estimation problem within the framework described above. That is, for every claim let us assume that there is a correct damages award (or, in particular, a correct award for pain and suffering or punitive damages) α , defined as the mean of the population of awards that would result from repeated adjudications, as described above (and that would reflect, for example, perfect

56. For example, using such differences results in the “canceling” of positive and negative values.

57. See SHARON L. LOHR, *SAMPLING: DESIGN AND ANALYSIS*, § 2.2 (2d ed. 2010), for a mathematical proof, and a lengthier discussion, of the bias-variance decomposition.

information regarding the applicable law and the facts of the case).⁵⁸ As discussed, we cannot know the correct award and must therefore arrive at a suitable *estimate* $\hat{\alpha}$ of the unknown correct award.

Thus, let us assume (for simplicity rather than necessity) that, on average, the estimate will equal the correct award, but that it entails some degree of variability σ^2 around the correct award. That is, $\hat{\alpha}$ is distributed with mean α and variance σ^2 . Notationally, $\hat{\alpha} \sim (\alpha, \sigma^2)$.

Our goal is to arrive at a damages estimate that is close to the correct award—that is “accurate” in the sense of minimizing error, as defined above by $MSE(\hat{\alpha}) = E[(\hat{\alpha} - \alpha)^2]$.

In statistics, and the sciences generally, determining a good estimator for an unknown quantity is very important and is the subject of much research. Let us use an analogy to better understand the problem and possible solutions. Consider a study by statisticians Bradley Efron and Carl Morris, in which they recorded batting average data for each of eighteen major league baseball players through his first forty-five official at bats of the 1970 season.⁵⁹ Suppose our goal is (as their goal was) to estimate the batting averages that would emerge for each player during the remainder of the season, using only the data collected. Let us focus on one player, Roberto Clemente, whose batting average for his first forty-five at bats was .400.⁶⁰

Initially, it might seem elementary that a player’s batting average for the first forty-five at bats would serve as the best estimator for his batting average for the remainder of the season. Thus, our estimate of Clemente’s remainder-of-the-season (or “after-forty-five”) batting average would be .400. But it can be shown mathematically and empirically that this is not in fact the best estimator. In particular, this estimator—Clemente’s batting average for his first forty-five at bats—can be improved by incorporating information regarding the batting averages for the first forty-five at bats of the other seventeen players.⁶¹ But why should this be? Why should the batting averages of Munson or Kessinger, who had “first-forty-five” batting averages of .156 and .289, respectively, have any influence on our estimate of Clemente’s after-forty-five batting average? Using other players’ first-forty-five batting averages would seem obvious in the absence of Clemente’s first-forty-five average; but we *have* Clemente’s first-forty-five batting average—

58. See Bavli, *supra* note 50, at 74–78. We can define the correct award using other measures of central tendency, such as the median, depending on how we want to characterize it. The mean, for example, is more sensitive than the median to extreme awards. For simplicity, the analysis in this article focuses only on the mean.

59. See Bradley Efron & Carl Morris, *Data Analysis Using Stein’s Estimator and Its Generalizations*, 70 J. AM. STATISTICAL ASS’N 311, 312–14 (1975).

60. *Id.* at 313.

61. See *id.* at 312–14.

and what could be better than that?

The explanation lies in the information we obtain from the first-forty-five batting averages of the other seventeen players. First, suppose we know nothing about batting averages. It seems intuitive that, in estimating Clemente's after-forty-five batting average, it is helpful to know whether his first-forty-five batting average, .400, is low, high, or average. In particular, it seems helpful to know that .400 is an extremely high batting average. As an extreme example, if we are told that, in a given period, a particular batting average is the highest ever achieved by that player, or in baseball in general, would we still use this average as our best estimate of the player's average in the next period? We would be wise to incorporate this information, since the player is unlikely to break baseball's all-time record twice in a row. More specifically, the usefulness of other-player information can be explained, in part, by considering the old idea (due to Sir Francis Galton) of "regression toward the mean," which, in its simplest form, states that if we obtain a relatively extreme measurement, then upon remeasurement, we are likely to obtain a new measurement closer to average.⁶² The reason is simple: an extreme measurement can be attributed to a combination of two things—the characteristics of the thing measured (e.g., the skill of a baseball player who was in fact a far superior batter than his peers) and random variation, or "luck." To the extent that the extreme measurement is due to randomness, it is likely to be less extreme upon remeasurement. Similarly, information regarding other players' performance provides context for Clemente's performance, and an estimator that incorporates such information is likely better than one based solely on the player's individual performance.⁶³

In short, if we believe that the players' batting averages are somehow bound together, we can make significant estimation improvements by incorporating other-player information.⁶⁴

Returning to the legal context, the value of a damages award (or, specifically, an award for pain and suffering or punitive damages) can be understood as attributable to 1) the application of the law to the

62. See generally Stephen M. Stigler, *Darwin, Galton and the Statistical Enlightenment*, 173 J. ROYAL STAT. SOC'Y 469 (2010).

63. We can consider two sources of variability: there is variation among the players (some are better than others, for example), and, given a particular player (and given his mean batting average in particular), there is variation in the player's batting average. "Regression toward the mean" thus operates in two ways: first, on the latter source, pulling a player's batting average in toward his mean batting average, and second, on the former source, pulling a player's batting average in toward the league's mean batting average. Incorporating other-player information allows us to account for the pull toward the league's mean batting average.

64. See generally Efron & Morris, *supra* note 59; Bradley Efron & Carl Morris, *Stein's Paradox in Statistics*, 236 SCI. AM. 119 (1977).

particular facts of the case, and 2) random variation. In order to gain a deeper understanding of the benefits of CCG, let us model an award “hierarchically” to account for both of the foregoing elements. As above, we assume that the actual award $\hat{\alpha}$ is distributed with mean α and variance σ^2 . But now, let us assume that α is itself random rather than constant, and that it is distributed with mean μ and variance τ^2 . Notationally, $\hat{\alpha} \sim (\alpha, \sigma^2)$ and $\alpha \sim (\mu, \tau^2)$. In words, the hierarchical model incorporates variability on two levels. First, the upper level ($\alpha \sim (\mu, \tau^2)$) indicates that cases are heterogeneous—that there can be factual variability across cases, and that the present case may be materially different from prior cases. I refer to this form of variability, τ^2 , as “claim variability.”⁶⁵ Second, the lower level ($\hat{\alpha} \sim (\alpha, \sigma^2)$) indicates that, given the facts of any single case (for example, the present case), there is variability (i.e., randomness) in the determination of damages. I refer to this form of variability, σ^2 , as “judgment variability.”⁶⁶

Again, our goal is to arrive at a good estimate $\hat{\alpha}$ of α . On one hand, we want the estimate to reflect the first form of variability—that meaningful factual differences (e.g., between the present case and prior cases) should result in different awards. On the other hand, we want to minimize the second form of variability—the randomness associated with the determination of damages in a given case. So how does incorporating prior-award information help? Unfortunately, since we generally do not *know* the correct damages award, it is impossible to distinguish the former form of variability from the latter. In other words, a tradeoff arises between minimizing bias—that our estimated damages award should, on average, be as close to the correct damages award as possible, reflecting the former form of variability—and minimizing variance, the degree of randomness associated with the estimated award. Thus, in a sense, incorporating prior-award information as guidance in determining a damages award allows the fact finder (implicitly) to strike a balance between bias and variance so as to minimize error. By accepting the possibility of introducing some bias due to material differences between the present case and prior cases, it is possible to gain far more in terms of accuracy, due to the reduction in award variability that follows from such guidance.

Furthermore, the incorporation of prior-award information should not be viewed as suffering from circularity—that is, as a method that is defeated by its reliance on information that entails the same flaws that

65. See Bavli, *supra* note 50, at 74–78.

66. See *id.* (discussing “judgment variability” and “claim variability”); see also Hillel J. Bavli, *Sampling and Reliability in Class Action Litigation*, 2016 CARDOZO L. REV. DE NOVO 207, 210–12 (2016).

the method is itself attempting to address. Rather, such methods should be understood as facilitating the improvement of awards by allowing for information sharing across cases. In the framework described, for example, an ideal method for reducing error from variability would generate a damages award by repeatedly adjudicating a case many times independently and using the mean of the repeated adjudications as the ultimate award. But such methods are generally impractical. Methods involving the incorporation of prior-award information can be understood as efficient alternatives. Such methods aim to control variability through information sharing across *different cases*, rather than across *replications of the same case*. The cost of efficiently reducing judgment variability, however, is the potential for bias arising from claim variability.

In order to gain a better understanding of the effects of prior-award information on error, let us consider two extreme scenarios. First, suppose that the prior-award information is “dogmatic” in the sense that it entirely dominates the damages award without leaving room for any influence from the facts of the present case itself—for example, suppose the jury decides dogmatically simply to apply the average of the previous awards as its award determination, without consideration of the facts of the current case. In this case, there would be no error from variability, since the average of the previous damage awards is constant; but to the extent that the average of the prior awards is different from the correct current award α , there would be substantial error from bias. On the other hand, suppose that the prior-award information has no influence on the current damages award—for example, suppose the jury decides dogmatically that it will arrive at an award determination based on the facts of the case alone, without influence from prior damage awards. In this case, assuming jury damage determinations are initially (i.e., without influence from prior awards) unbiased—that is, they achieve the correct award α on average—then there would be no error from bias, but substantial error from variability.

CCG improves the accuracy of awards for pain and suffering and punitive damages—award types that suffer from particularly high degrees of variability—by facilitating a balance between minimizing variability and introducing the possibility of bias. The quality of the balance is determined by the strength of the method for identifying comparable cases and distilling information for consideration by the trier of fact.

To be sure, let us consider two important issues. First, assuming an “appropriate” set of prior cases, what is the ideal level of influence that should be afforded to prior-award information? Second, how can we be certain that the prior cases identified are sufficiently comparable to the

present case so as to improve accuracy?

The first issue—the ideal influence of prior-award information, assuming an appropriate set of prior cases—depends on the judgment variability (σ^2) of the award in the present case and the claim variability (τ^2) of the prior cases. Higher judgment variability suggests weaker information obtained from the present case and greater influence of the prior awards; higher claim variability suggests weaker information obtained from the prior cases and greater influence of the present case.

For example, consider a tort case in which the evidence against the defendant is overwhelming and the damages incurred by the plaintiff are clear. Assume that the evidence is sufficiently strong that, if the case were litigated ten times independently, the damage awards in all ten replications would be approximately the same—say, \$100,000. In this case, the evidence is so great and the judgment variability so low (almost zero) that there is little to be gained from introducing prior-award information and the possibility of bias. On the other hand, consider a case in which, although there may be strong evidence regarding the facts of the case, there is only weak evidence regarding the appropriate damages award—as is almost always the case for awards for pain and suffering and punitive damages. If the case were repeatedly adjudicated ten times independently, the damage awards across the ten replications would vary wildly. In this case, the damages award is highly variable and there is much to be gained by introducing prior-award information. Further, the prior-award information is especially beneficial if claim variability is low and prior awards provide clear information. For example, assume that the facts and legal issues in the present case are clear, and that the court confidently identifies a set of ten comparable cases that involve similar facts—for example, similar injuries arising from similar circumstances. Assume further that the damages awarded in the ten cases are very similar—say, approximately \$50,000. In this case, in which judgment variability is high and claim variability is low, prior-award information should have a high degree of influence.⁶⁷

67. More formally, it can be shown that, if, as above, the damages award \hat{a} is distributed with mean a and variance σ^2 , and the correct damages award a is distributed with mean μ and variance τ^2 (thus, if we have judgment variability σ^2 and claim variability τ^2), then the optimal damages award weights current case information by $1/\sigma^2$ and prior-award information by $1/\tau^2$. In other words, the ideal influence of the prior awards (and particularly the mean of the prior awards) is inversely proportional to the claim variability (τ^2) of the prior awards relative to the judgment variability (σ^2) of the award in the present case. See generally Efron & Morris, *supra* note 59; W. James & Charles Stein, *Estimation with Quadratic Loss*, 1 PROC. OF THE FOURTH BERKELEY SYMP. ON MATHEMATICAL STAT. AND PROBABILITY 361 (1961). Note that since we cannot, in practice, know the value of μ , the mean of the correct awards, we instead estimate μ using the mean of the prior awards, and “shrink” toward this

It is worth reemphasizing here that it is incorrect to conclude, as numerous commentators have, that, because prior awards may suffer from the same arbitrariness that we aim to address in the present case, CCG would be ineffective in reducing (or, worse, would compound) award variability. Even assuming the (not-unlikely) scenario that judgment variability is approximately equal across cases (as I have assumed above), the foregoing analysis makes clear that the ideal influence of prior-award information can nevertheless be very high. In other words, CCG may nevertheless significantly reduce variability and improve accuracy. This makes sense, since CCG allows for the sharing of information across numerous (although variable) cases; it provides significant guidance where relevant information is otherwise scarce.

Consider an extreme example in which judgment variability is extremely high for all cases (the prior cases as well as the present case) and claim variability is extremely low. Here, although prior awards suffer from the same extreme arbitrariness that we aim to address in the present award, a court would benefit immensely from CCG methods. Specifically, because claim variability is extremely low, the prior-award information obtained here would be approximately equivalent to prior-award information that could be obtained from the “ideal” procedure described above, whereby a damages award is calculated by averaging over awards obtained from repeated independent adjudications (with different juries, judges, lawyers, presentations of the evidence, etc.) of a single case, the case at hand. The awards obtained from repeated adjudications would vary considerably, since judgment variability is extremely high. But the ultimate award, calculated by averaging all of the awards obtained from the repeated adjudications, is extremely precise—it is subject to almost no variability.

As suggested above, methods involving repeated adjudication are ideal but costly. CCG methods are efficient alternatives that come at the cost of possible bias that may arise due to “misalignment”—where the mean of the correct awards in the prior cases is different from the correct award in the subject case—and perhaps due to claim variability, or heterogeneity in the set of prior cases.⁶⁸ In our current example, however, bias is not a problem, because there is no misalignment problem and claim variability is assumed to be extremely low. Thus, although the prior awards are subject to extremely high judgment variability, in the aggregate they allow for an extremely accurate

value. We can estimate the variance terms as well. Notationally, the “shrinkage estimator” can be

written as $\hat{\alpha}_s = \frac{\hat{\alpha} + \frac{\mu}{\tau^2}}{\frac{1}{\sigma^2} + \frac{1}{\tau^2}}$.

68. See *infra* note 81.

outcome through the sharing of information across cases. Furthermore, we can relax the assumption of extremely low claim variability. Higher claim variability may cause bias; but the harm to accuracy associated with any introduction of bias is often well outweighed by the accuracy benefits associated with the reduction in judgment variability. In any case, the introduction of claim variability does not affect the conclusion that CCG methods are effective in reducing judgment variability and improving accuracy notwithstanding the potential arbitrariness associated with the prior awards.

The second issue—the accuracy benefits of CCG in light of the *actual* comparability of the prior awards—depends on whether the potential for bias is outweighed by the benefits of reducing variability. This is a valid concern in the sense that the accuracy benefits of CCG are indeed dependent on the material comparability of the prior cases to the present case.

For two reasons, however, the likelihood of identifying a set of cases sufficiently inappropriate, relative to the present case, to *harm* accuracy is very low. First, because awards for pain and suffering and punitive damages are highly variable, the potential accuracy benefits of CCG are great. As discussed above, the higher the variability of the award, the greater the potential benefits of CCG and the greater our tolerance for bias. The potential for harm from bias arising from misalignment would be more concerning, for example, if a court were interested in applying prior-award information to reduce the already-low variability of an award for a certain type of economic damages. However, the high variability of awards for pain and suffering and punitive damages—sufficiently high to prompt drastic (and highly biasing) measures such as damage caps—suggests that the benefits with respect to variability are likely to dominate any bias arising from misalignment (or any negative effects arising from high claim variability).

Second, the bias introduced by CCG methods is likely to be low. Statistically, the appropriateness of a set of comparable cases, relative to the present case, involves three major factors: 1) the “alignment” of the mean of the correct awards in the prior cases with the correct award in the present case, 2) the variability of the correct awards in the prior cases, and 3) the number of prior cases, or the “sample size.” In practice, a court should be concerned with the alignment of the material facts of the prior cases with those of the present case, the substantive breadth of the prior cases, and the sample size.

Note that we cannot perfectly predict how a trier of fact will in fact incorporate guidance from prior awards. For purposes of this discussion, and for the guidance I provide in Part V below regarding the development of comparable-case information, I assume that the trier of

fact will not act “irrationally” by affording more influence to highly dispersed prior awards than to minimally dispersed prior awards. It is not necessary for the trier of fact to behave perfectly rationally, but the foregoing assumption provides a good starting point for guidance. We trust the trier of fact to act approximately rationally in numerous legal contexts; and in developing guidance for the trier of fact, the court arguably should assume some degree of rationality, as it does in providing other forms of guidance.

Statistically (using fundamentals of “shrinkage estimation”), incorporating prior-award information will generally improve accuracy, as long as the prior awards are not tightly bound (i.e., with low variability) around a mean that is far from the correct award in the present case⁶⁹—in practice, as long as the prior awards are not tightly clustered around an award that reflects facts materially and substantially dissimilar to the present case.

High prior-award variability itself is (absent extreme conditions) unlikely to harm the damages award, since the variability of awards for pain and suffering and punitive damages is high, and, as suggested by the discussion above, the weight afforded the prior-award information (and, statistically, the mean of the prior awards in particular) is proportional to the inverse of the variability of the prior awards. This means that highly dispersed prior awards (e.g., prior awards that span a wide range of values) will have little influence on the award determination.

Additionally, although misaligned prior cases can cause significant bias, this concern fades when we consider the relatively remote conditions necessary for CCG to *harm* accuracy, in the sense of increasing award error. In particular, for CCG to harm accuracy, it would generally require a combination of significantly dissimilar cases and low prior-award variability—a combination that is highly unlikely if the prior cases are specifically selected to be comparable to the present case, and where the variability of the awards can often reflect (and perhaps should reflect) the court’s level of confidence regarding the comparability of the prior cases to the present case.⁷⁰

In short, we can expect that a reasonable method for identifying prior comparable cases will result in prior-award information that is likely to improve accuracy, and therefore reliability.⁷¹ Of course, the more comparable the prior cases are to the present case, the better the guidance for the fact finder, and the greater the accuracy benefits. In the

69. See generally Efron & Morris, *supra* note 59; James & Stein, *supra* note 67.

70. See *infra* Part V for a discussion of methods for selecting comparable cases.

71. A formal derivation of the particular mean and variability conditions that would cause a reduction in accuracy is beyond the scope of the current article.

following part, I discuss in greater detail the development of comparable-case information.

V. DEVELOPING COMPARABLE-CASE INFORMATION

In this part, I discuss a number of considerations for establishing comparable-case (or “prior-award”) information. I address two areas of concern: 1) identifying a set of materially comparable cases, and 2) distilling comparable-case information from these cases for consideration by the trier of fact.

A. *Identifying Comparable Cases*

To use CCG, a court must identify a set of comparable cases. As suggested above, a court should be concerned with three factors: the alignment of the prior cases with the subject case, the breadth of the prior cases, and the number of prior cases. The suitability of the cases and expected benefits of the guidance depend on the courts care in balancing these factors. For example, prior awards with lower breadth generally produce, in a sense, more specific, and therefore more influential, guidance; but lower breadth also results in a smaller sample size, whereas a larger sample provides more information and therefore better guidance. Additionally, the alignment of the prior cases affects the bias introduced by the prior awards; and the breadth of the prior cases, which affects the weight or influence of the prior awards, may reflect the court’s confidence in the alignment of the cases identified.

In terms of alignment and breadth, the most significant factor affecting a court’s identification of prior cases is likely its understanding of the subject case and which facts and issues are fundamental to an appropriate comparison.

Consider, for example, Judge Posner’s analysis in *Jutzi-Johnson*. That case involved allegations that a jail was negligent in failing to identify a prisoner as a suicide risk and take precautionary measures, resulting in the prisoner hanging himself. Commenting on the parties’ identification of comparison cases, Judge Posner explained:

The plaintiff cites three cases in which damages for pain and suffering ranging from \$600,000 to \$1 million were awarded, but in each one the pain and suffering continued for hours, not minutes. The defendant confined its search for comparable cases to other prison suicide cases, implying that prisoners experience pain and suffering differently from other persons, so that it makes more sense to compare Johnson’s pain and suffering to that of a

prisoner who suffered a toothache than to that of a free person who was strangled, and concluding absurdly that any award for pain and suffering in this case that exceeded \$5,000 would be excessive. The parties should have looked at awards in other cases involving asphyxiation, for example cases of drowning, which are numerous Had they done so, they would have come up with an award in the range of \$15,000 to \$150,000.⁷²

Judge Posner thus disagreed with the *alignment* of the plaintiff's cases, which involved pain and suffering that continued for hours rather than minutes. He also disagreed with the *alignment*, as well as the *breadth*, of the defendant's cases, which were "confined" to "other prison suicide cases"—as though "prisoners experience pain and suffering differently from other persons"—leading to "absurd" conclusions regarding a suitable award.⁷³ Finally, Judge Posner displayed concern for *sample size*, suggesting that widening the breadth of the prior cases to those involving other forms of asphyxiation would have resulted in "numerous" prior cases for comparison.⁷⁴ Sample size is important because prior awards are subject to judgment variability: if a court selects only a single case for comparison, then, even if the court is extremely confident that the prior case is indeed comparable, it would remain unclear whether the award in that case is due to the specific facts of the case or to randomness. A court may be justified in sacrificing some degree of "comparability" for purposes of increasing the sample size. Of course, in practice, a court may determine that the material facts of a case are too unique for comparison, and may accordingly forego any application of CCG.

With these factors in mind, the court should arrive at a process for selecting a suitable set of prior cases. Selection methods should remain flexible so that courts may cater such methods to the unique circumstances of the case at hand. Initially, courts can obtain guidance from comparative approaches used regularly to determine damages in other civil contexts.⁷⁵

Thus, consider a simple procedure in which the court asks each party to identify a set of comparable cases. Such a procedure is likely to result in a set of extreme cases, since each party selects cases to

72. *Jutzi-Johnson v. United States*, 263 F.3d 753, 760-61 (7th Cir. 2001).

73. *Id.*

74. *Id.*

75. *See infra* Part VI. Courts may also obtain guidance from the class action context, where courts sometimes select a set of "representative" or "bellwether" cases for settlement or litigation purposes. *See generally* Bavli, *supra* note 50.

maximize its own interests.⁷⁶ On the other hand, a method by which the court, rather than the parties, selects a set of comparable cases is likely to produce less variability, but is susceptible to a number of drawbacks, including the potential for judicial bias and undue burden on the court. Alternatively, the court may direct the parties to *agree* on a set of prior cases; but this may be difficult and inefficient without additional structure from the court.⁷⁷

Although the circumstances of a particular case may call for one of the methods above, a particularly useful approach is one in which the court selects a final set of cases from a pool identified by the litigants and the court. Variations of this method have been used in various civil contexts.⁷⁸ It allows for significant flexibility, and benefits from both a litigant-based selection process and the relative objectivity of the judge.

Finally, it is possible to include (implicitly) the trier of fact in the selection process by providing it with qualitative information regarding the set of comparable cases, thus allowing it to weight the prior-award information based, in part, on its assessment regarding the alignment of the prior cases. I discuss this possibility further in the following section.

B. Distilling Comparable-Case Information

Once the court has selected a suitable set of prior cases, it must decide what information should be considered by the trier of fact as guidance. A study by Professor Michael Saks et al. examined the effects of various forms of jury guidance on amounts awarded for pain and suffering in personal injury cases.⁷⁹ They found that, while damage caps performed poorly in terms of variability and distortion of award magnitude, various forms of distributional information—in particular, an “award interval,” an “award average-plus-interval,” and “award examples”—were effective in reducing variability while distorting award magnitude minimally or not at all.⁸⁰ These results support the model described in Part IV. In particular, in line with the analysis above, prior-award information is expected to cause a reduction in variability and minimal (if any) distortion of award magnitude where the distribution of prior-awards is similar to the distribution of awards that would be obtained in

76. See generally Loren H. Brown et al., *Bellwether Trial Selection in Multi-District Litigation: Empirical Evidence in Favor of Random Selection*, 47 AKRON L. REV. 663, 674–75 (2014).

77. See generally *id.* at 675.

78. See *infra* Part VI; see also Brown et al., *supra* note 76, at 672–76 (describing cases employing such procedures in the bellwether context). See generally Alexandra D. Lahav, *The Case for “Trial by Formula,”* 90 TEX. L. REV. 571 (2012).

79. See Saks et al., *supra* note 15.

80. See *id.* at 249–55. Saks et al. observed mixed results for an “award average” form. See *id.* at 253–55.

the absence of prior-award information. Indeed, Saks et al. obtained the values of the guidance provided to survey participants from the results of their pilot study. Specifically, they used the median of the pilot study for the award average, the 10th and 90th percentiles for the award interval, the 10th, 50th, and 90th percentiles for the award average-plus-interval, and the 10th, 35th, 65th, and 90th percentiles for the award examples.⁸¹

The discussion in Part IV suggests that courts should choose information forms that provide jurors with information regarding the “center” as well as the “spread,” or variability, of the distribution of prior awards.⁸² The award interval, award average-plus-interval, and award examples have the advantage, relative to providing, for example, only the award average, that jurors are provided information regarding the variability of the prior awards. This enables them to weight the prior-award information appropriately. The award examples form is particularly appealing, because it provides jurors with the most information, relative to the other information forms, regarding prior-award variability.

Further, in addition to numerical prior-award information, it may be beneficial for the court to allow the trier of fact to hear qualitative summary information regarding the prior cases. Providing the trier of fact with such information would strengthen its role in the possibly fact-intensive identification process, and would mitigate the effects of potential bias from misalignment. As suggested above, providing the trier of fact with qualitative case information implicitly includes it in the case-selection process by enabling it to weight the prior-award information based, in part, on its assessment of the comparability of the prior cases.⁸³

If the court allows the trier of fact to hear qualitative case information, it must determine an appropriate form for such information—for example, a brief description regarding the set of cases as a whole, or a set of short factual summaries regarding each individual case.

As described in Part VI below, there is substantial precedent for allowing a jury to hear comparable-case information in the form of

81. *Id.* at 248. The authors’ results for the *cap condition*, which caused a distortion in the size, or magnitude, of the award and mixed effects on variability (depending on the severity category) can be interpreted as supportive of the model herein, but may also have implications for the effects of prior-award information.

82. It is possible that, empirically, jurors will not in fact weight the prior-award information according to the relative variability of such information; but without significant evidence to the contrary, courts should arguably provide jurors with relevant variability information.

83. The court should balance the benefits of allowing qualitative case information with the costs (including attorneys’ fees, delays, etc.) of allowing such information.

evidence introduced as expert testimony. A jury may thereby “consider . . . arguments about the effect of [purported differences] on the validity of [the] comparison and . . . adjust its damage award accordingly.”⁸⁴

VI. CONCLUSION

In the previous parts, I addressed three major objections to CCG methods. First, I explained that CCG can reduce unpredictability and improve the accuracy of awards—notwithstanding the judgment variability associated with the prior awards—by allowing for the sharing of information across cases. Second, I discussed how information regarding awards in comparable cases can be understood not as an alternative to assessing damages for pain and suffering or an appropriate punitive damages award, but as fundamental to assessing such values. Finally, I explained that it is not necessary for the prior cases to be perfectly comparable to the subject case, and that, assuming a reasonable method for identifying comparable cases, it is unlikely—particularly for awards for pain and suffering and punitive damages—that they will diminish, rather than improve, accuracy.

Moreover, it is important to understand that using comparable-case information as guidance in determining awards for pain and suffering and punitive damages is not a remote concept that requires dramatic change to current practices. Indeed, courts have recognized the importance of comparable-case information as guidance in determining damage awards in a range of contexts.

First, there is substantial precedent from the judicial-review context for allowing information regarding prior awards to influence present awards and, in particular, for the use of comparable-case information as guidance in determining an appropriate award or range of awards. Indeed, courts have described the comparison of awards to prior awards for similar injuries as “[a] mainstay of the excessiveness determination,” noting that “[t]his use of comparison is a recognition that the evaluation of emotional damages is not readily susceptible to ‘rational analysis.’”⁸⁵ For example, courts in New York have held that, in determining whether an award is excessive, “[a] reviewing court *must* consider awards in comparable cases”⁸⁶—that the “determination of whether a compensatory damages award is excessive should not be conducted in a vacuum, but instead should include consideration of the amounts

84. *Syufy Enters. v. Am. Multicinema, Inc.*, 793 F.2d 990, 1003 (9th Cir. 1986).

85. *Salinas v. O’Neill*, 286 F.3d 827, 830–31 (5th Cir. 2002) (internal citations omitted).

86. *Alla v. Verkey*, 979 F. Supp. 2d 349, 372 (E.D.N.Y. 2013) (emphasis added).

awarded in other, comparable cases.”⁸⁷ The U.S. Supreme Court has itself indicated, in the context of determining whether a particular punitive damages award exceeded the amount permitted by the Due Process Clause of the Constitution, that “[c]omparing the punitive damages award and the civil or criminal penalties that could be imposed for comparable misconduct provides a[n] . . . indicium of excessiveness.”⁸⁸

Second, there is substantial precedent from the bench-trial context for CCG—and specifically, for consideration of comparable-case information by the trier of fact for guidance in determining awards for pain and suffering and punitive damages. In fact, consideration of comparable-case information is often expected. For example, in *Jutzi-Johnson*, Judge Posner remarked that most courts “treat the determination of how much damages for pain and suffering to award as a standardless, unguided exercise of discretion by the trier of fact, reviewable for abuse of discretion pursuant to no standard to guide the reviewing court either.”⁸⁹ He advised:

To minimize the arbitrary variance in awards bound to result from such a throw-up-the-hands approach, the trier of fact should . . . be informed of the amounts of pain and suffering damages awarded in similar cases. And when the trier of fact is a judge, he should be required as part of his Rule 52(a) obligation to set forth in his opinion the damages awards that he considered comparable. We make such comparisons routinely in reviewing pain and suffering awards, as do other courts. It would be a wise practice to follow at

87. *DiSorbo v. Hoy*, 343 F.3d 172, 183 (2d Cir. 2003) (internal quotation marks and citations omitted); see *Geressy v. Dig. Equip. Corp.*, 980 F. Supp. 640, 656 (E.D.N.Y. 1997) (“CPLR 5501(c)’s conception of reasonable compensation cannot exist in a vacuum. There needs to be some point of reference. With economic damages, the court may rely on traditional methods of economic analysis. As for the non-economic pain and suffering award, the reviewing court must begin by identifying some group of similar cases to serve as a referent. This task is difficult, particularly in cases exploring relatively new types of injuries and claims such as those in the instant case involving [repetitive stress injury (RSI)] claims against a keyboard manufacturer. Cases with similar causal agents, similarly-named diagnoses, or similar reductions in quality of life might serve as benchmarks.”), *aff’d in part sub nom. Madden v. Dig. Equip. Corp.*, 152 F.3d 919 (2d Cir. 1998).

88. See *BMW of N. Am., Inc. v. Gore*, 517 U.S. 559, 583 (1996); see also *Degorski v. Wilson*, No. 04 CV 3367, 2014 WL 3511220, at *1 (N.D. Ill. July 16, 2014) (“In assessing whether a punitive damage award is constitutionally appropriate, the Supreme Court has directed courts to focus their evaluation on three guideposts: (1) the reprehensibility of the defendant’s conduct; (2) the relationship between the amount of the punitive damages awarded and the harm or potential harm suffered by the Plaintiff; and (3) the difference between the punitive damages award and the civil penalties authorized or imposed in comparable cases.”).

89. *Jutzi-Johnson v. United States*, 263 F.3d 753, 759 (7th Cir. 2001).

the trial level as well.⁹⁰

Third, there is precedent from various civil contexts for asking *jurors* to make comparisons for purposes of determining damages. The comparisons often relate to goods or services for which there is an economic market, and the comparisons are undertaken for purposes of providing information regarding the market, which, in turn, provides information regarding the appropriate damages award.

For example, in the context of determining “just compensation” in eminent domain litigation, it is common for a court to ask the jury to undertake a “comparable sales” analysis, for which the jury hears testimony regarding the selling prices of comparable properties.⁹¹ In such cases, the court intends to provide the jury with information regarding the market to which the property at issue belongs, and thereby facilitate the jury’s assessment of just compensation and thus an appropriate compensatory damages award. Indeed, comparable sales evidence is often viewed as the “best evidence” of market value.⁹² Similarly, damage awards for breaches of contract for the sale of goods or services are often computed by looking at the market price of comparable goods or services.⁹³ And courts also rely on the jury to

90. *Id.*; see also *Arpin v. United States*, 521 F.3d 769, 776–77 (7th Cir. 2008) (explaining that the lower court “should have considered awards in similar cases, both in Illinois and elsewhere,” and that “[c]ourts may be able to derive guidance for calculating damages for loss of consortium from the approach that the Supreme Court has taken in recent years to the related question of assessing the constitutionality of punitive damages”); *Zurba v. United States*, 247 F. Supp. 2d 951, 961 (N.D. Ill. 2001) (considering awards in “comparable” cases), *aff’d*, 318 F.3d 736 (7th Cir. 2003); *Elk v. United States*, 87 Fed. Cl. 70, 97 (2009) (“Various cases, including those arising under the FTCA, suggest that judges faced with the prospect of determining pain and suffering awards should look to the awards in similar cases.”); *Cochran v. A/H Battery Assocs. and Arcorp Props.*, 909 F. Supp. 911, 917 (S.D.N.Y. 1995) (“[C]ourts have long reviewed damage awards in other similar cases to determine the scope of reasonableness in a particular case.”).

91. See *United States v. 1,129.75 Acres of Land*, 473 F.2d 996, 998 (8th Cir. 1973) (citing cases, and holding that, “[g]enerally, evidence of sales of comparable property is persuasive evidence of market value, either as direct proof or in support of a witness’s opinion”).

92. See *United States v. 24.48 Acres of Land*, 812 F.2d 216, 218 (5th Cir. 1987) (citing *United States v. 8.41 Acres*, 680 F.2d 388, 395 (5th Cir. 1982)) (“First, we reiterate the principle that the best evidence of market value is comparable sales.”); see also John J. Dvorske & Ann K. Wooster, *Condemnation of Property, Comparable Sales, Generally*, in 7 FEDERAL PROCEDURE, LAWYER’S EDITION § 14:90 (Dec. 2015).

93. See U.C.C. § 2-713 (AM. LAW INSTIT. 2014) (“[T]he measure of damages for non-delivery or repudiation by the seller [of goods] is the difference between the market price at the time when the buyer learned of the breach and the contract price together with any incidental and consequential damages provided in th[e] Article (Section 2-715), but less expenses saved in consequence of the seller’s breach.”); see also *Gulf Power Co. v. Coalsales II, LLC*, 522 F. App’x. 699, 707 (11th Cir. 2013) (holding, in a case involving a breach of a contract for the purchase and sale of coal, that “the appropriate measure of damages would be the market remedy provided for in § 672.713 [of the U.C.C.], in which damages are calculated by comparing the contract price to the fair market value of comparable goods on the open market”).

make comparisons in arriving at damages calculations in the antitrust context.⁹⁴

A detailed examination of the legal foundations of CCG is beyond the scope of this article. The legal context above is intended simply to show that there is a long history, across various contexts, of using comparable-case information as guidance in determining damages. Indeed, courts have applied such methods notwithstanding the objections addressed in this article. For example, consider the objection that courts will compound the arbitrariness of awards by obtaining guidance from prior awards that presumably suffer from the same arbitrariness that courts hope to address. This objection applies similarly to comparable-case information considered in the bench-trial context, but bench-trial judges nevertheless consider comparable-case information. Additionally, the objection that methods involving the use of prior awards “fail to address the fundamental issue of how one should initially assess the value of pain-and-suffering damages”⁹⁵ or arrive at an appropriate punitive damages award applies similarly to the use of comparable-case information in the judicial-review and bench-trial contexts; but such information is still used in these contexts. Finally, the objection that the validity of methods involving comparable-case information relies heavily on the presumption that the prior cases identified are indeed materially comparable to the present case applies similarly to all of the contexts discussed above,⁹⁶ but such methods are nevertheless used.

My aim in this article was to address the major objections to CCG methods by explaining simply, but formally, what many courts and commentators have recognized implicitly: that CCG is effective in reducing unpredictability and improving the accuracy of awards for pain and suffering and punitive damages by allowing for the sharing of information across cases.

94. For example, in *Syufy Enters. v. Am. Multicinema, Inc.*, 793 F.2d 990 (9th Cir. 1986), counterclaimant AMC based its damages calculation on comparisons with purportedly comparable markets. It presented such comparisons to the jury, which ultimately rendered an award accordingly. *See id.* at 1002–03. The Ninth Circuit rejected the argument that damages may not be computed based on comparable markets. *See id.* The court then rejected the argument that the comparison markets were not comparable to the market at issue and thereby inappropriately inflating AMC’s damages calculation. *See id.* at 1003. The court held that “[c]omparability is a question of fact” and that “[i]t was for the jury to consider [the opposing party’s] arguments about the effect of the [claimed differences] on the validity of comparison and to adjust its damage award accordingly. There was sufficient evidence to allow a jury to render a damage award for AMC on the basis of the comparison” *Id.*

95. Avraham, *supra* note 19, at 104.

96. The objection applies directly to the judicial-review and bench-trial contexts and analogously to the market-comparisons context.