Washington University Law Review

Volume 95 | Issue 4

2018

Income-Dependent Punitive Damages

Ronen Perry
University of Haifa

Elena Kantorowicz-Reznichenko
Erasmus School of Law

INCOME-DEPENDENT PUNITIVE DAMAGES

RONEN PERRY[*]
ELENA KANTOROWICZ-REZNICHENKO[**]

TABLE OF CONTENTS

* Professor of Law and Director, Aptowitzer Center for the Study of Risk, Liability, and Insurance, University of Haifa.

** Lecturer and Researcher, Rotterdam Institute of Law & Economics, Erasmus School of Law, Erasmus University Rotterdam.

We are grateful to Michael Faure, Louis Visscher, Ramsi Woodcock, participants in the 34th Annual Conference of the European Association of Law & Economics in London, and the editors of the *Washington University Law Review* for valuable comments on earlier drafts.

INTRODUCTION

Punitive damages are sums awarded to tort victims over and above their compensable harm.[1] Despite their relative rarity,[2] they have been very salient in the media,[3] preoccupied appellate courts,[4] and fascinated scholars for decades.[5] This prominence may be attributed, at least in part, to a combination of doctrinal idiosyncrasy and stupefying case outcomes. On the doctrinal level, punitive damages are a civil law remedy which is patently inconsistent with the traditional goals of civil law. From a case outcome perspective, the debate is fueled by "blockbuster awards,"[6] such as a $145 billion award in a class-action brought against tobacco companies in Florida[7] and a $28 billion award in an individual action

1. *See, e.g.*, James D. Ghiardi, *Punitive Damages in Wisconsin*, 60 MARQ. L. REV. 753, 760, 766 (1977) (explaining that punitive damages are awarded "over and above" any amounts necessary for compensation); Robert E. Riggs, *Constitutionalizing Punitive Damages: The Limits of Due Process*, 52 OHIO ST. L.J. 859, 893 (1991) (same).

2. All empirical studies found that punitive damages are very rarely awarded. Theodore Eisenberg & Martin T. Wells, *The Significant Association Between Punitive and Compensatory Damages in Blockbuster Cases: A Methodological Primer*, 3 J. EMPIRICAL LEGAL STUD. 175, 175 (2006) (summarizing empirical research); Catherine M. Sharkey, *Economic Analysis of Punitive Damages: Theory, Empirics, and Doctrine*, *in* RES. HANDBOOK ON ECON. TORTS 486, 500 (Jennifer Arlen ed., 2012) (same).

3. *See, e.g.*, Adam Liptak, *Foreign Courts Wary of U.S. Punitive Damages*, N.Y. TIMES (Mar. 26, 2008), http://www.nytimes.com/2008/03/26/us/26punitive.html.

4. *See, e.g.*, Atl Sounding Co. v. Townsend, 557 U.S. 404 (2009); Exxon Shipping Co. v. Baker, 554 U.S. 471 (2008); Philip Morris USA v. Williams, 549 U.S. 346 (2007); State Farm Mut. Auto. Ins. Co. v. Campbell, 538 U.S. 408 (2003); Cooper Indus., Inc. v. Leatherman Tool Grp., Inc., 532 U.S. 424 (2001); BMW of N. Am., Inc. v. Gore, 517 U.S. 559 (1996); Honda Motor Co. v. Oberg, 512 U.S. 415 (1994); TXO Prod. Corp. v. All Res. Corp., 509 U.S. 443 (1993); Pac Mut. Life Ins. Co. v. Haslip, 499 U.S. 1, 25 (1991); Browning-Ferris Indus. of Vt., Inc. v. Kelco Disposal, Inc., 492 U.S. 257 (1989); *see also infra* notes 7–8 and accompanying text.

5. *See, e.g.*, Kenneth S. Abraham & John C. Jeffries Jr., *Punitive Damages and the Rule of Law: The Role of Defendant's Wealth*, 18 J. LEGAL STUD. 415 (1989); Robert D. Cooter, *Punitive Damages for Deterrence: When and How Much?*, 40 ALA. L. REV. 1143 (1989); Theodore Eisenberg & Michael Heise, *Judge-Jury Difference in Punitive Damages Awards: Who Listens to the Supreme Court?*, 8 J. EMPIRICAL LEGAL STUD. 325, 327 (2011); Dorsey D. Ellis, *Fairness and Efficiency in the Law of Punitive Damages*, 56 S. CAL. L. REV. 1 (1982); Joni Hersch & W. Kip Viscusi, *Punitive Damages: How Judges and Juries Perform*, 33 J. LEGAL STUD. 1 (2004); Keith N. Hylton, *A Theory of Wealth and Punitive Damages*, 17 WIDENER L.J. 927 (2008); A. Mitchell Polinsky & Steven Shavell, *Punitive Damages: An Economic Analysis*, 111 HARV. L. REV. 869 (1998); Michael Rustad & Thomas Koenig, *The Historical Continuity of Punitive Damages Awards: Reforming The Tort Reformers*, 42 AM. U. L. REV. 1269 (1993); Gary T. Schwartz, *Deterrence and Punishment in the Common Law of Punitive Damages: A Comment*, 56 S. CAL. L. REV. 133 (1982); Anthony J. Sebok, *Punitive Damages: From Myth to Theory*, 92 IOWA L. REV. 957 (2007); Catherine M. Sharkey, *Punitive Damages as Societal Damages*, 113 YALE L.J. 347 (2003); Cass R. Sunstein et al., *Assessing Punitive Damages (with Notes on Cognition and Valuation in Law)*, 107 YALE L.J. 2071, 2081 (1998); *see also supra* notes 1–2.

6. *See* Hersch & Viscusi, *supra* note 5, at 4–10 (introducing the term "blockbuster awards"); Eisenberg & Wells, *supra* note 2 *passim* (responding to Hersch & Viscusi).

7. Liggett Grp. Inc. v. Engle, 853 So. 2d 434, 440–41 (Fla. 2006). The Florida Supreme Court observed that this was the largest punitive damages verdict in history, and concluded that it was

against Philip Morris in California.[8] While in both cases, as in many others, the extraordinary jury awards were ultimately reduced or overturned, they have surely left a notable mark.

The availability of punitive damages seems undisputed in most common law jurisdictions, but their measure remains controversial.[9] In particular, it is unclear whether and how courts and juries should take the defendant's wealth into account in assessing punitive damages.[10] This Article puts forward and defends an innovative yet simple method for incorporating this factor into the calculation. The proposal is based on an adaptation of a criminal law model, known as "day-fines," which has been primarily used in European and Latin-American legal systems.[11] A criminal day-fine model is based on two variables: "number of days" and "daily unit."[12] The former represents the gravity of the offense; measuring gravity in days makes the monetary sanction conceptually and substantively commensurate with incarceration. The latter reflects the offender's financial condition. A daily unit usually constitutes a fixed portion of the convicted delinquent's daily income. The two factors are multiplied to ascertain the total fine.

Building on global criminal law experience, we propose a new model for assessing punitive damages. In brief, if the gravity of the wrong seems to justify an extra-compensatory award, the scope of punitive damages will be determined in several steps. First, the court will determine the gravity of the wrong and translate it into corresponding "severity units" (analogous to the "number of days" in the criminal law model). Next, the court will assess the wrongdoer's daily income, broadly defined, or a particular fraction thereof. This is the "unit value" (analogous to the daily unit in the criminal law model). The product of these two variables constitutes "total damages." Lastly, if total damages are greater than compensatory damages in the particular case, the punitive award should equal the difference between total damages and compensatory damages. If total damages are lower than compensatory damages, punitive damages should be nil. In such cases the monetary sanction that would serve the twin goals of punitive damages, deterrence and retribution, is lower than

unconstitutional and violated Florida law. *Id.* at 456, 470. Moreover, the court decided to decertify the class. *Id.* at 440, 470.

8. Bullock v. Philip Morris USA, Inc., 131 Cal. Rptr. 3d 382, 391 (Cal. Ct. App. 2011). The plaintiff agreed to reduce the punitive award to $28 million to avoid a new trial. *Id.* at 391. However, the Court of Appeal of California granted a new trial on the matter of damages, and the punitive award was further reduced to $13.8 million. *Id.* at 392. The Court of Appeal held in 2011 that this amount was justified in light of the gravity of the defendant's misconduct. *Id.* at 406.

9. *See infra* Part I.

10. *See infra* notes 137–138, 239–245 and accompanying text.

11. *See infra* Section IV.A.

12. *See infra* Section IV.A.

compensatory damages. However, as long as tort law is committed to rectification of harm, compensatory damages will set the lower limit of the monetary award.

The Article unfolds in six parts. Part I outlines the development of the law governing punitive damages. Part II analyzes the possible rationales for this unique "middle-ground" doctrine, focusing on deterrence and retribution. Part III considers *whether* the defendant's wealth should be considered in assessing punitive damages in light of their underlying goals. Part IV demonstrates *how* the defendant's wealth can be integrated into the calculation. It extracts the foundations from European criminal justice systems and adapts the model to American civil law. Part V defends the proposed model from the relevant theoretical perspectives. Lastly, Part VI discusses potential hurdles to the implementation of the new model—constitutional constraints, statutory caps on punitive damages, and the need for special procedural tools.

I. THE LEGAL FRAMEWORK

Originating in England in the mid-eighteenth century,[13] punitive damages were soon imported into America. The first reported case was *Genay v. Norris*,[14] where the court held that a person poisoned by another was entitled to exemplary damages.[15] A few years later, in a breach of promise of marriage case, the court instructed the jury "not to estimate the damages by any particular proof of suffering or actual loss; but to give damages for *example's* sake, to prevent such offences in future."[16]

A fierce debate over the availability and legitimacy of punitive damages erupted in the mid-nineteenth century between Simon Greenleaf and Theodore Sedgwick. On the descriptive level, Greenleaf insisted that damages constituted compensation for an actual harm, and should be "precisely commensurate with injury; neither more nor less."[17] Sedgwick contended that in cases of fraud, malice, gross negligence, or oppression, the law permitted punitive damages, thereby blending the public and

13. *See* Wilkes v. Wood (1763) 98 Eng. Rep. 489, 498–99 (C.P.) ("[A] jury have it in their power to give damages for more than the injury received. Damages are designed not only as a satisfaction to the injured person, but likewise as a punishment . . . ."); Huckle v. Money (1763) 95 Eng. Rep. 768, 769 (K.B.) (introducing the term "exemplary damages"). The doctrine is firmly entrenched in English law: Kuddus v. Chief Constable of Leicestershire Constabulary [2002] 2 AC 122 (HL) (appeal taken from Eng.). Some argue that punitive damages may be traced back to the thirteenth century. Semra Mesulam, Note, *Collective Rewards and Limited Punishment: Solving the Punitive Damages Dilemma with Class*, 104 COLUM. L. REV. 1114, 1121 (2004).

14. 1 S.C.L. (1 Bay) 6 (1784).

15. *Id.* at 6.

16. Coryell v. Colbaugh, 1 N.J.L. 77, 77 (1791).

17. 2 SIMON GREENLEAF, A TREATISE ON THE LAW OF EVIDENCE 276 (9th ed. 1863).

private interests.[18] Greenleaf responded that Sedgwick misinterpreted the case law, confusing courts' willingness to let juries weigh intangible harms in assessing damages with recognition of non-compensatory damages.[19] On the prescriptive level, Greenleaf advocated a clear distinction between private and public law, insisting that a plaintiff in tort should not be permitted to vindicate the state's interests.[20] Sedgwick opined that a division between the public and private interests was "entirely fanciful and imaginary," and that "the sooner the idea [of damages as compensation] is got out of the head of a practical lawyer the better."[21]

The Supreme Court settled the dispute in *Day v. Woodworth*,[22] holding that the jury in a tort action could inflict exemplary, punitive, or vindictive damages on the defendant, based on the enormity of his wrong rather than the measure of the plaintiff's harm. The Court explained that despite past controversy, "repeated judicial decisions for more than a century are to be received as the best exposition of what the law is."[23] Of course, courts are still committed to the principle of reparation for actual harm, but most states allow punitive damages in exceptional cases, defined by the severity of the defendant's conduct.[24]

Since its recognition, the doctrine has expanded in three significant respects. First, by the early nineteenth century it had become clear that punitive damages were available not only under state common law, but also under Federal maritime law.[25] Second, by the end of the nineteenth century most jurisdictions allowed punitive damages awards not only against individuals, but also against corporations.[26] Still, there has been some controversy about the availability of the remedy against corporations

18. THEODORE SEDGWICK, A TREATISE ON THE MEASURE OF DAMAGES 666 (3rd ed. 1858).

19. 2 GREENLEAF, *supra* note 17, at 276–87 n.2; Simon Greenleaf, *The Rule of Damages in Actions Ex Delicto*, 9 L. REP. 529, 530–38 (1847).

20. Greenleaf, *supra* note 19, at 529–30. According to Greenleaf, "the state is competent to vindicate its own wrongs." *Id.* For further discussion of Greenleaf's position, see Rustad & Koenig, *supra* note 5, at 1299.

21. SEDGWICK, *supra* note 18, at 671–72.

22. 54 U.S. 363, 371 (1851).

23. *Id.*; *see also* Denver & Rio Grande Ry. v. Harris, 122 U.S. 597, 609–10 (1887) (same); McWilliams v. Bragg, 3 Wis. 424, 431 (1854) (same). *But cf.* Fay v. Parker, 53 N.H. 342, 382, 397 (1872) (holding that punishment is "out of place, irregular, anomalous, exceptional, unjust, unscientific, not to say absurd and ridiculous, when classed among civil remedies" and that the idea of punitive damages "is a monstrous heresy. It is an unsightly and an unhealthy excrescence, deforming the symmetry of the body of the law").

24. *See* Atl. Sounding Co. v. Townsend, 557 U.S. 404, 409–10 (2009) ("American courts have likewise permitted punitive damages awards in appropriate cases since at least 1784."); Pac. Mut. Life Ins. Co. v. Haslip, 499 U.S. 1, 25 (1991) ("[P]unitive or 'exemplary' damages have long been a part of Anglo-American law.").

25. The Amiable Nancy, 16 U.S. 546, 558 (1818).

26. Rustad & Koenig, *supra* note 5, at 1295.

liable under the doctrine of *respondeat superior*.[27] The controversy also persists in Federal maritime law. In the *Exxon Valdez* case, the district court instructed the jury that punitive damages could be awarded against a principal (particularly a corporation) because of an act by an agent, inter alia, where the agent was employed in a managerial capacity and was acting in the scope of employment.[28] The Ninth Circuit upheld this instruction.[29] Exxon, the defendant, argued that courts could not award punitive damages under maritime law against ship owners for actions by underlings not "directed," "countenanced," or "participated in" by the owners.[30] The Supreme Court was equally divided on this question, and therefore left the Ninth Circuit's opinion undisturbed without setting a precedent on this matter.[31] Third, while punitive damages were originally awarded in cases of malicious and mean-spirited conduct,[32] the doctrine has expanded in some jurisdictions to cases of recklessness[33] and even gross negligence.[34] These three developments laid the foundations for the unprecedented, though ultimately reduced, punitive damages award in the famous *Exxon Valdez* case.[35]

On the other hand, in the twentieth century the doctrine was somewhat restrained. In many states, the plaintiff was required to satisfy a higher standard of proof, such as "clear and convincing evidence," to obtain

27. *Id.* at 1295–97 (discussing this controversy); *see also* Ellis, *supra* note 5, at 63 ("[Some courts] follow the respondeat superior rule and hold that an employer may be liable for punitive damages for wrongful acts committed by employees in the course of their employment. [Others] follow the 'complicity rule' and limit vicarious punitive damage liability to those situations where wrongful acts were committed or specifically authorized or ratified by a managerial agent, or were committed by an unfit employee who was recklessly employed or retained."); RESTATEMENT (SECOND) OF TORTS § 909 (AM. LAW INST. 1979) (providing that "recovery of punitive damages against a principal for the acts of an agent" is limited to certain circumstances, e.g., when the principal authorized or ratified the act, or when the agent was employed in a managerial capacity); *In re* P&E Boat Rentals, 872 F.2d 642, 652 (5th Cir. 1989) (endorsing the Restatement); Protectus Alpha Navigation Co. v. N. Pac. Grain Growers, Inc., 767 F.2d 1379, 1387 (9th Cir. 1985) (same).

28. Exxon Shipping Co. v. Baker, 554 U.S. 471, 480 (2008).

29. *Id.* at 481.

30. *Id.* at 482–83.

31. *Id.* at 484.

32. Victor E. Schwartz et al., *Reining in Punitive Damages "Run Wild": Proposals for Reform by Courts and Legislatures*, 65 BROOK. L. REV. 1003, 1007 (1999); Sharkey, *supra* note 2, at 487.

33. *See* Ellis, *supra* note 5, at 20; Rustad & Koenig, *supra* note 5, at 1305–07; *see also* RESTATEMENT (SECOND) OF TORTS § 908(2) (AM. LAW INST. 1979) (providing that punitive damages may be awarded for reckless indifference to the rights of others).

34. *See, e.g.*, TXO Prod. Corp. v. All. Res. Corp., 419 S.E.2d 870, 887 (W. Va. 1992) ("[T]he punitive damages definition of malice has grown to include not only mean-spirited conduct, but also extremely negligent conduct that is likely to cause serious harm."); Rustad & Koenig, *supra* note 5, at 1305–07 (discussing this development).

35. Exxon Shipping Co. v. Baker, 554 U.S. 471 (2008).

punitive damages.[36] Moreover, in most states, punitive damages were subject to a statutory cap—a fixed upper limit or a single-digit punitive-to-compensatory-damages ratio.[37] A few state legislatures banned punitive damages generally, or in certain types of cases, such as actions against the state or public officials, medical malpractice cases, and claims against drug manufacturers.[38] Since the mid-1980s, a few state legislatures have also enacted split-recovery statutes, providing that a certain percentage of any punitive damages award shall be paid to the state.[39] Because one of the primary goals of these statutes was to prevent unusual windfalls to

36. *See* Lynda A. Sloane, Note, *The Split Award Statute: A Move Toward Effectuating the True Purpose of Punitive Damages*, 28 VAL. U. L. REV. 473, 483 (1993) (discussing the standard of proof).

37. *See, e.g.,* ALASKA STAT. § 09.17.020(f)–(g) (2017) ("[A]n award of punitive damages may not exceed the greater of (1) three times the amount of compensatory damages awarded to the plaintiff in the action; or (2) the sum of $500,000"; if the conduct was motivated by financial gain and the defendant knew the consequences, the award can reach "the greatest of (1) four times the amount of compensatory damages . . . ; (2) four times the aggregate amount of financial gain that the defendant received . . . ; or (3) the sum of $7,000,000."); COLO. REV. STAT. § 13-21-102(1)(a) (2017) ("The amount of such reasonable exemplary damages shall not exceed an amount which is equal to the amount of the actual damages awarded to the injured party."); FLA. STAT. § 768.73(1) (2017) ("[A]n award of punitive damages may not exceed the greater of: (1) Three times the amount of compensatory damages . . . or (2) The sum of $500,000"; if the conduct was "motivated solely by unreasonable financial gain," and the unreasonably dangerous nature of the conduct and high likelihood of injury were known to the defendant, the award can reach the greater of: (1) four times the amount of compensatory damages, or (2) $2 million.); IDAHO CODE § 6-1604(3) (2017) ("No judgment for punitive damages shall exceed the greater of two hundred fifty thousand dollars ($250,000) or an amount which is three (3) times the compensatory damages contained in such judgment."); N.D. CENT. CODE § 32-03.2-11 (2017) ("[T]he amount of exemplary damages may not exceed two times the amount of compensatory damages or two hundred fifty thousand dollars, whichever is greater."); *see also infra* Section VI.B.

38. *See, e.g.*, ALA. CODE § 6-11-26 (2017) ("Punitive damages may not be awarded against the State or any country or municipality thereof, or any agency thereof . . . ."); COLO. REV. STAT. § 13-21-102(5) (2017) ("[E]xemplary damages shall not be awarded in administrative or arbitration proceedings . . . ."); 735 ILL. COMP. STAT. ANN. 5/2-1115 (2017) (prohibiting punitive damages in legal or medical malpractice actions); 745 ILL. COMP. STAT. ANN. 10/2-102 (2017) (prohibiting punitive damages in actions against public officials); OHIO REV. CODE ANN. § 2307.80(C) (2017) (providing that punitive damages cannot be awarded in actions against manufacturers of drugs and medical devices); OR. REV. STAT. § 30.927, 31.740 (2017) (providing that punitive damages may not be awarded against drug manufacturers and health practitioners); W. VA. CODE § 29-12A-7 (2017) (prohibiting punitive damages in actions against a "political subdivision" and its employees).

39. *See, e.g.*, ALASKA STAT. § 09.17.020(j) (2017) (transferring 50% to the state); GA. CODE ANN. § 51-12-5.1(e)(2) (2017) (transferring 75% to the state); 735 ILL. COMP. STAT. ANN. 5/2-1207 (2017) (allowing courts to apportion punitive damages among the plaintiff, his or her attorney, and the state); IND. CODE ANN. § 34-51-3-6 (2017) (transferring 75%); IOWA CODE ANN. § 668A.1(2) (2017) (transferring 75% in some cases); MO. ANN. STAT. § 537.675 (2017) (transferring 50%); UTAH CODE ANN. § 78B-8-201(3) (2017) (transferring 50% of any amount in excess of $50,000); *see also* Scott Dodson, *Assessing the Practicality and Constitutionality of Alaska's Split-Recovery Punitive Damages Statute*, 49 DUKE L.J. 1335, 1337 (2000); Benjamin F. Evans, *"Split-Recovery" Survives: The Missouri Supreme Court Upholds the State's Power to Collect One-Half of Punitive Damage Awards*, 63 MO. L. REV. 511, 511–12 (1998). For a few years, a judge-made split-recovery rule was applied in Alabama. Life Ins. Co. of Ga. v. Johnson, 684 So. 2d 685, 698–99 (Ala. 1996). This rule was overruled in Life Ins. Co. of Ga. v. Johnson, 701 So. 2d 524, 534 (Ala. 1997) and repealed in ALA. CODE § 6-11-21(l) (2017).

plaintiffs, this trend has halted once the Supreme Court limited the punitive-to-compensatory damages ratio.[40]

Most importantly, it was held that the Due Process Clause of the Fourteenth Amendment prohibits the imposition of grossly excessive or arbitrary penalties on a wrongdoer.[41] A punitive damages award is therefore subject to substantive due process review. In *BMW v. Gore*[42] the Supreme Court held that in reviewing awards of punitive damages under the Due Process Clause, courts ought to consider three guideposts. The first is the degree of reprehensibility of the defendant's misconduct.[43] Factors relevant to this determination include the type of harm caused, victims' vulnerability, defendant's intentional malice or reckless disregard for health and safety of others, repetitive misconduct, and defendant's efforts to mitigate the harm caused.[44] The second guidepost is the disparity between the plaintiff's actual or potential harm and the punitive damages award.[45] In *State Farm Mutual Automobile Insurance Co. v. Campbell*,[46] the Court held that a single-digit ratio between punitive and compensatory damages was more likely to accord with due process than awards with ratios in the range of 500 to 1 (or even 145 to 1, as the jury awarded in *State Farm*). However, greater ratios could be consistent with the Due Process Clause where "a particularly egregious act has resulted in only a small amount of economic damages."[47] A stricter limit is imposed under Federal maritime law. In the *Exxon Valdez* case, the Court observed that in recent decades the median ratio of punitive to compensatory awards had remained less than 1 to 1, and that the percentage of cases with punitive awards had not markedly increased.[48] The Court concluded that awards at or below the empirically established median "would roughly express jurors' sense of reasonable penalties in cases with no earmarks of exceptional blameworthiness" which do not raise special under-enforcement problems.[49] Accordingly, and given the need for

40. *See infra* notes 45–47 and accompanying text.

41. State Farm Mut. Auto. Ins. Co. v. Campbell, 538 U.S. 408, 416 (2003); Cooper Indus., Inc. v. Leatherman Tool Group, Inc., 532 U.S. 424, 433–34 (2001); BMW of N. Am., Inc. v. Gore, 517 U.S. 559, 562 (1996); TXO Prod. Corp. v. All. Res. Corp., 509 U.S. 443, 453–455 (1993); *see also infra* Section VI.A.

42. 517 U.S. at 575–85.

43. *Id.* at 575–80.

44. Baker v. Exxon Mobile Corp., 490 F.3d 1066, 1085–89 (9th Cir. 2007).

45. *BMW*, 517 U.S. at 580–83.

46. 538 U.S. at 425.

47. *Id.* Perhaps TXO Prod. Corp. v. All. Res. Corp., 509 U.S. 443 (1993) is such an exceptional case: the court allowed a $10 million punitive award even though compensatory damages were only $19,000, arguably because the defendant "set out on a malicious and fraudulent course." *Id.* at 462.

48. Exxon Shipping Co. v. Baker, 554 U.S. 471, 497–500 (2008).

49. *Id.* at 512–13.

predictability, the Court held that a 1 to 1 ratio was a fair upper limit in such maritime cases.[50] Cases in which the defendant's conduct is more egregious than reckless are not so limited.[51] The third guidepost is the difference between the punitive damages award and the civil or criminal penalties authorized or imposed in comparable cases.[52]

II. JUSTIFICATIONS FOR PUNITIVE DAMAGES

A. From Compensatory to Non-Compensatory Rationales

One of the most fundamental principles in the modern law of torts is that damages should restore the victim to the condition he or she would have been in but for the tort (*restitutio in integrum*).[53] Punitive damages seem non-compensatory by definition. So although the doctrine's specific goals are yet to be explored, it appears inconsistent with the fundamental remedial goal.[54] A possible justification, which alleviates the conceptual problem, is the need to compensate for non-compensable losses such as hurt feelings, dignitary harm, and embarrassment.[55] The Supreme Court observed in *Cooper Industries, Inc. v. Leatherman Tool Group, Inc.*[56] that until the nineteenth century "punitive damages frequently operated to compensate for intangible injuries, compensation which was not otherwise available under the narrow conception of compensatory damages prevalent at the time."[57] However, the historical foundations of this argument seem dubious,[58] and at any rate, it has lost much of its appeal over time, with the

50. *Id.* at 513.

51. John P. Jones, *The Sky Has Not Fallen Yet on Punitive Damages in Admiralty Cases*, 83 TUL. L. REV. 1289, 1302 (2009).

52. BMW of N. Am., Inc. v. Gore, 517 U.S. 559, 583–85 (1996). For a while, it was unclear whether these guideposts should be explained to juries, or used solely by appellate courts discussing the constitutionality of punitive damages awards. But in Philip Morris USA v. Williams, 549 U.S. 346, 355 (2007), the court implied that juries should be advised of the proper guidelines.

53. RESTATEMENT (SECOND) OF TORTS § 901 (AM. LAW INST. 1979); *see also* United States v. Hathaley, 257 F.2d 920, 923 (10th Cir. 1958) ("The fundamental principle of damages is to restore the injured party, as nearly as possible, to the position he would have been in had it not been for the wrong of the other party."). *But see* John C.P. Goldberg, *Two Conceptions of Tort Damages: Fair v. Full Compensation*, 55 DEPAUL L. REV. 435 (2006) (arguing that the idea of "fair compensation" predated the current notion of "full compensation").

54. *See, e.g.*, Bennis v. Michigan, 516 U.S. 442, 469 n.13 (1996) (Stevens, J., dissenting) ("Tort law is tied to the goal of compensation (punitive damages being the notable exception).").

55. Exxon Shipping Co. v. Baker, 554 U.S. 471, 491–92 (2008) (discussing the compensatory rationale); David L. Walther & Thomas A. Plein, *Punitive Damages: A Critical Analysis:* Kink v. Combs, 49 MARQ. L. REV. 369, 370–71, 381 (1965) (same); Note, *Exemplary Damages in the Law of Torts*, 70 HARV. L. REV. 517, 520–21 (1951) (same).

56. 532 U.S. 424 (2001).

57. *Id.* at 437 n.11.

58. *See* Anthony J. Sebok, *What Did Punitive Damages Do? Why Misunderstanding the History of Punitive Damages Matters Today*, 78 CHI.-KENT L. REV. 163, 164, 204–05 (2003)

expansion of compensatory damages for intangible harms.[59] Somewhat related but less common was the argument that punitive damages were needed to compensate the victim for litigation costs, which were generally irrecoverable in American law.[60] This argument has garnered very little support in the case law. Thus, punitive damages are currently deemed separate and distinct from compensatory damages in most jurisdictions,[61] with very few exceptions.[62]

A doctrine that awards extra-compensatory damages is inconsistent with corrective justice theories that focus on rectification of the harm caused.[63] What then is the purpose of punitive damages? In the past, it was very frequently said that the aim of these awards was "to punish and deter." This phrase has been used in court decisions,[64] legal literature,[65] and jury instructions.[66] It is misleading because punishment is not a purpose but a mechanism; namely, the imposition of a sanction.[67] Imposing a sanction may have various goals, such as retribution, deterrence, appeasement of the victim, incapacitation of the wrongdoer, education, etc. Therefore, saying that punitive damages are meant to

("[P]unitive damages have never served the compensatory function attributed to them by the Court in *Cooper*.").

59. *See* Sloane, *supra* note 36, at 481–82 ("[T]he need for this function has waned with the onset of the award of damages for pain and suffering, mental distress, and most recently, hedonic damages.").

60. *See, e.g.*, Doroszka v. Lavine, 150 A. 692, 692–93 (Conn. 1930) ("[T]he purpose is not to punish the defendant . . . but to compensate the plaintiff . . . and so-called punitive or exemplary damages cannot exceed the amount of the plaintiff's expenses of litigation, less taxable costs."); James A. Breslo, Comment, *Taking the Punitive Damage Windfall Away from the Plaintiff: An Analysis*, 86 NW. U. L. REV. 1130, 1136 (1992) (reiterating this argument); Walther & Plein, *supra* note 55, at 381 (same).

61. Exxon Shipping Co. v. Baker, 554 U.S. 471, 492 (2008).

62. *See* Markey v. Santangelo, 485 A.2d 1305, 1308 (Conn. 1985) ("Punitive damages consist of a reasonable expense properly incurred in the litigation."); Peisner v. Detroit Free Press, Inc., 364 N.W.2d 600, 608–09 (Mich. 1984) (holding that punitive damages compensate for the "incremental injury to feelings attributable to defendant's malice.").

63. *See, e.g.*, ERNEST J. WEINRIB, THE IDEA OF PRIVATE LAW 135 n.25 (1995) (explaining that punitive damages are inconsistent with the corrective goal).

64. *See, e.g.*, BMW of N. Am., Inc. v. Gore, 517 U.S. 559, 568 (1996); Browning-Ferris Indus. of Vt., Inc. v. Kelco Disposal, Inc., 492 U.S. 257, 275 (1989); Newport v. Fact Concerts, Inc., 453 U.S. 247, 266–67 (1981); Gertz v. Robert Welch, Inc., 418 U.S. 323, 350 (1974); Moskovitz v. Mt. Sinai Med. Ctr., 635 N.E.2d 331, 343 (Ohio 1994); Hodges v. S. C. Toof & Co., 833 S.W.2d 896, 900 (Tenn. 1992); Hamilton Development Co. v. Broad Rock Club, Inc., 445 S.E.2d 140, 143 (Va. 1994).

65. *See, e.g.*, RESTATEMENT (SECOND) OF TORTS § 908(1) (AM. LAW INST. 1979); Howard A. Denemark, *Seeking Greater Fairness When Awarding Multiple Plaintiffs Punitive Damages for a Single Act by a Defendant*, 63 OHIO ST. L.J. 931, 935–36 (2002); Schwartz, *supra* note 5, at 134; Walther & Plein, *supra* note 55, at 372–73.

66. *See, e.g.*, RONALD W. EADES, JURY INSTRUCTIONS ON DAMAGES IN TORT ACTIONS 98 (3rd ed. 1993) (quoting a standard jury instruction whereby the purpose of punitive damages is to punish and deter).

67. *Cf.* George P. Fletcher, *The Place of Victims in the Theory of Retribution*, 3 BUFF. CRIM. L. REV. 51, 52 (1999) (explaining that retributive justice means "imposing punishment because it is deserved on the basis of having committed a crime.").

punish seems tautological. Yet the word "punishment" is often used more colloquially as a synonym for "retribution." And whenever courts say that punitive damages are intended to "punish and deter" they seem to suggest that punitive damages are aimed at retribution and deterrence. This is the most reasonable explanation for the classical punish/deter dichotomy.[68] In recent American decisions a more accurate terminology has been used. For example, in *State Farm Mutual Automobile Insurance Co. v. Campbell*[69] the Supreme Court held that punitive damages "are aimed at deterrence and retribution."[70] Case law is unclear as to whether extra-compensatory damages are justified only if they serve both goals, or if either one is sufficient.[71]

B. Deterrence

1. Outline

As explained above, courts have perceived deterrence as one of the two principal justifications for punitive damages.[72] The additional sanction may serve to deter the specific defendant from repeating the wrong and others from committing similar wrongs.[73] Two questions arise in this respect. The first is what types of wrongfulness punitive damages aim to deter. The simplistic answer, following Jeremy Bentham's utilitarian approach, conflates the end and the means. Deterrence is such an integral and distinctive feature of utilitarian and economic theories of law that a sanction must deter inefficient conduct. But American courts seem to have a somewhat different intuition. For instance, in *Cooper*,[74] the Supreme Court opined that the deterrent function of punitive damages was not exclusively efficiency-oriented: "Citizens and legislators may rightly insist that they are willing to tolerate some loss in economic efficiency in order to deter what they consider morally offensive conduct, albeit cost-beneficial morally offensive conduct . . . ."[75]

68. *See, e.g.*, Ellis, *supra* note 5, at 1, 4 (asserting that punitive damages serve deterrence and retribution); David G. Owen, *A Punitive Damages Overview: Functions, Problems and Reform*, 39 VILL. L. REV. 363, 373–74 (1994) (same); Sharkey, *supra* note 5, at 356–57 (same); Sunstein et al., *supra* note 5, at 2081 (same); Malcolm E. Wheeler, *The Constitutional Case for Reforming Punitive Damages Procedures*, 69 VA. L. REV. 269, 310 (1983) (same).

69. 538 U.S. 408 (2003).

70. *Id.* at 416–17; *see also* Exxon Shipping Co. v. Baker, 554 U.S. 471, 492 (2008); Pac. Mut. Life Ins. Co. v. Haslip, 499 U.S. 1, 19 (1991).

71. *See* Schwartz, *supra* note 5, at 134 (discussing this ambiguity).

72. *Supra* notes 64–66, 69–70 and accompanying text.

73. *See* Ellis, *supra* note 5, at 3 (discussing the deterrent effect of punitive damages).

74. Cooper Indus., Inc. v. Leatherman Tool Group, Inc., 532 U.S. 424 (2001).

75. *Id.* at 439–40 (quoting Marc Galanter & David Luban, *Poetic Justice: Punitive Damages and Legal Pluralism*, 42 AM. U. L. REV. 1393, 1450 (1993)).

The second question is why deterrence of unwarranted conduct requires extra-compensatory damages. As Gary Schwartz observed, those who point to deterrence as one of the main goals of punitive damages "are rendered almost useless by their obliviousness to the basic point that ordinary civil damages—in the course of providing compensation—concurrently function to deter."[76] Theorists have provided two types of answers to this challenge: one based on internalization of harm, the other on disgorgement of gain.[77]

2. Optimal Deterrence and Internalization of Harm

According to classic economic theory, efficient deterrence entails internalization by the wrongdoer of the social harm caused by his or her wrongful conduct. Only if the expected liability is equivalent to (or greater than) the expected externalized cost, will the potential injurer internalize that cost and take cost-effective precautions. Compensatory damages can secure optimal deterrence only if: (1) wrongdoers are always liable for harms caused by their wrongful conduct; and (2) compensatory damages are set in accordance with the social cost of the wrongful conduct. Punitive damages may be economically justified if one of these conditions is not met.

As regards the first condition, wrongdoers might escape liability for harms caused by their wrongful conduct.[78] Several factors—external to substantive tort law—let many negligent injurers off scot-free. Some victims are unaware that their harms are the result of wrongful conduct;[79] others do not have sufficient evidence to substantiate their case;[80] some may not sue because the expected costs of doing so outweigh the expected compensation,[81] and many refrain from suing if there is a familial, social, professional, or other connection between them and the injurer.[82] Thus, compensatory damages cannot induce potential injurers to take cost-effective precautions. According to economic theory, punitive damages may be used to overcome these under-enforcement problems.[83] Under this

76. Schwartz, *supra* note 5, at 137.

77. *See* Ellis, *supra* note 5, at 7 (presenting the two possible answers).

78. *See* Sharkey, *supra* note 2, at 488 (explaining that wrongdoers might escape liability).

79. *See* Gary T. Schwartz, *Empiricism and Tort Law*, 2002 U. ILL. L. REV. 1067, 1071 (2002) (discussing possible reasons for under-enforcement).

80. *Id.*

81. *Id.*

82. *See* Bruce Feldthusen, *The Civil Action for Sexual Battery: Therapeutic Jurisprudence?*, 25 OTTAWA L. REV. 203, 213 n.31, 214 & n.34, 218 (1993) (explaining that a personal relation between the injurer and the victim might prevent a lawsuit).

83. *See* WILLIAM M. LANDES & RICHARD A. POSNER, THE ECONOMIC STRUCTURE OF TORT LAW 160–65, 184–85, 223–24 (1987) (explaining that punitive damages can help overcome under-deterrence due to under-enforcement); STEVEN SHAVELL, ECONOMIC ANALYSIS OF ACCIDENT LAW

perception, total damages should be the actual harm multiplied by the reciprocal of the probability that the defendant *will* be found liable when he or she *should* be found liable; punitive damages must equal the excess of total damages over compensatory damages.[84]

This theory has relatively weak descriptive power. An under-enforcement rationale can hardly explain why punitive damages are unusual and hinge on the defendant's state of mind.[85] After all, under-enforcement exists in most cases, including those of negligence or no-fault liability—not only in cases of reprehensible conduct. Moreover, the extent of punitive damages is determined predominantly in accordance with the conduct's reprehensibility, not with the probability of escaping liability.[86] While the retributive yardstick is consistent with the constitutional framework (as we demonstrate below),[87] the economic model might entail departure therefrom. Experimental studies also reveal that in assessing punitive damages, potential jurors rely on the reprehensibility of the conduct and disregard detection rates.[88] These observations must be qualified, though, because heavier punitive awards have been deemed justifiable and constitutional when wrongdoing is hard to detect (increasing the chances of getting away with it),[89] or when the value of injury and the corresponding compensatory award are small (providing low incentives to sue).[90] Thus, arguably, although under-enforcement is not the primary factor, it is not totally ignored. The economic theory also raises serious practical concerns: it is extremely difficult to determine the actual detection rate on which the calculation is based.[91] Additionally, administrative and criminal law may be regarded as efforts to compensate

148 (1987) (same); Abraham & Jeffries, *supra* note 5, at 419 (same); Polinsky & Shavell, *supra* note 5, at 873–74 (same); Schwartz, *supra* note 5, at 139 (same); Sharkey, *supra* note 2, at 488–89 (same); Sharkey, *supra* note 5, at 365–68; Sunstein et al., *supra* note 5, at 2075, 2082 (same).

84. Cooter, *supra* note 5, at 1148 (discussing the multiplier method); Polinsky & Shavell, *supra* note 5, at 889, 911 (same). For further discussion of this method, see *supra* note 83.

85. *See* Benjamin C. Zipursky, *Rights, Wrongs, and Recourse in the Law of Torts*, 51 VAND. L. REV. 1, 95 (1998) (criticizing the economic theory of punitive damages); Schwartz, *supra* note 5, at 141–42 (same).

86. *But see* Mathias v. Accor Economy Lodging, Inc., 347 F.3d 672, 677 (7th Cir. 2003) (explaining that punitive damages serve to cancel out under-detection).

87. *See infra* Section II.C.

88. CASS R. SUNSTEIN ET AL., PUNITIVE DAMAGES: HOW JURIES DECIDE 135–39, 151–64 (2002).

89. BMW of N. Am., Inc. v. Gore, 517 U.S. 559, 582 (1996) ("A higher ratio may also be justified in cases in which the injury is hard to detect . . . .").

90. *Id.* ("[L]ow awards of compensatory damages may properly support a higher ratio . . . if, for example, a particularly egregious act has resulted in only a small amount of economic damages."); *see also* Exxon Shipping Co. v. Baker, 554 U.S. 471, 494, 513 (2008) (endorsing *BMW* on this point).

91. *See* Sharkey, *supra* note 2, at 496 (discussing the inability to determine the likelihood of detection).

for private under-enforcement of law; if tort law also aims to offset under-enforcement through punitive damages, over-deterrence may ensue.[92]

As regards the second condition, compensatory damages might not suffice if some of the actual harm caused by wrongful conduct is legally non-compensable.[93] Good examples are the loss of enjoyment of life in cases of wrongful death[94] or infliction of "insults and indignities."[95] In such cases, punitive damages may close the gap between the externalized social cost and compensable private harm.[96] This argument is admittedly problematic. From a descriptive perspective, the problem of under-compensation, like that of under-enforcement, is also characteristic of cases of negligence and strict liability, where punitive damages are not allowed. Moreover, in many jurisdictions punitive damages are precluded in cases of death, where they should be readily available according to the theoretical argument.[97] From a prescriptive perspective, if existing law fails to compensate for significant elements of harm, the proper strategy entails reforming or revising the law directly, "rather than straining for a surrogate result through reliance on punitive damages."[98]

3. Absolute Deterrence and Disgorgement of Gain

According to an alternative economic theory, compensatory damages might not provide the appropriate incentive when the wrongdoer derives illicit gains, financial or non-financial, from the wrongful conduct. In the leading English case of *Rookes v. Barnard*,[99] Lord Devlin opined that if a person expects personal profit from wrongdoing to exceed compensable harm to others, compensation is an insufficient deterrent: "Exemplary damages can properly be awarded whenever it is necessary to teach a wrongdoer that tort does not pay."[100] More recently, the Wisconsin Supreme Court held in *Jacque v. Steenberg Homes, Inc.*,[101] that punitive damages, by removing the profit from illegal activity, can help deter such

92. *See* Sunstein et al., *supra* note 5, at 2084.

93. *See, e.g.*, Ellis *supra* note 5, at 28 (contending that these non-compensable losses include relational economic and emotional harms).

94. *See* Sharkey, *supra* note 2, at 491 (discussing wrongful death).

95. *See* Schwartz, *supra* note 5, at 139 (discussing "insults and indignities").

96. *See* Mark A. Geistfeld, *Punitive Damages, Retribution, and Due Process*, 81 S. CAL. L. REV. 263 (2008) (considering wrongful death).

97. Schwartz, *supra* note 5, at 142–43.

98. *Id.* at 139–40; *see also* Polinsky & Shavell, *supra* note 5, at 939–41 (making this argument with respect to loss of enjoyment of life in wrongful death cases).

99. [1964] AC 1129 (HL) (appeal taken from Eng.).

100. *Id.* at 1227; *see also* Green Oil Co. v. Hornsby, 539 So. 2d 218, 223 (Ala. 1989) ("If the wrongful conduct was profitable to the defendant, the punitive damages should remove the profit and should be in excess of the profit, so that the defendant recognizes a loss.").

101. 563 N.W.2d 154, 165 (Wis. 1997).

conduct; "the defendant recognizes a loss" only when punitive damages exceed the profit created by the misconduct. This argument has found support in the law and economics literature.[102]

From an economic perspective, eliminating the wrongdoer's gains may be warranted in two types of cases. First, when the activity is undesirable and must be avoided altogether.[103] According to Polinsky & Shavell, this can happen when the injurer's gain is socially illicit, and therefore should not be taken into account in the social welfare calculation.[104] To the extent that the legitimate gain (which might be nil) is lower than the harm, the act should be deterred completely. Yet compensatory damages will not provide the necessary deterrence if the injurer's total gain exceeds the compensable harm. Complete deterrence requires disgorgement of the gain.[105] For example, if A's activity generates a $5,000 gain for A and a $1,000 loss to B, but A's activity is undesirable (the gain is socially illicit), compensatory damages will not prevent A from committing the wrongful conduct. Gains from harmful conduct may be socially illicit "when the injurer's utility derives from causing harm itself, as when a person punches another out of spite or defames another to see him suffer."[106] An activity may also be undesirable when the injurer's gain is greater than compensable harm, but the conduct has additional negative implications not addressed by tort law. Here too, complete deterrence is required, and may be achieved through disgorgement of gain.

Second, when transaction costs between the potential injurer and the potential victim are low, a negotiated ex ante agreement with respect to the interest at risk is more efficient than ex post tort litigation. The administrative costs of an agreement are lower than those of litigation, and an agreement better reflects subjective valuations than a court ruling. A compensatory damages model induces the potential injurer to unilaterally cause harm when the net personal gain exceeds that harm, although a negotiated agreement is preferable. Punitive damages based on the injurer's gain induce the potential injurer to seek voluntary transfer of entitlement, thereby giving effect to property rights.[107]

---

102. *See, e.g.*, Robert D. Cooter, *Economic Analysis of Punitive Damages*, 56 S. CAL. L. REV. 79, 98 (1982) ("[P]unitive damages should be computed at a level that offsets the illicit pleasure of noncompliance or the exceptional costs of compliance that motivated the injurer."); Hylton, *supra* note 5 *passim*.

103. Hylton, *supra* note 5, at 933–34; Sharkey, *supra* note 2, at 489, 492.

104. Polinsky & Shavell, *supra* note 5, at 909.

105. *Id.*

106. *Id.* Note, however, that one might challenge the exclusion of "illicit gains" from aggregate welfare calculations. *See, e.g.*, Gary S. Becker, *Crime and Punishment: An Economic Approach*, 76 J. POL. ECON. 169, 173 (1968) (assigning social value to offenders' gains).

107. Sharkey, *supra* note 2, at 493. Such incentives are known as "kickers" in criminal law and economics literature. Jules L. Coleman, *Crime, Kickers, and Transaction Structures*, *in* 27 CRIMINAL

Like harm-internalization theories, the gain elimination-model also has little support in the case law. The defendant's financial or non-financial gain is not correlated with reprehensibility. Additionally, it can neither explain nor justify the single-digit ratio guideline, which does not seem to have deterrence-oriented or retributive explanations at all.[108] Still, the Supreme Court in *Haslip*[109] considered "the profitability to the defendant of the wrongful conduct and the desirability of removing that profit and of having the defendant also sustain a loss" relevant in determining whether a punitive damages award was excessive. Note further that illicit gains are often handled through criminal law; in such cases, punitive damages may be justified only if for some reason the criminal system malfunctions.[110]

C. Retribution

Retribution may be defined as imposing a sanction that corresponds to individual moral desert.[111] The wrongdoer deserves to be punished on account of her wrongful conduct, and ought to be punished fairly regardless of the consequences of her punishment.[112] Thus, the concept of retribution entails two components: wrongfulness and just desert. The content of the wrong, for which a punishment is deserved, is not consensual, nor does it have to be. Retribution is a *form* of justice, not an independent *substantive* moral standard: It determines how justice should be done whenever a wrong is committed. The definition of wrongfulness is left to legal philosophers.

Retributive justice does not require that the sanction be identical to the wrong, unlike the primeval *lex talionis*.[113] Rather, it insists on proportionality between the severity of the sanction and the gravity of the wrong.[114] The sanction must meet the complementary tests of cardinal

---

JUSTICE: NOMOS 313, 318–19 (J. Ronald Pennock & John W. Chapman, eds., 1985) (discussing this literature).

108. Sharkey, *supra* note 2, at 495.

109. Pac. Mut. Life Ins. Co. v. Haslip, 499 U.S. 1, 21–22 (1991).

110. Schwartz, *supra* note 5, at 138–39.

111. *See* Fletcher, *supra* note 67, at 52 ("[Retributive justice means] imposing punishment because it is deserved on the basis of having committed a crime.").

112. *Id.* Retributive justice is therefore *retrospective*, in that it looks backward to the particular wrongdoing, not forward to the consequences of the sanction.

113. "The law of retaliation, which requires the infliction upon a wrongdoer of the same injury which he has caused to another." *Lex talionis*, BLACK'S LAW DICTIONARY (6th ed. 1990).

114. *See, e.g.*, TONY HONORÉ, RESPONSIBILITY AND FAULT 13, 83–84, 92, 123, 138 (1999) (explaining that retribution requires proportionality between the severity of the sanction and the gravity of the misconduct); Peter Cane, *Retribution, Proportionality, and Moral Luck in Tort Law*, *in* THE LAW OF OBLIGATIONS 141, 143, 160–61 (Peter Cane & Jane Stapleton eds., 1998) (same); Tony Honoré, *The Morality of Tort Law – Questions and Answers*, *in* PHILOSOPHICAL FOUNDATIONS OF TORT LAW 73, 87 (David G. Owen ed., 1995) (same); Jean Hampton, *Correcting Harms Versus Righting Wrongs: The Goal of Retribution*, 39 UCLA L. REV. 1659, 1690 (1992) (same); Jeffrie G.

(absolute) and ordinal (relative) proportionality.[115] According to the former, the sanction should not be too harsh or too lenient with respect to the *absolute* gravity of the wrong committed.[116] For example, it is unfair to impose a life sentence on a person who evaded a parking charge; similarly, it is unfair to impose a small fine on a cold-blooded murderer. According to the principle of ordinal proportionality, the sanction imposed for a certain wrong must reflect the *relative* gravity of the wrong: if wrong A is graver than wrong B, the sanction for wrong A must be more severe than the sanction for wrong B, and vice versa.[117] For example, the punishment for murder must always be more severe than the punishment for non-payment for parking. The principle of cardinal proportionality sets the upper and lower limits of the possible sanction in a given society regardless of the wrong's relative gravity. The principle of ordinal proportionality narrows these boundaries, so that the order of harshness of actual sanctions will correspond to the order of gravity of the given wrongs.[118]

The principle of proportionality is usually justified in terms of fairness.[119] However, as Cesare Beccaria and Jeremy Bentham noted, proportionality can also be defended in terms of marginal deterrence.[120] Put differently, even if the justice system is incapable of deterring all wrongs, it should aspire at least to prevent the wrongdoer from committing greater wrongs. Consider, for instance, an intoxicated driver who hits and injures a pedestrian. If the sanction is independent of the driver's response to the accident, he might be incentivized to escape from the scene to avoid liability. Escaping makes the offense graver, because it deprives the victim of potential help which can reduce the risk to life and limb. Imposing a more severe sanction for a "hit and run" offense will achieve marginal deterrence: even though a person is willing to drive intoxicated despite the risk of liability for injuring another, he or she may not be willing to pay the additional price of escaping. Similarly, consider a case of fraud. If the

---

Murphy, *Does Kant Have a Theory of Punishment?*, 87 COLUM. L. REV. 509, 530–32 (1987) (same); Note, *Punitive Damages and Libel Law*, 98 HARV. L. REV. 847, 851 (1985) (same).

115. *See* ANDREW VON HIRSCH, CENSURE AND SANCTIONS 29–46 (1993); ANDREW VON HIRSCH, PAST OR FUTURE CRIMES: DESERVEDNESS AND DANGEROUSNESS IN THE SENTENCING OF CRIMINALS 40–46 (1985); Cane, *supra* note 114, at 143, 161; Ellis, *supra* note 5, at 6–7.

116. *See* Cane, *supra* note 114, at 143.

117. *See id.*

118. *See* Ronen Perry, *The Role of Retributive Justice in the Common Law of Torts: A Descriptive Theory*, 73 TENN. L. REV. 177, 182–83 (2006) (discussing the distinction between cardinal and ordinal proportionality).

119. *See* Schwartz, *supra* note 5, at 134 (explaining that the goal of "punishment" is founded on conceptions of fairness).

120. CESARE BECCARIA, AN ESSAY ON CRIMES AND PUNISHMENTS 28–32 (Edward D. Ingraham trans., 1819) (1764); 2 JEREMY BENTHAM, AN INTRODUCTION TO THE PRINCIPLES OF MORALS AND LEGISLATION 14–15, 20 (2nd ed. 1823).

sanction for fraud is constant, the offender is incentivized to commit a fraud of a greater scale, with more victims and greater gains.

Wherever recognized, punitive damages are reserved for exceptional cases. Awarding non-compensatory damages is inconsistent with the corrective structure of tort law. Therefore, punitive damages are awarded merely on the margins, where courts feel that a compensatory award is an *extremely lenient* sanction with regard to the gravity of the defendant's conduct.[121] In other cases, the discrepancy between the gravity of the wrong and the severity of the sanction is left untouched for the sake of corrective justice. As observed by the Supreme Court, it should be presumed that a plaintiff has been made whole for his or her injuries by compensatory damages; punitive damages should only be awarded if the defendant's culpability, after having paid compensatory damages, is so reprehensible as to warrant the imposition of further sanctions to achieve retribution or deterrence.[122]

Of deterrence and retribution, the latter seems the dominant goal. It was held that an award of punitive damages must reflect the gravity of the respective wrong,[123] and that punitive damages are subject to the Due Process Clause, which prohibits the imposition of grossly excessive or arbitrary penalties.[124] Substantive due process review follows the *BMW v. Gore* guideposts.[125] The first guidepost—degree of reprehensibility, which is also the most important one,[126] can be easily explained in retributive terms. Determining the severity of the sanction in light of the degree of reprehensibility—absolute and relative—is essentially an exercise of retributive justice.[127] The third guidepost (relation to civil penalties in comparable cases) seems to aim at ordinal proportionality: similar wrongs deserve similar sanctions. The second guidepost (ratio of punitive damages to actual harm) may also be linked to the notion of retributive justice. Although it is separate from the first, the potential harm is in fact one of the primary indicators of the gravity of the defendant's conduct. However, the second guidepost is apparently intended to prevent the

121. *See, e.g.*, David Luban, *A Flawed Case Against Punitive Damages*, 87 GEO. L.J. 359, 364 (1998).

122. State Farm Mut. Auto. Ins. Co. v. Campbell, 538 U.S. 408, 419 (2003); BMW of N. Am., Inc. v. Gore, 517 U.S. 559, 575 (1996).

123. *See, e.g.*, Day v. Woodworth, 54 U.S. 363, 371 (1851).

124. *See supra* note 41; *infra* Section VI.A.

125. *See supra* notes 42–52 and accompanying text.

126. *BMW*, 517 U.S. at 575 ("[T]he most important indicium of the reasonableness of a punitive damages award is the degree of reprehensibility of the defendant's conduct."); Jacque v. Steenberg Homes, Inc., 563 N.W.2d 154, 164 (Wis. 1997) (quoting *BMW*).

127. Sharkey, *supra* note 2, at 495; *see also* Schwartz, *supra* note 5, at 144 (observing that the "punitive" nature explains "why the precise 'degree of reprehensibility' in the defendant's conduct is a key element in calculating the quantum of punitive damages").

plaintiff from acquiring an unreasonable windfall, and possibly to generate some certainty with respect to the extent of damages, not to ensure that the defendant's sanction is proportional. In any event, the first—retributive—guidepost is the dominant one.[128]

The dominance of the retributive justification is apparent not only in judicial rhetoric, but also in jury perceptions of the goals of punitive damages. Empirical evidence suggests that juries do not attempt to promote optimal deterrence but to "punish" wrongdoing, with at most a signal designed to ensure that certain misconduct will not happen again.[129] Ordinary people do not spontaneously think in terms of optimal deterrence when asked about appropriate punishment, and it is very hard to oblige them to think in these terms.[130] People assume the role of jurors with retributive intuitions, and it remains unclear whether and to what extent the courtroom can overcome them.[131]

Two comments are necessary at this point. First, egregious conduct may give rise to various types of sanctions, legal (criminal, administrative, civil, or disciplinary) and extra-legal (such as reputational harm). To avoid disproportion between the overall burden imposed on the defendant and the gravity of her wrong, these sanctions must be taken into account in deciding whether punitive damages may be awarded in a specific case, and in determining their amount.[132] Some jurisdictions have barred punitive damages in a civil action following or pending criminal conviction for the same conduct.[133] However, the prevailing view in the US is that criminal conviction does not bar punitive damages, although it should be taken into account in determining the size of the award.[134]

Second, an interesting question concerns judicial readiness to award punitive damages against firms. Firms are not persons, and when punitive damages are imposed on them, those who suffer may not be wrongdoers at

128. Experimental studies demonstrate that jurors are motivated primarily by a sense of moral outrage at reprehensible conduct. SUNSTEIN ET AL., *supra* note 88, at 236–38.

129. Sunstein et al., *supra* note 5, at 2085.

130. *Id.* at 2111.

131. *Id.*

132. *See* Schwartz, *supra* note 5, at 145 ("[A] punitive damage penal theory must explain why the criminal law does not adequately fulfill society's punishment objectives.").

133. *See* Walther & Plein, *supra* note 55, at 384; *see also* Perm. Civ. App. 9670/07 Jane Doe v. John Doe (unpublished, 2009) (Isr.) (holding that if the wrongdoer has already been sentenced for the same misconduct in a criminal process, punitive damages will be utterly exceptional).

134. *See, e.g.*, Pac. Mut. Life Ins. Co. v. Haslip, 499 U.S. 1, 22 (1991) (holding that the imposition of criminal sanctions on the defendant for its conduct should be considered in assessing the reasonableness of a punitive damages award); *In re* Exxon Valdez, 270 F.3d 1215, 1226 (9th Cir. 2001) ("[A] prior criminal sanction does not . . . bar punitive damages."); RESTATEMENT (SECOND) OF TORTS § 908 cmt. a (AM. LAW INST. 1979) ("The awarding of punitive damages is not prevented by a prior criminal conviction for the same act, which is relevant only to the amount of the award . . . ."); Annotation, *Assault: Criminal Liability as Barring or Mitigating Recovery of Punitive Damages*, 98 A.L.R.3d 870 (1980).

all.[135] A punitive damages award may end up injuring not the firm so much as stockholders (if the firm absorbs the cost or a fraction of it, and profits consequently plummet), consumers (if the firm spreads the cost ex post through price increases), and employees (if the firm is forced to cut back, or goes into liquidation), who had nothing to do with the underlying wrong. It is far from clear that juries awarding punitive damages are aware of this point or can be easily convinced of its validity.[136] At any rate, this article focuses on individual defendants, so any concerns arising from the imposition of monetary sanctions on corporate defendants should be reserved for future research.

III. TAKING INTO ACCOUNT THE DEFENDANT'S INCOME

Many state courts and legislatures have recognized the relevance of the defendant's wealth in determining the extent of punitive damages,[137] and this is often reflected in applicable jury instructions.[138] This Part demonstrates that taking the defendant's wealth into account may be consistent with the twin goals of punitive damages.

A. Deterrence

Intuitively, the deterrent effect of a monetary sanction depends on the offender's wealth: "A six-figure punitive judgment against a multi-billionaire would have the same deterrent impact as a parking ticket to the average person."[139] However, the classic economic argument is that the

135. *See* Schwartz, *supra* note 5, at 144 ("[T]he entire notion of punishment-as-punishment becomes deeply problematic when applied to the corporate form.").

136. Sunstein et al., *supra* note 5, at 2114 n.157.

137. *See, e.g.*, FLA. STAT. § 768.72(1) (2017); IOWA CODE § 668A.1 (2017); KAN. STAT. ANN. § 60-3701(b)(6) (2017); MINN. STAT. § 549.20(3) (2017); MISS. CODE ANN. § 11-1-65(1)(f)(ii)(3) (2017); MO. ANN. STAT. § 510.263(2) (2017); NEV. REV. STAT. ANN. § 42.005(4) (2017); OHIO REV. CODE ANN. § 2307.80(B)(6) (2017); Green Oil Co. v. Hornsby, 539 So. 2d 218, 222 (Ala. 1989) ("The defendant's financial position is . . . a consideration essential to a post-judgment critique of a punitive damages award."); Neal v. Farmers Ins. Exch., 582 P.2d 980, 990 (Cal. 1978) ("Also to be considered is the wealth of the particular defendant; obviously, the function of deterrence . . . will not be served if the wealth of the defendant allows him to absorb the award with little or no discomfort."); Liggett Group Inc. v. Engle, 853 So. 2d 434, 458 (Fla. Dist. Ct. App. 2006) ("A defendant's financial capacity is a crucial factor in determining the appropriateness of a punitive damages award."), *quashed in irrelevant part*, 945 So. 2d 1246 (Fla. 2006); Tetuan v. A.H. Robins Co., Inc., 738 P.2d 1210, 1242 (Kan. 1987) (holding that the punitive award was not excessive in light of the defendant's sales and earnings); Palmer v. A.H. Robins Co., Inc., 684 P.2d 187, 216, 221 (Colo. 1984) (holding that the punitive award was not excessive in light of the "financial ability of the offender to pay").

138. ARK. MODEL JURY INSTRUCTIONS—CIVIL § 2218 (2016) ("In arriving at the amount of punitive damages you may consider the financial condition of (defendant), as shown by the evidence."); CAL. CIVIL JURY INSTRUCTIONS § 3947, 3949 (2017) (requiring consideration of the defendant's financial condition in determining the amount necessary to punish and deter); WIS. JURY INSTRUCTIONS: CIVIL § 1707.1 (2017) (same).

139. Hylton, *supra* note 5, 927, 928.

defendant's wealth is irrelevant for deterring socially undesirable conduct.[140] Recall that punitive damages may serve two economic purposes.[141] First, they may secure internalization of the social costs of wrongful conduct, thereby affecting potential injurers' levels of precaution without eradicating socially beneficial activities. For example, if the injurer causes the victim a $100 loss, but the probability of finding the injurer liable is 0.25, total damages in the amount of $100/0.25=$400 (compensatory damages of $100, and punitive damages of $300) will secure internalization. Second, punitive damages may eliminate defendants' gains from wrongful conduct, thereby eradicating socially undesirable activities. For example, if the injurer imposes a loss of $100 on the victim and gains $500 from the conduct, but the activity is undesirable, the court should impose a compensatory award of $100 and a punitive award of at least $400 to eliminate the gain and eradicate the activity.

If the purpose of punitive damages is to internalize the social costs of wrongful conduct, the defendant's wealth seems irrelevant in the determination of a punitive judgment. The compensatory and punitive components of damages, which together internalize the full social costs of the defendant's conduct, are both independent of the defendant's wealth.[142] And whether the defendant is wealthy or poor, his or her personal cost-benefit analysis (expected liability versus the cost of avoiding it) is the same. Thus, the defendant's wealth seems irrelevant to efficient deterrence. The multi-billionaire and the pauper must pay the same punitive judgment.[143] Similarly, if the defendant is driven by a desire to obtain a certain profit or to avoid a certain cost, the punitive judgment can be set at a level that prevents him or her from obtaining that profit or avoiding that cost. As long as the profit or cost can be defined independently of the defendant's wealth, there should be no need to consider or even know the defendant's wealth in determining the proper punitive award.[144]

However, a more sophisticated analysis demonstrates that in some circumstances wealth may be relevant.[145] To begin with, the wrongdoer's wealth should be taken into account when "either the victim's loss or the defendant's gain from wrongdoing is unobservable and correlated with the

140. Abraham & Jeffries, *supra* note 5, at 415; Polinsky & Shavell, *supra* note 5, at 913.

141. *See also* Hylton, *supra* note 5, at 932–934 (discussing the two economic justifications).

142. Cooter, *supra* note 5, at 1176–77; Polinsky & Shavell, *supra* note 5, at 913.

143. Abraham & Jeffries, *supra* note 5, 417 and *passim*; Brief for Mitchell Polinsky et al. as Amici Curiae Supporting Petitioner, Philip Morris USA v. Williams, 549 U.S. 346 (2011) (No.05-1256), 2007 U.S. LEXIS 1332.

144. Hylton, *supra* note 5, at 929.

145. *Id.* at 930.

defendant's wealth."[146] In such cases, the defendant's wealth serves as a proxy for a particular variable which is relevant in the assessment of punitive damages.

If we use punitive damages to secure full internalization of the social costs of wrongful conduct, information on the defendant's wealth may be relevant when the external costs are not observable and wealth can be used as a proxy or help estimate external costs. But since in most cases the external costs are directly observable, information on the defendant's wealth will not be necessary to ensure full internalization.[147]

If we use punitive damages to disgorge gains and eradicate undesirable activity, information on the defendant's wealth may be relevant when the gain is not observable and is correlated with the defendant's wealth.[148] The wrongful gain is unobservable if it cannot be evaluated by an exogenously determined price, and it is correlated with wealth to the extent that the willingness to pay for obtaining it increases with wealth. Specifically, when a person spitefully causes harm to another *or* intentionally appropriates the other's unique property for the injurer's own enjoyment, the gain is equal to the maximum the injurer would be willing to pay to impose the same loss on the victim. In such cases, no objective signals of the injurer's gain are available. However, it is highly plausible that the injurer's gain (willingness to pay) is correlated with his or her wealth, because the value of money tends to decline as wealth increases.[149] Put differently, in cases of spiteful pleasure or intentional appropriation the gain is unobservable and correlated with wealth, so punitive damages should be correlated with the defendant's wealth.[150] For example, to offset the utility a rich person would obtain from slandering someone he or she disliked, we might need to impose $10,000 in punitive damages, whereas to deter a person with only modest assets, $1,000 in punitive damages might suffice.[151] In sum, because the victim's loss is typically observable, wealth will tend to be a relevant factor only when the legal system aims to eliminate the defendant's gain and eradicate the wrongful activity.

In addition to the case in which the defendant's illicit gain is unobservable and correlated with his or her wealth, the wrongdoer's wealth may also be relevant in assessing punitive damages if (1) potential injurers are risk-averse, and (2) they do not have access to liability insurance covering punitive damages. In such cases, appropriate

146. *Id.* at 930, 937.
147. *Id.* at 936.
148. *Id.* at 936–38.
149. *Id.* at 938–42; Polinsky & Shavell, *supra* note 5, at 914.
150. *See also* Abraham & Jeffries, *supra* note 5, at 418 (explaining that wealth may be relevant in setting punitive damages awards to deter those who cause harm out of spite or malice).
151. Polinsky & Shavell, *supra* note 5, at 914.

deterrence will be accomplished with a lower level of damages than if potential injurers are risk-neutral. The more risk-averse an individual is, the lower the optimal level of damages. Assuming that less wealthy individuals are more risk-averse than the wealthier ones, the optimal level of punitive damages will be lower for poorer individuals. Thus, punitive damages should be higher for wealthier individuals. However, if the purpose of punitive damages is to secure internalization of harm, then even the wealthiest individuals should not be obliged to pay punitive damages exceeding the level determined by the multiplier formula.[152]

Note that from an economic perspective, wealth should play no role in the determination of a punitive judgment against corporations. The reason is that a corporation will be rationally motivated by the incremental monetary benefits and costs of its actions, and those benefits and costs are independent of the corporation's wealth.[153] Corporations do not act out of noneconomic motivations which sometimes direct the conduct of individuals, such as a desire to experience a spiteful pleasure,[154] and they are generally risk-neutral. If damages payable by a corporation exceed the harm multiplied by the reciprocal of the probability of enforcement, the corporation will be led to take excessive precautions.[155] Moreover, imposing punitive damages that depend on wealth imposes a tax on corporate size and success, discouraging growth and development, especially in industries involving high liability costs (such as the pharmaceutical and aircraft industries).[156]

B. Retribution

In recent years, several scholars have opined that retributive justice must take into account the subjective experience of punishment.[157] The idea is that wrongdoers may experience sanctions differently, and that their punishment should reflect these experiential differences to achieve proportionality between the gravity of the wrong and the severity of the sanction.[158] If wrongdoers suffer more or less than deserved because of

152. *Id.* at 913.

153. Hylton, *supra* note 5, at 944–45; Polinsky & Shavell, *supra* note 5, at 910–14.

154. Abraham & Jeffries, *supra* note 5, at 418–19; Hylton, *supra* note 5, at 945; Polinsky & Shavell, *supra* note 5, at 914. Corporate officials might of course act out of noneconomic motivations. In such cases, individually assessed punitive damage may be due. Presumably, these officials can expect even more severe sanctions in the realm of employment.

155. Polinsky & Shavell, *supra* note 5, at 911. This is apparently true only for strict liability.

156. *Id.* at 911–12.

157. Dan Markel & Chad Flanders, *Bentham on Stilts: The Bare Relevance of Subjectivity to Retributive Justice*, 98 CALIF. L. REV. 907, 915–16 (2010).

158. *See generally*, Adam J. Kolber, *The Subjective Experience of Punishment*, 109 COLUM. L. REV. 182 *passim* (2009). Bronsteen, Buccafusco, and Mazur adopt the same normative assumption but focus on the impact of "hedonic adaptation" on the experienced severity of the punishment, and

their experience and preferences, the differences between the deserved sanctions and the perceived sanctions are unjust from a retributive perspective.[159] Strict retributive subjectivism, advocated by Adam Kolber, requires calibration of punishment to each wrongdoer's idiosyncratic psychological profile—experiences, capacities, and baselines.[160] This version has been contested in legal literature.[161] Yet there is little disagreement with the view that from a retributive perspective, "punishment ranges or modes should draw on generalizations informed by human psychology."[162]

Strict subjectivists and more traditional retributivists seem to agree that if monetary sanctions aim at retributive justice, they cannot be independent of wrongdoers' wealth.[163] The community aims to make the wrongdoer suffer in a manner proportional to the wrong done. To this end, we need to estimate the effect of sanctions on well-being, and to assess the impact of monetary sanctions, we need to know how wealthy wrongdoers are.[164] The Supreme Court correctly observed in *BMW* that "a fixed dollar award will punish a poor person more than a wealthy one."[165] This understanding hinges on the principle of the diminishing marginal utility of individual wealth, depicted in Figure 1. The intuitive explanation for this principle is that any addition to a person's wealth is used to satisfy less important and less urgent needs.[166]

---

conclude that such adaptation mitigates the differences between more and less severe sentences. John Bronsteen et al., *Happiness and Punishment*, 76 U. CHI. L. REV. 1037, 1070–71 (2009) [hereinafter Bronsteen et al., *Happiness and Punishment*]; *see also* Bronsteen et al., *Retribution and the Experience of Punishment*, 98 CALIF. L. REV. 1463, 1469 (2010) [hereinafter Bronsteen et al., *Retribution*] ("[T]he decisions about which punishments to impose and the decisions associated with proportionality must depend on knowledge about how those punishments are typically experienced by offenders.").

159. Bronsteen et al., *Happiness and Punishment*, *supra* note 158, at 1039, 1069–72; Kolber, *supra* note 158, at 188, 216, 236.

160. Kolber, *supra* note 158, at 185–87, 216, 236.

161. *See, e.g.*, Markel & Flanders, *supra* note 157 *passim* (challenging the subjective accounts of retribution); Kenneth W. Simons, *Retributivists Need Not and Should Not Endorse the Subjectivist Account of Punishment*, 109 COLUM. L. REV. SIDEBAR 1 *passim* (2009) (same).

162. Markel & Flanders, *supra* note 157, at 956; *see also* Bronsteen et al., *Retribution*, *supra* note 158, at 1464 ("[T]o the extent that [a certain psychological phenomenon] affects the experience of punishment that the typical person is expected to have, [it] is relevant to punishment.").

163. *See, e.g.*, Bronsteen et al., *Retribution*, *supra* note 158, at 1475 ("The reason that noncontingent fines communicate insufficient condemnation is that they are typically experienced as insufficiently negative when they are very small relative to the offender's wealth."); Kolber, *supra* note 158, at 226 (making a similar argument); Markel & Flanders, *supra* note 157, at 956–57 (same); Simons, *supra* note 161, at 6 n.11 ("Kolber aptly criticizes, as crude and unjustifiable, the widespread current practice of imposing an invariant, one-size-fits-all criminal fine for a particular crime, without regard to the offender's wealth or income."). For Jeremy Bentham's view on this matter, see *infra* note 180.

164. Perry, *supra* note 118, at 188 n.57.

165. BMW of N. Am., Inc. v. Gore, 517 U.S. 559, 591 (1996).

166. SHAVELL, *supra* note 83, at 186.

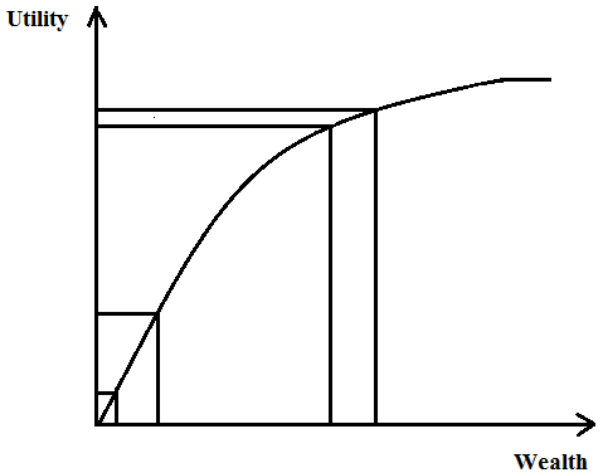

Figure 1. Utility Function

The private utility function does not have an accurate mathematical formulation. Steven Shavell chose to demonstrate the diminishing marginal utility of wealth with a negative exponential function of this form: $y=a(1-e^{(-x/b)})$ (x representing wealth and y representing utility).[167] But in our view this choice is somewhat problematic because it assumes an upper limit to the private utility of wealth (the asymptote $y=a$ in the positive region). An increasing concave function with no horizontal asymptotes, such as $y=a\times log(x/b+1)$, would give a more accurate description of the phenomenon.[168] Persons whose private utility functions are concave are risk-averse, although characterizing a person by his or her attitude to risk is less relevant in ex post application of retributive justice, when the wrongdoer has to bear a certain burden and not a probabilistic risk.

167. *Id.* at 188–89.

168. A logarithmic function $y=a\times log(x/b+1)$ displays four important properties. First, no wealth generates no utility. In mathematical terms, the function passes through the origin. Second, more wealth always generates more utility. In mathematical terms, the function is increasing. Third, each additional $1 has a lower perceived value. This is the diminishing marginal utility of wealth. In mathematical terms the function is concave in the positive region. Fourth, although marginal utility decreases, the utility that people can derive from wealth is unbounded. In mathematical terms the function has no maxima or asymptote. The negative exponential function $y=a(1-e(-x/b))$ has the first three properties but lacks the fourth. For a discussion of this feature see, e.g., Peter C. Fishburn, *Unbounded Utility Functions in Expected Utility Theory*, 90 Q. J. ECON. 163 (1976). In addition to its formal appropriateness, the logarithmic function is simple to understand and easy to use in numerical examples.

Now assume Moe, Larry, and Curly, have a similar utility function, *y=10,000log(x/10,000+1)*. Moe has $10,000,000, Larry $1,000,000, and Curly $100,000. Each commits the same intentional wrong. Imposing a monetary sanction of $100,000 for this wrong will result in a loss of 10,414 utility units for Curly (the value of his entire wealth), 452 utility units for Larry, and only 44 for Moe.[169] If punishment is designed to inflict a certain loss of utility on the actor, a wealthy wrongdoer should be required to bear a larger monetary sanction than an indigent wrongdoer.[170] To the extent that punitive damages serve a retributive goal, it makes sense to consider the defendant's wealth in calculating the punitive award.[171]

To be sure, the shape of individual utility functions (the values of *a* and *b* in the logarithmic formulation) varies. Clearly, liability for $10,000 harms a person more than liability for $1,000, but whether liability for $10,000 harms Jack more than it harms Jill cannot be ascertained based on their relative wealth. The wealthier defendants may value their money more than the less wealthy; hence the assumption that larger penalties are always needed to inflict the same impact on wealthier defendants is somewhat speculative.[172] However, the fact that individual utility functions vary does not undermine the understanding that marginal utility normally decreases, and that the utility function is concave for most people. Generally, therefore, the wrongdoer's wealth is expected to have a real impact on the experienced severity of the monetary sanction.

The remaining question is how relative wealth should be reflected in the severity of the punishment. Strict subjectivists might require consideration of individual utility curves. But this is impractical and theoretically unsound.[173] More traditional retributivists will probably endorse an ex ante, objective, rule-based determination of monetary sanctions,[174] integrating the need for proportionate punishment with the understanding that the experienced severity of these sanctions hinges on the wrongdoer's wealth. This method is practical, and while the outcome might be imprecise, it will be fair—in the sense of roughly treating the equals equally, and the unequals unequally, and also in the sense of certainty and predictability of punishments.

169. Utility losses are calculated as follows (*x* being the wrongdoer's initial wealth): *10,000log(x/10,000+1)−10,000log((x−100,000)/10,000+1)*.

170. Hylton, *supra* note 5, at 928 n.2.

171. Schwartz, *supra* note 5, at 144.

172. Abraham & Jeffries, *supra* note 5, at 421.

173. As explained above, strict subjectivism is rejected by most theorists. *See supra* Section III.B.

174. Markel & Flanders, *supra* note 157, at 957, 987.

IV. THE PROPOSED MODEL

A. The Criminal Law Lead

Criminal lawmakers and scholars have long realized that monetary sanctions such as fines have several advantages over incarceration. Most importantly, fines have lower administrative costs, they are transferred to the state thereby increasing rather than decreasing its funds, and they compensate society in addition to punishing offenders.[175] With this understanding, lawmakers and scholars have tried to develop an optimal fining method for the criminal justice system. The most common method is fixed fines—where the extent of the fine is based on the severity of the offense. This method has two apparent flaws. First, wealthier individuals might be less responsive to changes in fine amounts, due to the diminishing marginal utility of income. This theoretical insight has empirical support.[176] The deterrent effect of fixed fines might therefore be inconsistent. Second, because a fixed fine for a particular offense has a different impact on people with different income, the severity of the sanction imposed on the offender is in fact not correlated with the gravity of the offense. This violates the principle of retribution. In other words, if two offenders perceive the monetary sanction imposed for a particular offense differently due to a difference in personal wealth, then for at least one of them the sanction is not truly proportionate to the offense.

A unique feature of several European criminal law systems is the use of a day-fine model. A day-fine system systematically integrates not only the gravity of the offense, but also the offender's wealth. It is based on a two-stage process. In the first stage, the court assesses the severity of the offense, and determines the corresponding "number of days" for which a fine must be paid: The graver the offense, the greater the number of days. In the second stage, the court estimates the offender's financial condition and sets the "daily unit." This unit usually constitutes a fraction of the offender's daily income. For example, the daily unit in Finland is roughly 50% of the offender's net daily income.[177] A certain amount may be

175. Gary S. Becker, *Crime and Punishment: An Economic Approach*, 76 J. POL. ECON. 169, 179–80 (1968). Becker explained that the costs of punishment can be reduced by replacing imprisonment with pecuniary sanctions. Unlike fines, imprisonment imposes costs both on the offender and on society (primarily the costs of building and maintaining prisons and of hiring and training guards). Thus, according to economic theory, fines should be used to the extent possible before turning to other forms of penalty.

176. *See* Avner Bar-Ilan & Bruce Sacerdote, *The Response of Criminals and Noncriminals to Fines*, 47 J. L. & ECON. 1, 2, 16 (2004) (concluding that wealthier criminals are less responsive to fine increases).

177. CRIM. CODE ch.2 (a), § 2(2) (Fin.) provides that "[o]ne sixtieth of the average monthly income of the person fined, less the taxes and fees defined by a Decree and a fixed deduction for basic

deducted to allow for the basic needs of the offender and of his or her dependents.[178] The total fine imposed thus equals the number of days, which depends on the severity of the offense, multiplied by the daily unit, which depends on the offender's financial condition. Although the nominal amount paid may differ across offenders who have committed the same offense, the relative burden of the punishment is similar.

Assume, for example, that two similar offenders (in terms of criminal history), X and Y, committed a similar crime of larceny. However, the net daily income of offender X is $100 and the daily income of offender Y is $20. In the first stage the judge will determine the number of fine-days that fit the severity of the crime. Assume ten days comply with the gravity of the particular larceny. In the second stage, depending on each offender's daily income, the court determines the daily unit of the fine. This may be set at 50% of the net daily income, as explained above.[179] In this case, the daily unit of the fine is $50 for offender X and $10 for offender Y. Finally, the court needs to multiply the number of days by the daily unit. As a result, the fine imposed on offender X is 10×$50=$500. Offender Y will pay 10×$10=$100 for the same crime. This example demonstrates that the nominal fine is different for the two offenders. Nevertheless, the relative burden imposed on them by this sanction is similar, namely 50% of their ten-day income.

The idea of wealth-dependent criminal punishment was mentioned by Jeremy Bentham in the nineteenth century,[180] but was not endorsed by any legal system before the twentieth century. The first country to introduce day-fines was Finland in 1921.[181] Other Scandinavian countries followed

consumption, is deemed to be a reasonable amount of a day fine." *See also* Tapio Lappi-Seppälä, *Imprisonment and Penal Policy in Finland*, *in* 54 SCANDINAVIAN STUD. L. 333, 336 (Peter Wahlgren ed., 2009) (describing the Finnish day-fines system).

178. CRIM. CODE ch.2 (a), § 2(2) (Fin.).

179. For simplicity's sake, we do not take into account the possible deduction to allow for "basic needs."

180. 2 JEREMY BENTHAM, THEORY OF LEGISLATION 129, 143 (Charles M. Atkinson trans., 1914) (1802):

> Punishments the same in name are not always the same in reality. Age, sex, rank, fortune . . . ought to modify the penalties imposed for offences of the same character . . . [A] given pecuniary penalty might prove a bagatelle to the rich man, and act of cruel oppression when inflicted on his poorer neighbor . . . .
>
> Punishment ought . . . to correspond with [the offenders'] various measures of sensitivity . . . otherwise, punishments nominally the same, being found too severe for some persons and too mild for others, will now overshoot the mark . . . A fine, fixed by law at a given amount, can never be a punishment possessing this kind of equality, by reason of the great disparity in the fortunes of different offenders.

181. Hans-Jörg Albrecht, *Sanction Policies and Alternative Measures to Incarceration: European Experiences with Intermediate and Alternative Criminal Penalties*, UNITED NATIONS ASIA & FAR EAST INST. FOR PREVENTION CRIME AND TREATMENT OFFENDERS RESOURCE MATERIAL SERIES NO. 80 at 28, 33 (2009), http://www.unafei.or.jp/english/pdf/RS_No80/No80_07VE_Albrecht.pdf.

suit (e.g., Sweden in 1931 and Denmark in 1939).[182] However, only several decades later did other European countries adopt a day-fine model. These include Germany and Austria (1975), Hungary (1978), France and Portugal (1983),[183] Spain (1995), Poland (1997), and Switzerland (2007).[184] Yet, to this day, some European countries, such as Belgium, Iceland, Italy, the Netherlands, and Norway, still maintain the fixed-fine system.[185] The UK legislature adopted a day-fine system in 1991. While judicial resistance caused its abolition after a few months, a debate on reintroducing day-fines is ongoing.[186] This method was also endorsed in the early twentieth century by many Latin American systems, including Argentina, Colombia, the Dominican Republic, Ecuador, Guatemala, Honduras, Nicaragua, Paraguay, Uruguay, and Venezuela.[187]

Several jurisdictions in the United States experimented with day-fines,[188] but only a few states have retained them, and to a very limited extent compared to European jurisdictions.[189] The first jurisdiction to experiment with day-fines for low-level offenses was Richmond County in Staten Island, New York, in 1988.[190] The results of this experiment were positive, and similar experiments were held in Milwaukee (Wisconsin), Maricopa County (Arizona), Polk County (Iowa), Bridgeport (Connecticut), and several counties in Oregon.[191] At least three states—Alabama, Alaska, and Oklahoma—allow day-fines in limited circumstances.[192] Unfortunately, the American experience with day-fines has been so limited and sporadic that one cannot make general observations or conclusions about its reception and success, or compare it with the vast and long-standing European experience.

The particulars of a day-fine regime may vary. Jurisdictions differ, inter alia, in the types of offenses for which day-fines can be imposed, the scope of the income included in the calculation, the fraction used to calculate the

182. *Id.* at 34; EDWIN W. ZEDLEWSKI, U.S. DEP'T OF JUSTICE, DOCUMENT NO. NCJ 230401, ALTERNATIVES TO CUSTODIAL SUPERVISION: THE DAY FINE 3 (2010), https://www.ncjrs.gov/pdffiles1/nij/grants/230401.pdf.

183. Albrecht, *supra* note 181.

184. Elena Kantorowicz-Reznichenko, *Day-Fines: Should the Rich Pay More?*, 11 REV. L. & ECON. 481, 485 (2015).

185. Albrecht, *supra* note 181 at 34–35.

186. ANDREW ASHWORTH, SENTENCING AND CRIMINAL JUSTICE 348–50 (6th ed. 2015).

187. ZEDLEWSKI, *supra* note 182, at 3..

188. JUDITH GREENE & CHARLES WORZELLA, DAY FINES IN AMERICAN COURTS: THE STATEN ISLAND AND MILWAUKEE EXPERIMENTS (Douglas C. McDonald ed., 1992), https://csgjusticecenter.org/wp-content/uploads/2013/07/1992-day-fines-in-Staten-Island-Milwaukee.pdf.

189. *See* ZEDLEWSKI, *supra* note 182, at 5–6.

190. *Id.* at 5.

191. *Id* at 5–6.

192. *See, e.g.*, ALA. CODE § 12-25-32(2)(b)(8) (2017) (allowing day-fines in limited circumstances); ALASKA STAT. § 12.55.036 (repealed in 2009) (same); OKLA. STAT. tit. 22, § 991a(A)(y) (2017) (same).

daily unit, the upper and lower nominal limits on a daily unit, the minimum and maximum number of days on the fining spectrum, and the ratio between day-fines and prison times in case of default. For example, while there is no uniformity across countries with respect to imposing a nominal cap on the daily unit (Finland and Denmark do not have statutory caps), all systems that employ day-fines limit the number of days in the calculation.[193] This practice is consistent with other forms of (non-monetary) punishment wherein legislatures usually set upper limits on the duration of the suffering, such as imprisonment (aside from life imprisonment)[194] or community service.[195]

B. Tort Law Adaptation

As demonstrated in Part II, punitive damages serve two goals which are usually associated with criminal law: deterrence and retribution. As explained in Part III, the defendants' wealth may be relevant in achieving these goals, specifically when illicit gains are unobservable or when the perceived "punitive bite" of a punitive damages award depends on one's level of wealth. These two facts make day-fines, which aim to achieve the same goals with the same acknowledgement of the importance of wrongdoers' wealth, an appealing reference point in restructuring punitive damages. In this section, we develop a model of income-dependent punitive damages. Admittedly, in some jurisdictions, implementing this model may require legislative intervention, but such intervention is definitely not unprecedented in this area of law.[196] In Parts V and VI, we investigate whether this model serves the goals described in Parts II and III and can overcome constitutional, statutory, and procedural constraints.

To adapt the day-fine model to tort law, three steps must be taken. First, if the judge or jury concludes that punitive damages may be warranted, it must assess the appropriate level of the overall monetary sanction. To do so, it must determine the number of "severity units," reflecting the relative gravity of the specific wrong, calculate the "daily unit," which is the wrongdoer's daily income or a preset fraction

193. For a summary of the differences, see Kantorowicz-Reznichenko, *supra* note 184, at 486.

194. *See, e.g.*, FLA. STAT. § 775.082 (2017) ("A person who has been convicted of a designated misdemeanor may be sentenced as follows: (a) For a misdemeanor of the first degree, by a definite term of imprisonment not exceeding 1 year; (b) For a misdemeanor of the second degree, by a definite term of imprisonment not exceeding 60 days."); N.Y. PENAL LAW § 70.15 (2017) ("A sentence of imprisonment for a class A misdemeanor shall be a definite sentence. When such a sentence is imposed the term shall be fixed by the court, and shall not exceed one year . . . .").

195. *See, e.g.*, Art. 1:22c para. 2 SR (Neth.). (setting a 240-hour limit on community service); Community Service by Offenders (Hours of Work) (Scotland) Order 1996, SI 1938, ¶ 3 (setting a 300-hour limit).

196. *See supra* notes 37–39.

thereof,[197] and multiply the two variables. The product constitutes "total damages," the upper limit of the defendant's liability, to the extent that it is greater than compensatory damages (as explained below). If clear guidelines are set in advance, and relevant evidence about the wrongdoer's income is readily available, this calculation should be straightforward.

To provide certainty and avoid inconsistent (hence unfair) penalties, the law must determine in advance how total damages in cases requiring a punitive award are to be calculated. The importance of predetermined guidelines cannot be overstated. Any penalty must be reasonably predictable in its severity to enable every person to know what the stakes are in choosing one course of action or another.[198] Moreover, a punitive scheme must be consistent, imposing similar sanctions for similar wrongs.[199] Unpredictability of high punitive awards is in tension with their punitive function because of the implication of unfairness that an eccentrically high punitive verdict carries.[200] The current legal framework—with the single-digit limit on the punitive-to-compensatory damages ratio[201]—promises some level of certainty.[202] However, empirical studies demonstrate that this guidepost is insufficient. For example, one study found that the punitive-to-compensatory damages ratio was higher in jury trials than in bench trials,[203] undermining the twin goals of punitive damages. Moreover, the punitive-to-compensatory damages ratio can hardly be defended in terms of the theoretical justifications of punitive damages.[204] A superior model will align with these justifications and provide greater predictability.

Thus, lawmakers must first set up an index—a formula or guidelines for associating the severity of the conduct with the number of severity units that the wrongdoer must pay. To fulfill the proportionality requirements, the number of severity units should increase with the reprehensibility of the misconduct.[205] Such guidelines were introduced in the American day-fine pilots in the 1980s.[206] For instance, in the Staten

197. In criminal day-fine systems, the "daily unit" is often based on a certain portion of the offender's daily income. Lawmakers may similarly limit the unit value in the proposed punitive damages model.

198. Exxon Shipping Co. v. Baker, 554 U.S. 471, 502 (2008); *see also* Schwartz, *supra* note 5, at 141 ("From a deterrence standpoint, it confounds understanding to permit such vast uncertainty as to the level of the expected penalty.").

199. *Exxon Shipping*, 554 U.S. at 502.

200. *Id.* at 499–501.

201. State Farm Mut. Auto. Ins. Co. v. Campbell, 538 U.S. 408, 426 (2003).

202. *See* Eisenberg & Heise, *supra* note 5, at 327 (finding that the variation in punitive damages awards is largely a function of variation in the underlying compensatory award).

203. *Id.* at 328.

204. Sharkey, *supra* note 2, at 500.

205. *See infra* Section V.B.

206. GREENE & WORZELLA, *supra* note 188 at 13, 24.

Island experiment, the number of fine-days that could be imposed for offenses against the person ranged from 17–23 days in cases of minor injury to an acquaintance to 81–109 days in cases of a substantial injury to a stranger.[207] A similarly structured index can be used in tort actions. Note, however, that a wrongdoer is already liable for the victim's harm, so some correlation between the severity of the injury and that of the sanction exists. More importantly, the key factor in assessing punitive damages is the reprehensibility of the conduct, not the severity of the injury in itself.

Next, lawmakers must determine how a unit value, and more specifically the defendant's daily income, should be evaluated. Particularly, they must decide whether and how to take into account passive income and possibly the defendant's net use value of his or her own property, such as a household or a car. In theory, to avoid under-deterrence and under-punishment, the daily share should encompass all sources of wealth, not only income from work. A comprehensive income measure is used in the German and Swedish day-fines models.[208] Admittedly, evaluating a person's daily income might not be easy, so our proposal is viable only if courts can obtain information about defendants' income from reliable sources, such as the tax authorities. We shall return to this concern below.[209] Lastly, if the basic unit is not the wrongdoer's entire daily income, the exact percentage and any deductions must also be determined in advance.

The second step is the deduction of compensatory damages from total damages. Punitive damages are sums awarded to tort victims over and above their compensable harm. If total damages are greater than compensatory damages in a particular case, the punitive award should equal the difference between total damages and compensatory damages. If total damages are lower than compensatory damages, punitive damages should be nil. In such cases the monetary sanction that would serve the twin goals of punitive damages (deterrence and retribution) is lower than compensatory damages. However, as long as tort law is committed to rectification of harm, compensatory damages set the lower limit of the monetary award.

To illustrate the proposed model, assume that Z was assaulted in a bar fight, and incurred a $200 injury. Assume further that the number of severity units attached to recklessly inflicting minor personal injury under

207. *Id.* at 24.

208. *See* STRAFGESETZBUCH [StGB] [PENAL CODE], § 40, para. 3, *translation at* https://www.gesetze-im-internet.de/englisch_stgb/englisch_stgb.html (Ger.) ("The income of the offender, his assets and other relevant assessment factors may be estimated when setting the amount of a daily unit."); Lina Eriksson & Robert E. Goodin, *The Measuring Rod of Time: The Example of Swedish Day-fines*, 24 J. APPLIED PHIL. 125, 130 (2007) (discussing Swedish law).

209. *See infra* Section VI.C.

the preset guidelines is ten, and that the unit value is 50% of the wrongdoer's daily income. Now, if the injury was caused by X, whose daily income is $100, total damages are $100×10×0.5=$500. X will have to pay $200 in compensatory damages and $500-$200=$300 in punitive damages. If the injury was caused by Y, whose daily income is $20, total damages are $20×10×0.5=$100. Y will have to pay compensatory damages in the amount of $200 and no punitive damages.

While current doctrine enables courts to award extra-compensatory damages to secure proper deterrence and retribution, it does not allow courts to award sub-compensatory damages to secure proper deterrence and retribution. If law and policymakers decide that the compensatory goal may be subject to deterrence and retribution, courts may be allowed to award sub-compensatory damages when the compensatory level exceeds "total damages." In addition to all arguments made above, which focused on general and individual deterrence, awarding sub-compensatory damages may sometimes be justified in terms of marginal deterrence, as recognized by economic theorists in the related area of criminal fines.[210] Low-income potential wrongdoers might not be able to fully pay compensatory damages. These might be induced to commit more severe wrongs, because the expected benefit will increase, whereas the expected cost—capped by their solvency—will not. A reform allowing sub-compensatory damages has some merit, but it transcends the boundaries of this Article.

The third step is the consideration of other applicable sanctions. Because "total damages" set an upper limit to liability, courts awarding punitive damages must determine whether this limit has already been reached through criminal fines and other civil sanctions imposed on the defendant for the same conduct.[211] If the aggregate of all monetary sanctions exceeds "total damages," punitive damages should not be awarded at all. For example, assume total damages equal $10,000 and compensatory damages are $4,000. If a criminal fine in the amount of $6,000 was imposed on the wrongdoer for the same wrongdoing, the court should not award punitive damages.

210. C.Y. Cyrus Chu & Neville Jiang, *Are Fines More Efficient than Imprisonment?*, 51 J. PUB. ECON. 391 *passim* (1993).

211. Pac. Mut. Life Ins. Co. v. Haslip, 499 U.S. 1, 21–22 (1991).

V. THEORETICAL DEFENSE OF THE MODEL

A. Deterrence

As Part III demonstrated, the defendant's wealth may be relevant to deterrence if punitive damages are imposed to secure internalization, and potential injurers are risk-averse and do not have access to liability insurance, or if they are imposed to disgorge an illicit gain, which is unobservable and correlated with the defendant's wealth. Under current law, the extent of punitive damages awards depends on the three guideposts set forth in *BMW v. Gore*.[212] In reviewing punitive damages awards, courts ought to consider the reprehensibility of the defendant's misconduct; the disparity between the plaintiff's actual or potential harm and the punitive damages award; and the difference between the punitive damages award and the civil or criminal penalties authorized or imposed in comparable cases. Apart from examining the level of reprehensibility, these guideposts focus on factors which are external to the defendant. Therefore, similar wrongs are expected to result in equal punitive damages awards in absolute terms.

If socially illicit gain is correlated with wealth[213] and the marginal utility of individual wealth is diminishing,[214] a wealthier wrongdoer's willingness to pay for committing the wrong is higher than that of a low-income injurer. In such circumstances, a fixed punitive damages award, just like fixed fines in criminal law, might under-deter the wealthier.[215] Under-deterrence will occur when the defendant's willingness to pay for the pleasure of causing harm exceeds the sum of compensatory and punitive damages. The problem of under-deterrence of wealthier offenders was one of the justifications for introducing day-fines in the criminal justice system,[216] and it is equally relevant in the case of punitive damages.

212. *See supra* notes 42–52 and accompanying text.

213. Hylton, *supra* note 5, at 936–38.

214. *See* ROBERT S. PINDYCK & DANIEL L. RUBINFELD, MICROECONOMICS 167–68 (8th ed. 2012) (discussing diminishing marginal utility of wealth).

215. This argument is theoretical and based on the utility function. To the best of our knowledge, there are no empirical studies systematically comparing the deterrent effect of fixed fines and day-fines. Most studies on day-fines focus on payment rates, which are very important in the criminal law context, because offenders who fail to pay their fines are often imprisoned. One of the empirically established advantages of day-fines is higher payment rates. *See, e.g.*, LAURA L. WINTERFIELD & SALLY T. HILLSMAN, THE STATEN ISLAND DAY-FINE PROJECT, NAT'L INST. OF JUSTICE: RESEARCH IN BRIEF 1, 3 (1993). Forasmuch as imprisonment is not available in tort law, reducing default rates is irrelevant to our analysis. Older studies demonstrated that day-fines were better in reducing reconviction rates than probation and imprisonment with respect to property crimes of non-career criminals. *See, e.g.*, Hans-Jörg Albrecht & Elmer H. Johnson, *Fines and Justice Administration: The Experience of the Federal Republic of Germany*, 4 INT'L J. COMP. & APPLIED CRIM. JUST. 3, 12–13 (1980).

216. *See, e.g.*, Eriksson & Goodin, *supra* note 208, at 129 (discussing the Swedish system).

Assuming a certain activity, such as defamation, is undesirable behavior that tort law seeks to eradicate, we would like to set "price discrimination" where punitive damages equal the minimum extent of damages that can deter such behavior. In terms of marginal utility, to deter a wrongdoer from causing harm, the utility loss from "total damages" needs to be equal to or higher than the wrongdoer's utility gain from his misconduct. And if the gain is correlated with wealth,[217] "total damages" ought to increase with the level of wealth to achieve the same utility loss.[218] Figure 2 illustrates this for two wrongdoers with a similar utility function and different levels of wealth.

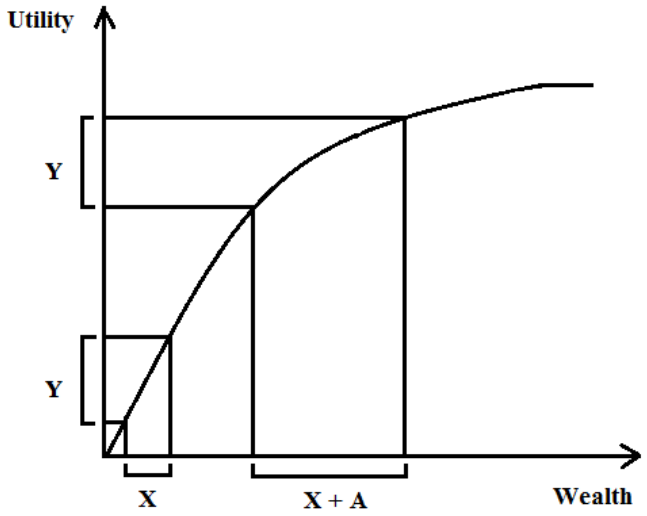

Figure 2. Equalizing Utility Loss

The Figure demonstrates that in order to impose an equal utility loss Y (which fits the particular wrong), "total damages" should equal X for the poorer wrongdoer, and X+A for the wealthier. Sufficient deterrence might entail variance in the unit value. In other words, if the purpose of the reform is proper price discrimination, the portion of wealth taken into account in calculating "total damages" should vary in accordance with the defendant's personal wealth and utility function. However, courts cannot ascertain individual utility functions, and are therefore unable to calculate

217. Polinsky & Shavell, *supra* note 5, at 913.

218. Our analysis is limited to individuals. As explained above, corporations are generally risk-neutral and are not motivated by socially illicit gains. Therefore, punitive damages imposed on corporations should be independent of their wealth. *See supra* notes 153–156 and accompanying text.

the impact of monetary sanctions on different individuals in terms of utility. Using equal portions of wealth (or income) when determining "total damages," as in the criminal day-fines model, is a reasonable second-best choice.

An additional advantage of an income-dependent punitive damages system is the enhancement of marginal deterrence. If punitive damages are fixed for similar misconducts resulting in comparable harms, marginal deterrence has a limit—reached when the monetary sanction equals the defendant's wealth. Once "total damages" equal or exceed the defendant's entire wealth, he or she no longer has an incentive to choose a less severe misconduct. In the economics literature of criminal law, this limitation justifies the imposition of imprisonment.[219] Nevertheless, imprisonment is not available in tort law, rendering alternative guarantees for marginal deterrence crucial. The proposed model translates the gravity of the misconduct into the number of severity units (which is independent of wealth). Therefore, the more egregious the conduct, the greater the number of severity units. Since the absolute extent of punitive damages is determined by a proportion of the defendant's income, marginal deterrence is not limited by the defendant's wealth. This advantage does not hold if "total damages" are lower than compensatory damages. In such circumstances, punitive damages will be nil and remain nil at least with some increases in the level of reprehensibility.

Finally, taking into account the wrongdoer's wealth is necessary to avoid over-deterrence in performing activities which are not undesirable per se. Sometimes, potential wrongdoers cannot obtain insurance coverage for punitive damages, either because the market does not provide such coverage,[220] possibly due to moral-hazard problems, or because the law prohibits such coverage on public policy grounds.[221] As explained above, poorer wrongdoers may be more risk-averse than wealthier ones.[222] Thus, in the absence of insurance, imposing an identical monetary sanction will have a different impact in terms of utility on people with different levels of wealth. The expected utility loss generated by a particular monetary sanction will be greater for poorer potential wrongdoers, resulting in greater willingness to invest in excessive precautions or to engage in

219. Chu & Jiang, *supra* note 210 *passim*; Steven Shavell, *Criminal Law and the Optimal Use of Nonmonetary Sanctions as a Deterrent*, 85 COLUM. L. REV. 1232 (1985).

220. *See, e.g.*, Benjamin P. Cooper, *The Lawyer's Duty to Inform His Client of His Own Malpractice*, 61 BAYLOR L. REV. 174, 212–13 (2009) ("The possibility of punitive damages is particularly significant since many malpractice insurers do not cover punitive damages.").

221. *See* Michael A. Rosenhouse, Annotation, *Liability Insurance Coverage as Extending to Liability for Punitive or Exemplary Damages*, 16 A.L.R.4th 11 (1982) (discussing relevant case law). For a more recent analysis of case law, see Fairfield Ins. Co. v. Stephens Martin Paving, L.P., 246 S.W.3d 653, 688 (Tex. 2008) (Hecht, J., concurring).

222. *Supra* Section III.B; *see also* Polinsky & Shavell, *supra* note 5, at 913.

suboptimal levels of activity. While the proposed model falls short of eliminating this problem, it may help ameliorate it.

B. Retribution

Retribution seems to pose a more serious challenge to our model. To recap, it is a form of justice that stresses the proportionality between the gravity of the wrong and the severity of the sanction.[223] For income-dependent punitive damages to meet the standard of retribution, they must fulfill the requirements of cardinal and ordinal proportionality. At first sight, some difficulties are evident. With respect to cardinal proportionality, there might be rare cases in which an exceptionally wealthy wrongdoer commits a minor tort. Based on the suggested model, "total damages" might be very high in absolute terms, and seem incompatible with the gravity of the conduct. The European experience with day-fines in criminal law provides a prominent example. In 2001, a director of Nokia was given a $103,600 fine for speeding in Finland.[224] Similarly, our model does not seem to comply with the principle of ordinal proportionality if we look at "total damages" in absolute terms. For instance, the proposed model might oblige a wealthy wrongdoer to pay $10,000 in damages for publicly insulting another person, while an indigent wrongdoer would only pay $1,000 for physically assaulting another. This seems to go against ordinal proportionality, which requires a graver misconduct to be punished by a more severe sanction (higher damages).

There are two possible rejoinders to these concerns. First, the proposed model is unique: it requires careful interpretation of the theory rather than straightforward application. Forasmuch as this model consists of two independent components—the number of severity units and the unit value, compliance with the proportionality test should not be evaluated by the nominal value of "total damages" but in view of the first component (severity). In other words, in an income-dependent punitive damages system, the gravity of the misconduct and the severity of the sanction are captured by the number of severity units. With respect to wealth-dependent sanctions, such as day-fines in criminal law and income-dependent punitive damages in torts, the number of days or severity units, rather than the nominal awards, may serve the expressive function of punishment. For instance, Swedish media reports the number of fine-days

223. *See supra* Section II.C.

224. *Nokia Boss Gets Record Speeding Fine*, BBC NEWS (Jan. 14, 2002, 2:29 PM), https://perma.cc/RQ7F-B4GF.

imposed on convicted offenders rather than focusing on the total amount of the fine.[225]

Lawmakers can establish guidelines attaching specific ranges of severity units to different types of misconduct, depending on their gravity. Alternatively, courts may use the factors, recognized as relevant in determining the degree of reprehensibility, as a tool for determining the number of severity units on a case-by-case basis. These factors include the type of harm, the victim's vulnerability, the injurer's intentional malice, the repetitive nature of the conduct, etc.[226] If this suggested interpretation is followed, the requirements of cardinal and ordinal proportionality are met. With respect to cardinal proportionality, it is possible to set upper and lower limits to the number of severity units attached to each type of wrongdoing. For instance, in the Staten Island experiment with day-fines, the range of fine-days was 5–120.[227] This range was deemed appropriate considering the gravity of the misdemeanors included in the experiment. Moreover, a specific range was set for each offense, allowing flexibility and judicial discretion in light of mitigating and aggravating circumstances in specific cases.[228] To the extent that the number of severity units is within a range that "fits" the absolute gravity of the wrong, cardinal proportionality is attained. Similarly, an index that makes the number (or range) of severity units contingent on the reprehensibility of the conduct ensures that the punitive award reflects the relative gravity of the conduct, as required by ordinal proportionality.[229]

An additional rejoinder is related to our general discussion of the need to take the defendant's wealth into account in assessing punitive damages.[230] Many retributivists believe that the wrongdoer's wealth should not be disregarded when pecuniary penalties are imposed.[231] Retributive justice requires that the level of the defendant's suffering be proportional to the wrong he has committed. Under the current system and the three guideposts,[232] two wrongdoers liable for an equally reprehensible misconduct might face the same punitive damages award in nominal

225. Eriksson & Goodin, *supra* note 208, at 131. Thus, when a person was convicted of two minor drug offenses and DUI, the media reported his sentence as 70 days of a 50 kronor fine. Matilda Ermeland, *Dagsböter för Knarkbrott*, SUNDSVALLS TIDNING (Nov. 17, 2015), https://perma.cc/MA6W-ZRPC.

226. *See* Baker v. Exxon Mobile Corp., 490 F.3d 1066, 1085–89 (9th Cir. 2007) (discussing the relevant reprehensibility factors), *rev'd on other grounds*, 554 U.S. 471 (2016).

227. Greene, *supra* note 206, at 22.

228. *Id.* at 23–24.

229. A system in which mitigating and aggravating factors are considered on a case-by-case basis may also be used in a tort adaptation. The downside is additional complication and higher administrative costs.

230. *Supra* Section III.B.

231. *See supra* notes 163–164 and accompanying text.

232. *See supra* notes 42–52 and accompanying text.

terms. However, every given amount constitutes a different portion of each wrongdoer's income or wealth. A $1,000 award constitutes 10% of a $10,000 monthly income, and 80% of a $1,250 monthly income. Thus, it is reasonable to assume that the suffering of the person with the lower income is greater than that of the person with the higher income. If the level of reprehensibility is equal, then from a retributive justice perspective, a nominally equal punitive award cannot be deemed proportionate to the wrong for both wrongdoers. The proposed model, albeit not without flaws, offers a potential solution. Given the diminishing marginal utility of wealth (see Figure 2), obliging the wealthier wrongdoer to pay the same portion of his or her income as the poorer wrongdoer closes the gap between the levels of suffering which would exist if the defendant's wealth were ignored. Aligning the monetary sanction with the wrongdoer's wealth is more likely to achieve proportionality between the gravity of the wrong and the severity of the sanction as actually experienced by wrongdoers.

The question is which portion of a wrongdoer's wealth should be set as the basis for the daily unit of income-dependent punitive damages in order to achieve proportionality between the gravity of the conduct and the severity of the sanction. Intuitively, a loss of 50% of one's wealth will impose a different burden on people with different initial wealth, even if the utility function is identical for all. To illustrate this point, let us return to the case of Moe, Larry, and Curly.[233] As before, assume the three have an identical utility function, $y=10{,}000log(x/10{,}000+1)$. Moe has $10,000,000, Larry $1,000,000, and Curly $100,000. If total damages constitute 50% of each wrongdoer's wealth, Moe will be obliged to pay $5,000,000, Larry $500,000, and Curly $50,000 (assuming the actual harm does not exceed $50,000). Inserting these amounts into the utility function will result in a loss of 3,006 utility units for Moe, 2,968 utility units for Larry, and 2,632 utility units for Curly.[234] This example demonstrates two important features of the proposed model. First, it closes the gaps between utility losses incurred by perpetrators of similar wrongs who have different levels of wealth, created by a system based on proportionality between the extent of punitive damages and the gravity of the wrong. Second, this model does not close these gaps completely. Put differently, even income-dependent punitive damages do not guarantee that equal wrongdoing will result in equally burdensome sanctions. And variance in individual utility functions surely exacerbates the problem.

233. *See supra* Section III.B.

234. Utility losses are calculated as follows (x being the wrongdoer's initial wealth): $10{,}000log(x/10{,}000+1)-10{,}000log(0.5x/10{,}000+1)$.

Ideally, "total damages" should be adjusted in accordance with each wrongdoer's personal wealth and utility function. However, this is clearly impractical: courts cannot ascertain individual utility functions and are unable to calculate the impact of monetary sanctions on different individuals in terms of utility. The proposed model uses a flat fraction of the wrongdoer's daily income as the unit value. For example, "total damages" for two wrongdoers responsible for an equally reprehensible wrong may be calculated by multiplying 50% of each wrongdoer's daily income (the unit value) by the number of severity units. This is clearly a second-best solution from a retributive perspective. To be sure, it does not provide perfect individualization of total damages. However, it brings the assessment of punitive damages closer to that required by the principles of retributive justice. And it has the important advantages of practicability and predictability.

VI. CHALLENGES AND CONCERNS

Even if the proposed model achieves the twin goals of punitive damages, deterrence and retribution, it needs to be evaluated in light of existing constitutional and statutory constraints. In this Part we first analyze potential constitutional challenges. Next, we investigate the legislative reforms that adopting the proposed model may entail. Finally, we review the rules on information discovery in civil cases, in order to assess the practicability of an income-based calculation of damages.

A. Constitutional Constraints

The constitutional challenge concerns the prohibition on excessive sanctions. Although the Excessive Fines Clause of the Eighth Amendment[235] does not apply to punitive damages awards in civil cases between private parties,[236] punitive damages are subject to the Due Process Clause of the Fourteenth Amendment.[237] The latter forbids states from depriving persons of "life, liberty, or property, without due process of law." According to the Supreme Court, a punitive damages award violates the Due Process Clause if it is grossly excessive ("substantive due

235. U.S. CONST. amend. VIII.

236. Browning-Ferris Indus. of Vt., Inc. v. Kelco Disposal, Inc., 492 U.S. 257, 263–64, 275–76 (1989).

237. U.S. CONST. amend. XIV, § 1; State Farm Mut. Auto. Ins. Co. v. Campbell, 538 U.S. 408, 416 (2003); Cooper Indus., Inc. v. Leatherman Tool Group, Inc., 532 U.S. 424, 433–34 (2001); BMW of N. Am., Inc. v. Gore, 517 U.S. 559, 562–63 (1996); TXO Prod. Corp. v. All. Res. Corp., 509 U.S. 443, 453–55 (1993).

process"), or imposed without a fair procedure ("procedural due process").[238]

Taking the defendant's wealth into account in assessing punitive damages does not in itself violate the Due Process Clause. In fact, starting in the 1980s, courts have used the defendant's wealth to refute fears that punitive awards are excessive.[239] In *Pacific Mutual Life Insurance Co. v. Haslip*,[240] the Supreme Court mentioned several factors taken into account in determining whether a punitive damages award was excessive or inadequate, including the three guideposts subsequently endorsed in *BMW*, and the "financial position" of the defendant. The Court held that these factors imposed a "sufficiently definite and meaningful constraint on the discretion of . . . fact finders in awarding punitive damages."[241] In *TXO Production Corp. v. Alliance Resources Corp.*,[242] the Court found that although the punitive award was very large, it was not grossly excessive, in part due to the defendant's wealth. In *BMW*, the Court explained that the defendant's wealth was a relevant consideration, because a fixed amount of punitive damages would punish wealthy and poor wrongdoers differently.[243] In line with these rulings, some state legislatures have explicitly instructed or permitted courts to consider the defendant's financial condition and net worth when assessing excessiveness of punitive damages.[244]

True, in *State Farm v. Campbell*[245] the Supreme Court opined that "[t]he wealth of a defendant cannot justify an otherwise unconstitutional punitive damages award." Because it is clear that no court can justify or affirm an unconstitutional award, the Court's statement could be understood as a direction to lower courts not to use wealth as the primary basis for determining a punitive award. The defendant's wealth is not totally irrelevant, but relying on it is subject to the general constitutional constraints. We turn, therefore, to the challenges that the three *BMW v. Gore*[246] guideposts might present to the proposed model.

238. *See* Philip Morris USA v. Williams, 549 U.S. 346, 353–55 (2007) (explaining that punitive damages are subject to both substantive and procedural due process).

239. *See, e.g.*, Hatrock v. Edward D. Jones & Co., 750 F.2d 767, 773 (9th Cir. 1984) (holding that the punitive award was not excessive because it represented only 0.33% of the defendant's gross revenues).

240. 499 U.S. 1, 21–22 (1991).

241. *Id.* at 22.

242. 509 U.S. 443, 462 (1993).

243. BMW of N. Am., Inc. v. Gore, 517 U.S. 559, 591 (1996) (Breyer, J., concurring).

244. *See, e.g.*, FLA. STAT. § 768.72(1) (2017); IOWA CODE § 668A.1 (2017); KAN. STAT. ANN. § 60-3701(b)(6) (2017); MINN. STAT. § 549.20(3) (2017); MISS. CODE ANN. § 11-1-65(1)(f)(ii)(3) (2017); MO. REV. STAT. § 510.263(2) (2017); NEV. REV. STAT. § 42.005(4) (2017).

245. 538 U.S. 408, 427 (2003).

246. 517 U.S. at 575–85.

First, punitive damages should be aligned with the degree of reprehensibility of the defendant's misconduct. For instance, the Court in *BMW v. Gore* criticized Alabama's statute for not making a distinction between serious misrepresentation ("tricking the elderly out of their life savings") and minor representation (failing to disclose minor repairs of a new car before sale).[247] As mentioned above, reprehensibility of the conduct can be expressed in "severity units" (or a number of days). A suitable range of "severity units" for each level of reprehensibility or for each category of misconduct can be provided by preset guidelines. Thus, the difference between serious and minor misrepresentations can be reflected in different ranges of severity units. Moreover, fine-tuning the monetary sanction on the basis of wrongdoers' wealth generates more accurate proportionality between the gravity of the wrong and the severity of the sanction as actually experienced by wrongdoers. To the extent that the impact of sanctions on wrongdoers is important in determining their severity, the proposed model seems more consistent with the constitutional imperative.

Second, *BMW* requires consideration of the ratio between punitive and compensatory damages. This guidepost seems to reduce the proposed model's feasibility. Although the Supreme Court has refused to draw a clear line between constitutionally acceptable and unacceptable ratios,[248] it has provided instructive observations. For example, in *Haslip*,[249] and then again in *BMW*,[250] the Court concluded that an award of more than four times the amount of compensatory damages might be close to constitutional impropriety.[251] In *State Farm* it opined that single-digit ratios were more likely to conform to due process than higher ratios.[252] Because compensatory damages depend on the plaintiff's harm, and punitive damages under the proposed model depend on the defendant's wealth, the ratio between punitive and compensatory damages will occasionally exceed the acceptable ratio. For instance, if an especially wealthy wrongdoer has caused minor harm, applying the proposed model

247. *Id.* at 589–90.

248. *See, e.g.*, *State Farm*, 538 U.S. at 425 ("We decline again to impose a bright-line ratio which a punitive damages award cannot exceed."); *BMW*, 517 U.S. at 582 ("Of course, we have consistently rejected the notion that the constitutional line is marked by a simple mathematical formula, even one that compares actual and potential damages to the punitive award . . . . It is appropriate, therefore, to reiterate our rejection of a categorical approach."); TXO Prod. Corp. v. All. Res. Corp., 509 U.S. 443, 458 (1993) ("[W]e are not prepared to enshrine petitioner's comparative approach in a 'test' for assessing the constitutionality of punitive damages awards."); Pac. Mut. Life Ins. Co. v. Haslip, 499 U.S. 1, 18 (1991) ("We need not . . . draw a mathematical bright line between the constitutionally acceptable and the constitutionally unacceptable that would fit every case.").

249. *Haslip*, 499 U.S. at 23–24.

250. *BMW*, 517 U.S. at 581.

251. *See also State Farm*, 538 U.S. at 426 (endorsing *Haslip* and *BMW*).

252. *Id.*

might yield a very high ratio. Allowing punitive awards above the *State Farm* single-digit guideline based on the defendant's wealth requires modification, or at least reinterpretation, of *State Farm* itself. This would not be revolutionary, because the Court has never adopted a rigid range of acceptable ratios, and has already recognized circumstances justifying higher ratios.[253]

Third, courts should evaluate punitive damages awards in light of other criminal and civil penalties imposed or authorized in comparable cases.[254] As long as the American criminal justice system applies fixed fines rather than day-fines, punitive damages under the proposed model might be incompatible with criminal sanctions imposed for similar misconduct. In an income-dependent punitive damages scheme, reprehensibility is expressed by "severity units" (or the number of days), not by the total amount of damages. In fixed-fines criminal systems, reprehensibility is expressed by the payable amount. Punitive damages under our model will not correlate with relevant criminal fines. Thus, implementing this model seems to require some adjustment of the third *BMW* guidepost. However, the Court in *State Farm* downplayed this guidepost. It explained that criminal sanctions may be used to determine "the seriousness with which a State views the wrongful action," but have "less utility" in determining the dollar amount of punitive damages.[255] Put differently, criminal sanctions may be considered because they indicate the relative gravity of the wrong, not because punitive damages must be similar in amount to relevant criminal sanctions. The proposed model ensures correlation between the punitive award and the gravity of the wrong, as reflected by relevant criminal sanctions, so the third guidepost is satisfied. But even if the third guidepost were interpreted to require correlation between the amount of punitive damages and relevant criminal sanctions, this would not be a significant obstacle. To begin with, damages can be compared not only to criminal fines, but also to prison terms. Correlation between the severity of income-dependent punitive damages and that of incarceration for a comparable wrong (where relevant) can be more easily attained. If the lack of correlation between punitive damages and criminal fines is still deemed inconsistent with the third guidepost, the latter should be qualified. Presumably, criminal sanctions have served as a benchmark due to the lack of clear guidelines for assessing punitive damages. Once the proposed

253. *BMW*, 517 U.S. at 582 ("Indeed, low awards of compensatory damages may properly support a higher ratio than high compensatory awards, if, for example, a particularly egregious act has resulted in only a small amount of economic damages. A higher ratio may also be justified in cases in which the injury is hard to detect or the monetary value of noneconomic harm might have been difficult to determine."); *see also State Farm*, 538 U.S. at 425 (quoting *BMW*).

254. *BMW*, 517 U.S. at 583–85.

255. *State Farm*, 538 U.S. at 428.

model provides clear guidelines, reliance on external measures becomes redundant anyway.

Last but not least, the Supreme Court has repeatedly mentioned in its due process analysis the need to avoid "arbitrary deprivation of property,"[256] and to ensure that the probability and extent of punitive damages are predictable.[257] Punitive damages raise concerns in this regard. For example, the Court in *State Farm* pointed out that jury instructions leave juries wide discretion as to the amount of punitive damages.[258] When such vague instructions are combined with evidence about the defendant's net worth, verdicts might reflect jurors' biases against big businesses, especially those without a strong local presence.[259] The proposed model, if properly implemented, provides concrete and clear guidelines, limits juries' discretion and susceptibility to bias, avoids arbitrariness, and introduces certainty and predictability into the system. Reprehensibility is expressed in severity units (or number of days), and an appropriate range of units can be set in advance for each category of misconduct. The defendant's wealth is also incorporated into the calculation in a systematic and predictable manner. For example, the legislature can set the unit value at 50% of the defendant's daily income. The special structure of income-dependent damages thus has the potential to satisfy the requirements of due process in a better manner than the existing system.

B. Statutory Caps

Another obstacle to implementing the proposed model may be statutory caps on punitive damages. Many states limit the extent of punitive damages by setting an absolute cap or a maximal punitive-to-compensatory damages ratio.[260] Although some state courts have found these limits unconstitutional,[261] they are still very common. Unfortunately, they may frustrate any attempt to secure deterrence and retributive justice through punitive damages. Specifically, if the wrong is egregious or if the wrongdoer is especially wealthy, "total damages" might exceed the

256. *Id.* at 417 ("[P]unitive damages pose an acute danger of arbitrary deprivation of property. Jury instructions typically leave the jury with wide discretion in choosing amounts . . . ."); *BMW*, 517 U.S. at 586 (Breyer, J., concurring) ("A 'grossly excessive' punitive award amounts to an 'arbitrary deprivation of property without due process of law.'") (quoting TXO Prod. Corp. v. All. Res. Grp., 509 U.S. 443, 454 (1993)).

257. *State Farm*, 538 U.S. at 418 ("A State can have no legitimate interest in deliberately making the law so arbitrary that citizens will be unable to avoid punishment based solely upon bias or whim.").

258. *Id.* at 417.

259. *See id.* at 418.

260. For a representative list of statutory caps, see *supra* note 37.

261. *See, e.g.*, Lewellen v. Franklin, 441 S.W.3d 136, 144–45 (Mo. 2014) (finding the Missouri statutory cap unconstitutional).

statutory cap. In such cases, the punitive award will be insufficient: it will under-deter potential wrongdoers and be too lenient considering the gravity of the wrong.

The problem might not be as severe as it appears at first glance. In many states, the caps are differential: the more reprehensible the conduct, the higher the cap. For example, the Florida punitive damages legislation normally caps the awards at three times the amount of compensatory damages or $500,000 (whichever is greater).[262] But a higher cap applies if the wrongful conduct was unreasonably dangerous and motivated by unreasonable gain, and the defendant was aware of the risk.[263] The cap is lifted altogether in cases of intentional causation of harm.[264] This flexibility may suffice to overcome the theoretical hurdle in most cases. Admittedly, however, it is far from perfect.

Adapting the legislative framework to the structure and purposes of the proposed model is probably more desirable. There are several ways to limit punitive damages in an income-dependent system. Most countries that implement day-fines in criminal law impose statutory limits on the daily unit,[265] and this can also be done in an income-dependent punitive damages scheme. However, such limits give up on efficient deterrence and proper retribution in the very same cases that inspired the invention of wealth-dependent sanctions—those of exceptionally wealthy defendants. Relatively high caps on daily units, such as the €30,000 limit on the daily unit in the German day-fine system,[266] can surely ameliorate the problem, but they are still hard to justify. Alternatively, the legislature can limit the number of severity units. All systems that employ day-fines in criminal law limit the number of days in the calculation.[267] As long as these limits maintain ordinal proportionality between the gravity of the wrong and the severity of the sanction, they seem less problematic than limiting the daily unit. They also contribute to predictability. Finally, the legislature can remove the caps altogether, allowing unqualified application of the proposed model.

262. FLA STAT. § 768.73(1)(a) (2017)

263. *Id.* § 768.73(1)(b).

264. *Id.* § 768.73(1)(c).

265. Kantorowicz-Reznichenko, *supra* note 193, at 486. Only Finland and Denmark do not impose any statutory cap on the daily unit. *Id.*

266. STRAFGESETZBUCH [StGB] [PENAL CODE], § 40, para. 2, *translation at* https://www.gesetze-im-internet.de/englisch_stgb/englisch_stgb.html (Ger.).

267. Kantorowicz-Reznichenko, *supra* note 193, at 486.

C. Procedural Tools

The viability of the proposed model hinges on reasonable access to the defendant's financial information. To ascertain the second variable in the formula, the court needs to assess the defendant's net worth. To calculate "total damages," the court needs to multiply a fraction of the defendant's daily income, broadly defined (the unit value), by the number of severity units. Wrong assessment of the financial condition of the defendant would lead to under- or overdeterrence. Furthermore, punitive damages will no longer represent a proportionate punishment for the specific wrong and wrongdoer. Therefore, it is crucial to review whether the current legal framework provides procedural instruments to access information about the defendant's financial condition.

Rule 26 of the Federal Rules of Civil Procedure governs discovery in civil cases tried in U.S. district courts.[268] According to the general principle, parties can obtain discovery regarding non-privileged information that is "relevant to any party's claim or defense and proportional to the needs of the case," considering the burden of discovery and its likely benefit.[269] Specifically, the rule instructs the claimant to make available documents relevant to the computation of damages.[270] However, while the scope of information eligible for discovery is broad, courts have some discretion, and may balance the interests of justice against the parties' privacy. Thus, even though income tax returns are nonprivileged, courts may limit their discovery.

For instance, in *Wiesenberger v. W. E. Hutton & Co.*[271] the plaintiff brought an action against financial advisers for considerable loss caused by their flawed advice.[272] The defendants requested disclosure of the plaintiff's income tax returns, to demonstrate that the poor investment yielded tax savings that could partly offset the loss. The court denied the request, holding that unless "clearly required in the interests of justice," litigants should not be obliged to disclose income tax returns.[273] It observed that in this particular case, the relevant information could be obtained from other sources, such as the plaintiff's books and bank statements, making the income tax returns unnecessary.[274] The court added

268. FED. R. CIV. P. 26.
269. *Id.* § 26(b)(1).
270. *Id.* § 26(a)(1)(A)(iii).
271. 35 F.R.D. 556 (S.D.N.Y. 1964).
272. *Id.* at 557.
273. *Id.*
274. *Id.* at 558.

that if the alternative sources were proven insufficient, the defendant would be able to reapply for discovery of the income tax returns.[275] Note further that while Rule 26 explicitly mentions discovery of documents underlying the computation of damages by the claimant, and not by the defendant, the general principle of discovery applies to both.

In the proposed model, the defendant's financial information in general, and income tax returns in particular, are directly relevant to the assessment of "total damages." Without this information, the second variable in the formula, the unit value, cannot be calculated. Furthermore, producing these documents would not normally be costly or burdensome, because they already exist and are in the defendant's possession. Privacy concerns may surely arise, but the importance of the defendant's financial information for the computation of damages, and the ability to minimize any infringement of privacy, will probably tip the scale for disclosure. As long as the defendant's wealth can be properly ascertained, the practicability obstacle is surmounted.

State courts have their own rules on discovery in civil cases, and these ought to be examined separately. As explained above, many states have recognized the relevance of the defendant's wealth in determining the extent of punitive damages.[276] This already implies that courts have reasonable access to defendants' financial information. At any rate, explicit procedural rules permitting discovery of financial information exist in many jurisdictions. For instance, the California Civil Code stipulates that the defendant's financial condition is relevant in the assessment of punitive damages, and that the court can order the defendant to disclose his financial information.[277] It is very common to require discovery of the defendant's financial information, income tax returns in particular, only when it becomes necessary, namely after establishing that punitive damages are prima facie allowed.[278] The rationales for this

275. *Id.*

276. *See supra* note 137 and accompanying text.

277. CAL. CIV. CODE § 3295 (2017).

278. *See, e.g.,* ALASKA STAT. § 09.17.020(e) (2017) (allowing discovery only after "the fact finder has determined that an award of punitive damages is allowed . . . ."); FLA. STAT. § 768.72(1) (2017) (allowing discovery "after the pleading concerning punitive damages is permitted."); IOWA CODE § 668A.1(3) (2017) ("[A] claim for punitive damages shall not form the basis for discovery of the wealth . . . on behalf of the party from whom punitive damages are claimed until such time as the claimant has established that sufficient admissible evidence exists to support a prima facie case establishing the requirements [for punitive damages]."); Bryan v. Thos. Best & Sons, Inc., 453 A.2d 107, 108 (Del. Super. Ct. 1982) ("[I]f . . . a factual basis for punitive damages existed, then plaintiff would be permitted to discover defendants' financial condition."); Hetter v. Eighth Judicial Dist. Court of State in & for the Cty. of Clark, 874 P.2d 762, 766 (Nev. 1994) ("[B]efore tax returns or financial records are discoverable on the issue of punitive damages, the plaintiff must demonstrate some factual basis for its punitive damage claim."); Rupert v. Sellers, 368 N.Y.S.2d 904, 912 (N.Y. App. Div. 1975) ("[T]he Court should take a special verdict as to whether . . . plaintiff is entitled to punitive damages.

restriction may be protecting defendants from harassment and undue pressure from plaintiffs and minimizing infringement of privacy.[279]

The proposed model does not seem to pose special practical challenges for the courts. Courts in many states are experienced in processing requests for disclosure of financial information and in assessing defendants' wealth in the context of punitive damages. However, the structure of the proposed model raises a unique procedural question: the computation of "total damages," which requires information about the defendant's financial condition, is necessary not only for determining the extent of punitive damages, but also for deciding whether they are deserved (if total damages exceed compensatory damages). This problem can be resolved. Punitive damages are generally reserved for exceptional cases, where the defendant's culpability, after having paid compensatory damages, is so reprehensible as to warrant the imposition of further sanctions to achieve retribution or deterrence.[280] Courts can make a prima facie determination about the desirability of punitive damages in light of the reprehensibility of the conduct (which controls the number of severity units) and cumulative experience in implementing the proposed model.[281] A positive conclusion will be followed by disclosure of the defendant's financial information, enabling calculation of the unit value and final assessment of the desirability and extent of punitive damages.

CONCLUSION

This Article has set forth an innovative model for incorporating defendants' wealth into the assessment of punitive damages, based on the notion of day-fines, commonly used in criminal law systems around the

Not until plaintiff obtains such a special verdict that he is entitled to punitive damages is it necessary or important for him to know defendant's wealth.").

279. *See, e.g.,* Cobb v. Superior Court of L.A., 160 Cal. Rptr. 561, 565 (Ct. App. 1979) ("[Courts need] to balance the right of a defendant to be protected from harassment and to have his right to privacy protected against intrusion into his financial affairs against the right to legitimate discovery when punitive damages are properly a part of a plaintiff's case."); *Bryan*, 453 A.2d at 108 ("In cases involving the requested disclosure of a defendant's financial condition, however, consideration must be given to the defendant's right to privacy and his right to protection from harassment . . . ."); Tennant v. Charlton, 377 So.2d 1169–70 (Fla. 1979) ("If plaintiffs were allowed unlimited discovery of defendants' financial resources in cases where there is no actual factual basis for an award of punitive damages, the personal and private financial affairs of defendants would be unnecessarily exposed and . . . the threat of such exposure might be used . . . to coerce settlements from innocent defendants."); Moran v. Int'l Playtex, Inc., 480 N.Y.S.2d 6, 8 (App. Div. 1984) ("[A] defendant should be protected from the harassing effects of net worth discovery as long as the claim for punitive damages has not been transformed into a special verdict establishing such liability.").

280. State Farm Mut. Auto. Ins. Co. v. Campbell, 538 U.S. 408, 419 (2003); BMW of N. Am., Inc. v. Gore, 517 U.S. 559, 575 (1996).

281. *See supra* note 278 and accompanying text.

world. Part I presented the historical development of the law pertaining to punitive damages. It started with the preliminary controversy about the availability of punitive damages in light of their apparent incompatibility with fundamental principles of civil law. Next it discussed the doctrine's expansion in nineteenth-century case law and the reactionary restraint by twentieth-century state legislation. Finally, it explained that punitive damages are subject to constitutional due process review.

Part II identified the doctrine's main rationales. It refuted the argument that punitive damages are aimed at compensating non-compensable harm and focused on the twin goals of deterrence and retribution. The main question with respect to deterrence was why compensatory damages, which normally ensure internalization of expected loss by potential injurers, were insufficient. Two types of answers seem theoretically sound: (1) punitive damages may be used to secure full internalization when compensatory damages are insufficient due to under-enforcement problems; and (2) punitive damages may be used to eliminate wrongdoers' illicit gains. Neither answer neatly fits the constitutional framework. Thus, retribution seems to be the dominant goal.

Part III considered whether taking into account the defendant's wealth in assessing punitive damages could be consistent with their underlying goals. From a deterrence perspective, the wrongdoer's wealth may be relevant when the victim's loss or the defendant's gain from wrongdoing is unobservable and correlated with the defendant's wealth. The wrongdoer's wealth may also be relevant in assessing punitive damages if they are intended to overcome under-enforcement problems, while potential injurers are risk averse and have no insurance coverage for punitive damages. From a retributive perspective, the sanction must be proportionate to the wrong. According to many theorists, a sanction's severity depends on its effect on the wrongdoer's well-being. And to assess the impact of monetary sanctions on wrongdoers, it is important to know how wealthy they are due to the diminishing marginal utility of individual wealth.

Part IV demonstrated how the defendant's wealth can be integrated into the calculation. After explaining the European experience with day-fines, it proposed a derivative model for assessing punitive damages. First, the court must determine the gravity of the wrong and translate it into corresponding severity units. Then the court must assess the unit value—the wrongdoer's daily income, broadly defined, or a particular fraction thereof. Total damages equal the product of these two variables. If they are greater than compensatory damages, the punitive award will be the difference between total damages and compensatory damages. If total

damages are lower than compensatory damages, punitive damages will be nil.

Part V defended the model from the two relevant theoretical perspectives. From a deterrence perspective, income-dependent punitive damages are defensible in two types of circumstances, mentioned above. First, they may be justified when under-enforcement problems require extra-compensatory damages, and liability insurance does not cover such damages. Second, they are defensible when the wrongdoer's illicit gain is unobservable and correlates with wealth. From a retributive perspective, income-dependent punitive damages may fulfill the requirements of cardinal and ordinal proportionality. Accounting for the wrongdoers' wealth is in fact more consistent with retributive justice since the actual level of the defendant's suffering is not ignored. Part V demonstrated that from both perspectives, the proposed model provides only a second-best solution. Although it does not offer perfect individualization of the monetary sanction, it better serves deterrence and retribution than current law, and has the important advantages of practicability and predictability.

Part VI discussed potential hurdles to the implementation of the new model. It first analyzed potential constitutional challenges. Particularly, it inquired whether the model comports with substantive and procedural due process. Part IV then examined whether statutory caps on punitive damages might thwart the proposed model and therefore need to be modified or repealed. Finally, Part VI reviewed the rules on information discovery in civil cases, in order to assess the practicability of an income-based calculation of damages.

This Article has not resolved all theoretical and practical problems that the proposed model might raise. For example, it does not systematically examine whether and subject to which modifications the model may be applicable to punitive damages awards against corporations. Because some of the main theoretical bases for income-dependent punitive damages hinge on defendants' concave utility functions, it is important to determine whether corporations are also risk-averse.[282] More importantly, a punitive damages award may end up injuring not the firm so much as third parties who had nothing to do with the wrong. For example, if the firm absorbs the sanction, its profits will drop, and individual stockholders might incur a capital loss as well as lower dividends. Adjusting the punitive award to the corporation's wealth or income might have an unwarranted impact on

282. *See, e.g.*, J. Eric Bickel, *Some Determinants of Corporate Risk Aversion*, 3 DECISION ANALYSIS 233 *passim* (2006) (discussing "the degree of risk aversion induced by three rationales for corporate risk management: the cost of financial distress, costly external finance, and the principal-agent relationship between shareholders and management").

third parties with totally different utility functions and levels of wealth. Still, the Article takes an important step in structuring a feasible and theoretically defensible model for assessing punitive damages with wrongdoers' financial condition in mind.