
ARTICLE

HOW SHOULD PUNITIVE DAMAGES WORK?

DAN MARKEL *

In a companion article, I argued that the purpose of punitive damages should be to advance—in part—the public’s interest in retributive justice. These “retributive damages” should be an expressly intermediate sanction, independent of other remedial or penal options. The companion article provided the basic structure of these retributive damages; however, the theoretical nature of the proposal did no more than touch on how they would operate in practice.

This Article addresses the next question: how should punitive damages, including retributive damages, work? This question is especially timely in light of the Supreme Court’s recent decision in Philip Morris USA v. Williams, which held that juries may not consider the harms to nonparties in determining punitive damages awards.

To make punitive damages work, we must first separate retributive damages from other extracompensatory damages meant to achieve cost internaliza-

* Assistant Professor, Florida State University College of Law. A.B., Harvard; M.Phil., University of Cambridge; J.D., Harvard. I am grateful to the many people who have shared comments and conversations on this punitive damages project as a whole, and especially to the folks who read earlier versions of this particular piece within that endeavor: Kelli Alces, Miriam Baer, Stephanos Bibas, Michael Cahill, Bennett Capers, Danielle Citron, Tom Colby, Robin Craig, Brian Galle, Tom Galligan, Mark Geistfeld, David Hoffman, Keith Hylton, Adam Kolber, Erik Knutsen, Julian Ku, Katy Kuh, Alyssa Lathrop, Ethan Leib, Wayne Logan, Marcia McCormick, Richard Myers, Michael O’Hear, Mitch Polinsky, BJ Priestler, Marty Redish, Adam Richardson, Alice Ristroph, Cathy Sharkey, Jason Solomon, Tony Sebok, Manuel Utset, Steve Vladeck, Verity Winship, Ekow Yankah, and Ben Zipursky. I also thank participants at workshops at FSU, Houston, Brooklyn, SMU, SEALS, Marquette, and CLEA. Despite all this help, errors remain mine; if you spot them, I’d be grateful for comments. Please send them, along with any other reactions, to markel@post.harvard.edu. Last, and as always, I’m grateful to Wendi Adelson for her enduring patience, love, and laughter.

tion or to vindicate the victim's dignity interests. Because these three purposes are distinct, conflating them carries the danger of both under- and overprotection of various defendants. Once we understand these purposes and the distinctions between them, we should be able to map them on to our existing institutional design for civil damages. This Article begins that important task, first, by explaining why and how defendants should enjoy certain procedural protections depending on which purpose the damages serve, and second, by addressing two critical implementation issues associated with this pluralistic scheme of extracompensatory damages: insurance and settlement.

INTRODUCTION	1386
I. THE ACTION IN PUNITIVE DAMAGES	1390
A. <i>Recent Developments in Punitive Damages</i>	
<i>Law and Scholarship</i>	1390
1. The Law	1390
2. The Normative Scholarship	1393
B. <i>The Basic Structure of Retributive Damages: A Recap</i>	1398
II. PLURALISM ABOUT PURPOSES	1403
A. <i>Cost Internalization and Deterrence Damages</i>	1404
1. Overview	1404
2. Deterrence Damages After <i>Philip Morris</i>	1407
3. Standard of Review for Deterrence Damages	1410
4. Who Receives Deterrence Damages	1411
5. Reconciling Deterrence Damages with	
Retributive Damages	1412
B. <i>Aggravated Damages for Victim Vindication</i>	1414
1. Overview	1414
2. Possible Limits on Aggravated Damages	1417
3. Reconciling Aggravated Damages and	
Retributive Damages	1418
4. Reconciling Aggravated Damages and	
Deterrence Damages	1419
5. Standard of Review for Aggravated Damages	1419
C. <i>A Pluralistic Extracompensatory Damages Scheme:</i>	
<i>Applications</i>	1420
III. PROCEDURAL PROTECTIONS, PLURALISM, AND	
PUNITIVE DAMAGES	1422
A. <i>The Debate over Punitive Damages and</i>	
<i>Procedural Safeguards</i>	1423
1. Some Background	1423
2. Pluralism and Procedural Safeguards	1424
a. <i>Aggravated Damages and</i>	
<i>Deterrence Damages</i>	1425

b. <i>Retributive Damages</i>	1427
i. Colby's Error	1428
ii. Galanter and Luban's Errors	1433
iii. Why an Intermediate Compromise?	1435
B. <i>Retributive Damages as Intermediate Sanction:</i>	
<i>Applying the Logic to Procedural Safeguards</i>	1436
1. Standard of Proof	1436
2. Availability and Standard of Appellate Review	1437
3. Is a Unanimous Jury Required?	1438
a. <i>Judges or Juries</i>	1438
b. <i>Unanimity</i>	1441
4. Procedural Bifurcation	1442
5. Confrontation of Adverse Witnesses and Compulsory Process	1442
6. Access to Counsel	1443
7. Privilege Against Self-Incrimination	1445
8. Duplicative Punishment Concerns in Simple Litigation Contexts	1450
a. <i>Double Jeopardy Doctrine and Punitive Damages</i>	1450
b. <i>The Duplicative Punishment Problem</i>	1452
c. <i>Solving the Overkill Problem in the Simple Context</i>	1454
i. Should Retributive Damages Follow Criminal Prosecution?	1454
ii. Should Criminal Prosecution Follow Retributive Damages?	1457
iii. Can Retributive Damages Coincide with Criminal Prosecution?	1459
9. Excessiveness Review Under the Eighth Amendment	1461
IV. IMPLEMENTING EXTRACOMPENSATORY DAMAGES	1462
A. <i>Should Insurance Be Available—If So, When?</i>	1462
1. Insurance for Insurable Reckless Risks	1463
2. Does Access to Insurance Blunt the Retributive Condemnation?	1469
B. <i>Settlement, Transparency, and Accountability</i>	1471
1. The Danger of Sweetheart Deals to Retributive Damages	1471
2. Settlement, Insurance, and the Public Interest	1476
CONCLUSION	1478

APPENDIX: INSTRUCTIONS FOR ASSESSING EXTRACOMPENSATORY

DAMAGES	1479
A. <i>Retributive Damages</i>	1479
B. <i>Aggravated Damages for Repairing</i> <i>Personal Dignity Harms</i>	1481
C. <i>Deterrence Damages</i>	1482

INTRODUCTION

What are punitive damages for? In a recent article, I argued that states should understand and restructure punitive damages, in part, to advance the public's interest in retributive justice.¹ For clarity's sake, I called such damages "retributive damages" to distinguish them from extracompensatory damages designed to pursue other goals, such as cost internalization or victim vindication. Although that article explained the normative rationale and basic structure for retributive damages as an intermediate sanction and why society should want that sanction independent of other remedial or penal options, the theoretical nature of the proposal merely scratched the surface of how such damages should operate in practice.

This Article, the second in a series, addresses the next logical question: how should punitive damages work? Both questions are especially timely in light of the Supreme Court's recent decision in *Philip Morris USA v. Williams*.² This Article focuses on a range of important implementation issues left previously unaddressed—e.g., are any procedural safeguards for defendants facing punitive damages necessary and, if so, which ones and why? How should such damages interact with criminal prosecutions? Should an insurance market be permitted? How should settlement be regulated?

A discussion of these implementation issues, however, should only occur after explaining why "punitive damages" should be relabeled "extracompensatory damages," as well as why extracompensatory

¹ See Dan Markel, *Retributive Damages: A Theory of Punitive Damages as Intermediate Sanction*, 94 CORNELL L. REV. 239 (2009) [hereinafter Markel, *Retributive Damages*].

² 549 U.S. 346 (2007). In *Philip Morris*, the Court held that the Due Process Clause forbids juries from figuring in the harms to nonparties in determining the amount of punitive damages that a defendant must pay. The Court also addressed punitive damages more recently in *Exxon Shipping Co. v. Baker*, 128 S. Ct. 2605 (2008). Because the *Exxon* case was resolved under federal maritime law, however, its significance for this project is not nearly as substantial as the Court's constitutional decision in *Philip Morris*.

damages should treat retributive damages separately from damages meant to pursue nonretributive goals, such as compensating the injury to a victim's dignity or facilitating the pursuit of cost internalization to the extent permitted after *Philip Morris*. Because these purposes are distinct, a jurisdiction that conflates them risks both under- and overprotecting various defendants. Once we correctly understand these distinctive purposes, our institutional design should map them appropriately. This Article begins that task, so that states can build a pluralistic framework that is both attractive from a policy perspective and compatible with constitutional values and doctrine.

Unfolding in four Parts, the Article begins in Part I by furnishing some background to the recent law and scholarship on the purposes of punitive damages. Readers of the earlier companion article may profitably skim that Part, the role of which is primarily to set the stage for the discussion in this Article. Part II offers a structure that tries to disaggregate the distinct purposes of extracompensatory damages in a manner that is both constitutionally feasible and attractive from a policy perspective. Rather than argue that retributive damages should play an exclusive or primary role,³ I urge a pluralistic approach for state legislatures to consider. Under such a legislative framework, extracompensatory damages would be available separately, if necessary, for retributive, cost internalization, and personal dignity repair (or what I more frequently call victim vindication) purposes. Correspondingly, there would be three kinds of extracompensatory damages permitted under the relevant statutes: retributive, deterrence, and aggravated damages.⁴ Much work has already been performed on the conceptualization and implementation of deterrence and aggravated damages.⁵ While this Article builds on those achievements, its com-

³ Cf. Marc Galanter & David Luban, *Poetic Justice: Punitive Damages and Legal Pluralism*, 42 AM. U. L. REV. 1393, 1451 (1993) ("Efficiency plays no role in the normative universe of punitive damages as we conceive of it.").

⁴ I use the terms "deterrence damages" and "aggravated damages" in ways that might seem idiosyncratic to some readers. Deterrence damages signify those damages meant to facilitate optimal deterrence or "cost internalization," not "complete deterrence." Aggravated damages are meant to account for, and thereby vindicate, that part of a victim's injury, not already compensated in all jurisdictions, for insults to one's personal dignity. See *infra* Part II.

⁵ See, e.g., *Ciraolo v. City of N.Y.*, 216 F.3d 236, 244-47 (2d Cir. 2000) (Calabresi, J., concurring) (advocating nonpunitive extracompensatory damages for the sake of cost internalization); Bruce Chapman & Michael Trebilcock, *Punitive Damages: Divergence in Search of a Rationale*, 40 ALA. L. REV. 741, 805-26 (1989) (discussing cost internalization and aggravated damages for compensation of dignity harms); Thomas C. Galligan, Jr., *Augmented Awards: The Efficient Evolution of Punitive Damages*, 51 LA. L. REV. 3, 12-13

parative advantage and its focus are on the procedural safeguards and mechanisms necessary to implement the public interest in retributive justice, a point that has been demonstrably underexamined in the existing literature.⁶ States should employ a damages scheme that sharpens juries' and judges' decision making by disaggregating the retributive and nonretributive functions that are normally conflated in awards of punitive damages.

Part III turns to the issue of which procedural safeguards should be in place when awarding these types of extracompensatory damages. A number of scholars have argued that punitive damages, insofar as they serve public retributive goals, are unconstitutional because civil defendants lack procedural safeguards such as those provided in criminal cases.⁷ The problem with this claim is that it mistakenly sug-

(1990) (describing cost internalization damages); A. Mitchell Polinsky & Steven Shavell, *Punitive Damages: An Economic Analysis*, 111 HARV. L. REV. 869, 873-76 (1998) (developing a set of principles for determining cost internalization damages); Catherine M. Sharkey, *Punitive Damages as Societal Damages*, 113 YALE L.J. 347, 401-02 (2003) (discussing how a "compensatory societal damages" framework improves upon limitations of Polinsky and Shavell's method for ensuring cost internalization); Mark A. Geistfeld, *Punitive Damages, Retribution, and Due Process*, 81 S. CAL. L. REV. 263, 269-74, 296-309 (2008) (conceiving the role of punitive damages primarily in terms of individual victim vindication and proposing methods to ensure that such vindication satisfies due process); Anthony J. Sebok, *Punitive Damages: From Myth to Theory*, 92 IOWA L. REV. 957, 961, 1023-29 (2007) (arguing that punitive damages should be regarded as "a form of state-sanctioned revenge"); Benjamin C. Zipursky, *A Theory of Punitive Damages*, 84 TEX. L. REV. 105 (2005) (offering a conceptual and interpretive account of punitive damages as victim vindication). More recently, Tom Colby has joined the group defending a victim vindication interpretation. See Thomas B. Colby, *Clearing the Smoke from Philip Morris v. Williams: The Past, Present, and Future of Punitive Damages*, 118 YALE L.J. 392, 434 (2008) ("Punitive damages vindicate the dignity of an individual victim by allowing her to punish the defendant for committing a humiliating or insulting tort upon her.").

⁶ See Cass R. Sunstein et al., *Assessing Punitive Damages (with Notes on Cognition and Valuation in Law)*, 107 YALE L.J. 2071, 2085 (1998) ("Regrettably, the legal culture lacks a full normative account of the relationship between retributive goals and punitive damages."). My goal in *Retributive Damages*, *supra* note 1, was primarily to address that lacuna.

⁷ See, e.g., Thomas B. Colby, *Beyond the Multiple Punishment Problem: Punitive Damages as Punishment for Individual, Private Wrongs*, 87 MINN. L. REV. 583, 643-50 (2003) (arguing that punitive damages for public purposes and without procedural safeguards would be "indistinguishable from criminal punishment"); Jeffrey W. Grass, *The Penal Dimensions of Punitive Damages*, 12 HASTINGS CONST. L.Q. 241, 242 (1985) ("By punishing defendants rather than compensating plaintiffs [the] purpose [of punitive damages] falls squarely under the ambit of criminal law." (footnote omitted)); John Calvin Jeffries, Jr., *A Comment on the Constitutionality of Punitive Damages*, 72 VA. L. REV. 139, 139 (1986) (noting the lack of procedural safeguards in awarding punitive damages); Malcolm E. Wheeler, *The Constitutional Case for Reforming Punitive Damages Procedures*, 69 VA.

gests that criminal procedural safeguards apply like a binary switch that toggles between “on” and “off.” In fact, the extent of protection provided by many procedural safeguards operates on a continuum marked by the severity of the punishment imposed.⁸ The same logic should inform the design of safeguards for defendants facing retributive damages—or so I argue.

After introducing the debate over this issue, Part III examines which safeguards, if any, would be necessary, from a constitutional point of view, for aggravated and deterrence damages. When defendants are facing nonretributive damages, the same precautionary measures do not necessarily apply; however, certain measures might still be warranted to ensure fidelity to federalism principles and basic procedural fairness. More importantly, if states adopt a retributive damages regime, they should, if they have not already done so, also add some heightened procedural safeguards for defendants facing retributive damages.

As explained in Part III, the extent of such protections should fall roughly between the kind of protection we confer upon defendants in cases involving compensatory damages and the kind of protection we confer upon defendants in criminal cases involving modest sanctions such as criminal fines. But consistency with constitutional mandates is not the only goal—the level of procedural protections should also be faithful to the basic values underlying retributive justice. These organizing principles are applied, albeit in a preliminary fashion, to matters such as exposure to duplicative punishment for the same misconduct, standards of proof, standards of review, the privilege against self-incrimination, access to counsel, jury trial rights, and bifurcation of evidence going to liability and evidence of a defendant’s financial condition.

Finally, Part IV addresses two basic issues related to implementing this disaggregated scheme of extracompensatory damages. Specifically, it examines how this structure would intersect with insurance and settlements. There are, of course, other issues that warrant discussion: First, is it intelligible and attractive to impose retributive damages against corporate defendants and other entities? Second, how does one deal with the array of federalism, duplicative punishment, and bankruptcy concerns of defendants whose misconduct af-

L. REV. 269, 270 & n.5 (1983) (citing sources that discuss the application of criminal procedural safeguards to defendants facing the imposition of punitive damages).

⁸ See *infra* Section III.A.

fects multiple parties across jurisdictional lines? I leave these related areas of inquiry to resolve in the next article in this series, *Punitive Damages and Complex Litigation*.⁹ In the meantime, my current goal is to lend clarity to the basic muddle of punitive damages law and policy by suggesting some solutions for what might be called the “simple litigation” context—i.e., where *X* commits *Y* misconduct and, in so doing, threatens or causes injury only to *Z*.

I. THE ACTION IN PUNITIVE DAMAGES

A. *Recent Developments in Punitive Damages Law and Scholarship*

1. The Law

Without rehashing the entire history of punitive damages in America,¹⁰ it is important to appreciate the major developments in the field of American punitive damages law and theory.¹¹ Early Anglo-American courts awarded “exemplary” damages for a range of purposes; in some cases, they served as compensation to a plaintiff for suffering “intangible wrongs,”¹² while in other cases these damages served as punishment designed to make an example of the defendant.¹³ On the conventional account, the compensatory function of punitive damages has waned as the scope of compensatory damages has expanded to include such “intangibles” as hurt feelings and indignities.¹⁴

⁹ Dan Markel, *Punitive Damages and Complex Litigation* (unpublished manuscript, on file with author) [hereinafter Markel, *Punitive Damages*]; see also Dan Markel & Gregg Polsky, *Taxing Punitive Damages* (unpublished manuscript, on file with author) (addressing proper taxation policies associated with this proposed punitive damages scheme and other related tax issues).

¹⁰ See *Pac. Mut. Life Ins. Co. v. Haslip*, 499 U.S. 1, 25-28 (1991) (Scalia, J., concurring in the judgment) (furnishing a brief history of punitive damages).

¹¹ This Section draws upon the discussion in Part I of Markel, *Retributive Damages*, *supra* note 1. Readers familiar with that work may skim or skip this Section. Those seeking a more comprehensive introduction to the topic might also consider looking at Anthony J. Sebok, *Punitive Damages in the United States*, in *PUNITIVE DAMAGES* (Helmut Koziol & Vanessa Wilcox eds., forthcoming 2009).

¹² See Martin H. Redish & Andrew L. Mathews, *Why Punitive Damages Are Unconstitutional*, 53 EMORY L.J. 1, 13-16 (2004) (discussing early English cases in which the plaintiff demonstrated a dignitary harm that would otherwise remain uncompensated in the absence of exemplary damages).

¹³ See *Exxon Shipping Co. v. Baker*, 128 S. Ct. 2605, 2620 (2008) (providing examples of the application of this punitive damages function).

¹⁴ *Cooper Indus., Inc. v. Leatherman Tool Group, Inc.*, 532 U.S. 424, 438 n.11 (2001).

There remains concern among some scholars, however, that, contrary to the conventional wisdom, the scope of compensatory damages in contemporary tort law does not encompass the “aggravated” nature of the injury to a plaintiff when her dignity has been insulted by the defendant’s misconduct.¹⁵

In the last two decades, the Supreme Court has emphasized that punitive damages should be principally understood as “quasi-criminal” “private fines” designed to punish the defendant and deter the misconduct at issue.¹⁶ Unfortunately, although courts frequently invoke the purposes of retribution and deterrence,¹⁷ they often offer little analysis of these purposes or their implications. For instance, courts rarely instruct juries to parse the amount of money necessary to punish the defendant and the amount necessary to achieve deterrence. Moreover, they rarely distinguish between optimal deterrence (aiming at cost internalization) and complete deterrence (aiming at preventing the commission of similar misconduct in the future).¹⁸

¹⁵ See Anthony J. Sebok, *What Did Punitive Damages Do? Why Misunderstanding the History of Punitive Damages Matters Today*, 78 CHI.-KENT L. REV. 163, 205 (2003) (“If punitive damages served a compensatory function [in early cases], it would have been for a category of injury that is still not considered compensable by contemporary tort law, namely the injury of insult that wounds or dishonors.”).

¹⁶ See, e.g., *Cooper Indus.*, 532 U.S. at 432; *Gertz v. Robert Welch, Inc.*, 418 U.S. 323, 350 (1974) (describing punitive damages as “private fines” designed to punish and deter “reprehensible conduct”). Nevertheless, if a state passed an enabling statute authorizing punitive damages expressly for the purposes of compensating a plaintiff or society, the Supreme Court would probably not say that such purposes are inherently improper or unconstitutional. Cf. Sharkey, *supra* note 5, at 391-92 (elaborating a proposal for compensatory societal damages).

¹⁷ See, e.g., *Exxon Shipping Co.*, 128 S. Ct. at 2621 (2008) (“[T]he consensus today is that punitives are aimed not at compensation but principally at retribution and deterring harmful conduct.”); *Philip Morris USA v. Williams*, 549 U.S. 346, 352 (2007) (“This Court has long made clear that ‘[p]unitive damages may properly be imposed to further a State’s legitimate interests in punishing unlawful conduct and deterring its repetition.’” (alteration in original) (quoting *BMW of N. Am., Inc. v. Gore*, 517 U.S. 559, 568 (1996))).

¹⁸ For those unfamiliar with these terms, under an optimal deterrence or cost internalization regime, a defendant who pays the costs of her tortious activity should be able to continue pursuing that activity. The decision to seek cost internalization for an activity is predicated on a determination that the gains to the defendant are socially licit. The basis for concerns about cost internalization is further described by Judge Calabresi in *Ciraolo v. City of New York*. See *infra* note 68. By contrast, complete deterrence endeavors to remove the incentive for the defendant to undertake that conduct altogether because the gains from such conduct are deemed illicit. See Keith N. Hylton, *Punitive Damages and the Economic Theory of Penalties*, 87 GEO. L.J. 421, 421 (1998). Complete deterrence is a goal that, if realized, would entail that zero instances of the particular misconduct would occur. Others have used different terms to distinguish

As this consensus on the purposes of punitive damages—retribution and deterrence—has emerged, the Supreme Court has begun to establish a constitutional framework for regulating such damages. The Court's requirements can be summed up in six rules.

First, when courts review the constitutionality of punitive damages awards, the most important factor that they must consider is the degree of reprehensibility of the defendant's misconduct.¹⁹ Second, reviewing courts must also consider whether "the disparity between the actual or potential harm suffered by the plaintiff and the punitive damages award" is constitutionally excessive.²⁰ More controversially, in *State Farm v. Campbell*, the Court established a presumption that "in practice, few awards exceeding a single-digit ratio between punitive and compensatory damages, to a significant degree, will satisfy due process."²¹ Third, reviewing courts should consider "the disparity between the punitive damages award and the 'civil penalties authorized or imposed in comparable cases.'"²² Fourth, reviewing courts, under the Supreme Court's recent *Philip Morris* decision, must ensure that the jury does not impose on defendants an amount that includes the harms to nonparties to the litigation.²³ One might see this rule as related to the Court's stated interest in ensuring that one state refrain from punishing defendants for conduct lawfully performed in another state.²⁴ Fifth, judicial review of a jury's award of punitive damages

between optimal and complete deterrence. See, e.g., GUIDO CALABRESI, THE COSTS OF ACCIDENTS: A LEGAL AND ECONOMIC ANALYSIS 68-69 (1970) (distinguishing between general (permissive) deterrence and specific (prohibitory) deterrence); Robert Cooter, *Prices and Sanctions*, 84 COLUM. L. REV. 1523, 1524-31 (1984) (distinguishing between the "pricing" and "sanctioning" of different types of behavior). Lastly, it is worth noting that complete deterrence, which calls for a sentence designed to signal that the conduct is prohibited, should not be confused with complete enforcement, which would call for sufficient resources to reduce the misconduct to zero.

¹⁹ See *State Farm Mut. Auto. Ins. Co. v. Campbell*, 538 U.S. 408, 419 (2003) (restating the rule that punitive damages only be awarded where a defendant's conduct is so reprehensible that it justified an award in addition to compensatory damages). The Court has further specified a number of factors that contribute to a determination of reprehensibility. See *BMW of N. Am., Inc. v. Gore*, 517 U.S. 559, 575-77 (1996) (listing factors that indicate a greater degree of reprehensibility).

²⁰ *State Farm*, 538 U.S. at 418.

²¹ *Id.* at 425.

²² *Id.* at 428 (quoting *Gore*, 517 U.S. at 575).

²³ See *Philip Morris USA v. Williams*, 549 U.S. 346, 353 (2007).

²⁴ *State Farm*, 538 U.S. at 421.

must be available.²⁵ Finally, appellate review of punitive damages must apply a *de novo* standard of review of the jury's award, at least in a federal case.²⁶

Importantly, although the Court developed these rules to improve fair notice and proportionality for defendants facing these sanctions,²⁷ the Court has not extended to defendants the protections normally applicable in the criminal law context. Indeed, defendants have no safeguards established under the Constitution beyond what has been discussed,²⁸ though most states have introduced a flurry of caps, multipliers, and other limits on punitive damages.²⁹

2. The Normative Scholarship

The complex and rapidly evolving nature of punitive damages law has attracted the attention of scholars from a variety of disciplines.³⁰ In terms of normative approaches to punitive damages, a number of scholars, such as Professors Polinsky and Shavell, think that extracompensatory damages should focus on advancing the goal of optimal deterrence (or what I also call "cost internalization" or "deterrence").³¹ Under this framework, which I discuss in greater detail in Part II, a defendant's culpability or state of mind is immaterial to her obligation to pay for the harms that she causes. Instead, what matters is whether any likelihood exists that the defendant would evade paying compensation for the harms that she caused. If there is such a possibility, then the amount of punitive damages should be calibrated accord-

²⁵ See *Honda Motor Co. v. Oberg*, 512 U.S. 415, 432 (1994) (holding that Oregon's denial of adequate judicial review of punitive damages violated the Fourteenth Amendment Due Process Clause).

²⁶ *Cooper Indus., Inc. v. Leatherman Tool Group, Inc.*, 532 U.S. 424, 435-36 (2001).

²⁷ See *State Farm*, 538 U.S. at 416-17.

²⁸ See Markel, *Retributive Damages*, *supra* note 1, at 249-53.

²⁹ See Michael L. Rustad, *The Closing of Punitive Damages' Iron Cage*, 38 LOY. L.A. L. REV. 1297 app. A (2005) (surveying punitive damages practices across states and the District of Columbia).

³⁰ See Markel, *Retributive Damages*, *supra* note 1, at 242-43 nn.4-9 (citing different approaches taken by scholars in considering extracompensatory damages).

³¹ See Polinsky & Shavell, *supra* note 5, at 897-98 (arguing that courts should use damages determinations to facilitate optimal deterrence by applying the "punitive damages multiplier" developed by the authors). I recognize that by conflating deterrence with optimal deterrence (or cost internalization), I am implicitly obscuring the work of some economists who view this law through the prism of complete deterrence. See, e.g., Hylton, *supra* note 18, at 423 (arguing that the optimal deterrence model should be used in limited cases and that complete deterrence should be the goal in most situations). See *infra* note 58 for a partial explanation of the reasons for this move.

ingly.³² A number of judges have embraced the basic insight undergirding this approach.³³ However, as Professor Sharkey points out, a total cost internalization approach would not necessarily provide victims compensation for their losses.³⁴

In contrast to the cost internalization school, other scholars analyze punitive damages law in terms of how such damages might vindicate a victim's dignity and autonomy interests, which may have been injured or insulted by the defendant's misconduct.³⁵ Since these victim vindication approaches effectively legitimize the utilization of enhanced awards to repair the injury that the defendant's misconduct caused to the plaintiff's dignity, these damages are more precisely la-

³² See Polinsky & Shavell, *supra* note 5, at 887-96 ("[I]f a defendant can sometimes escape liability for the harm for which he is responsible, the proper magnitude of damages is the harm the defendant has caused, multiplied by a factor reflecting the probability of his escaping liability." (emphasis omitted)). But see Keith N. Hylton & Thomas J. Miceli, *Should Tort Damages Be Multiplied?*, 21 J.L. ECON. & ORG. 388 (2005) (registering skepticism toward the use of the multiplier approach in the context of civil damages based on concerns regarding the supply of lawsuits and the cost of litigation). Professor Sharkey's proposal for compensatory societal damages was designed to address perceived shortcomings with the Polinsky-Shavell model for achieving cost internalization. See Sharkey, *supra* note 5, at 368-70 (identifying problems with the use of a strict punitive damages multiplier, such as the failure to include cases involving "diffuse" harms). The suggestions for compensating society for defendants' more diffuse harms to society, however, are not likely to survive *Philip Morris* for reasons similar to the argument that I make about Professor Colby's claims in Part II. That is, notwithstanding the avowedly nonpunitive rationale behind compensatory societal damages, my sense is that the Supreme Court is unlikely to think that such an approach satisfies due process because it allows an award of damages for harms against persons or entities that a defendant cannot litigate against specifically. See Michael B. Kelly, *Do Punitive Damages Compensate Society?*, 41 SAN DIEGO L. REV. 1429, 1433-35 (2004) (raising concerns that Sharkey's proposal will founder on due process grounds).

³³ See Sharkey, *supra* note 5, at 372 n.71 (collecting cases in which courts have expressed their approval of the multiplier approach).

³⁴ See *id.* at 390-91; see also *infra* note 91.

³⁵ See, e.g., Colby, *supra* note 5, at 434; Galanter & Luban, *supra* note 3, at 1432 (discussing the view that culpably wronging a person expresses that the victim is of less value than the wrongdoer); Geistfeld, *supra* note 5, at 269-74 (advancing the idea of punitive damages primarily in terms of the private interest in victim vindication of "tort rights"); John C.P. Goldberg, *Tort Law for Federalists (and the Rest of Us): Private Law in Disguise*, 28 HARV. J.L. & PUB. POL'Y 3, 7 (2004) ("What is at stake in [punitive damages] is not [a state's] interests in obtaining retribution on behalf of its citizens or in deterring sharp business practices, but the [plaintiffs'] interest in vindicating their rights not to be mistreated in the way that they were. . . . [These rights include] providing [the plaintiffs] with *satisfaction*—a remedy adequate to acknowledge and avenge [the defendant's] predatory conduct towards them."); Sebok, *supra* note 5, at 1007-15 ("[T]he private right whose violation grounds [a punitive damages] award is the private right not to have one's dignity violated."); Zipursky, *supra* note 5, at 151-53 (discussing the plaintiff's right "to be punitive" toward a defendant liable for a willful wrong).

beled “aggravated” damages, as they are called in various common law jurisdictions.³⁶ Some victim vindication theorists have defended large parts of extant punitive damages common law on the ground that these practices serve as vehicles by which victims or their allies can persuade juries to avenge victims’ interests through ad hoc, and therefore unpredictable, awards of money damages.³⁷ Indeed, for some social justice tort theorists, common law, jury-driven punitive damages practices serve as means for ordinary people to fight malfeasant entities and their lobbyists seeking business-friendly tort reform.³⁸

Some of these scholars, such as Galanter and Luban, drawing on the work of Jean Hampton’s victim vindication justification for punishment,³⁹ view themselves as committed to the goals or values of retributive justice.⁴⁰ But as emphasized in the insightful interpretive ac-

³⁶ See Chapman & Trebilcock, *supra* note 5, at 763 (“Where there is already injury in place that the law recognizes as damages, this added ‘insult’ to injury would count more accurately as ‘aggravated,’ than as punitive, damages.”).

³⁷ See Kaimipono David Wenger & David A. Hoffman, *Nullificatory Juries*, 2003 WIS. L. REV. 1115, 1119 (defending the role of juries in “protect[ing] us from rule by legal economists” through “relatively unconstrained punitive awards”). Galanter and Luban also endorse (at least implicitly) a jury’s imposing punitive damages against a defendant, in a single case, for all the harm that the defendant’s misconduct caused persons in similar situations. See, e.g., Galanter & Luban, *supra* note 3, at 1436-38 (providing examples of “expressive defeat” of defendants through punitive damages). Galanter and Luban also think that judges should extend “great deference” to jury determinations because of juries’ special competence in articulating “the community’s ‘message’ through the medium of damages.” *Id.* at 1439. My view circumscribes jury decision making considerably more.

³⁸ See Rustad, *supra* note 29, at 1301 (characterizing tort reform of punitive damages as “special legislation to help corporate America”). See generally THOMAS H. KOENIG & MICHAEL L. RUSTAD, IN DEFENSE OF TORT LAW (2001) (providing a paradigmatic account of the social justice theory of tort law); Richard L. Abel, *Questioning the Counter-Majoritarian Thesis: The Case of Torts*, 49 DEPAUL L. REV. 533 (1999) (arguing that legislatures are beholden to special interests, while courts are more likely to be focused on the common good in the tort law context); see also David F. Partlett, *The Republican Model and Punitive Damages*, 41 SAN DIEGO L. REV. 1409 (2004) (defending a robust role for juries in punitive damages awards on the basis of republican theory).

³⁹ See Jean Hampton, *Correcting Harms Versus Righting Wrongs: The Goal of Retribution*, 39 UCLA L. REV. 1659, 1685-98 (1992) (arguing that conduct that expresses disrespect and does damage to “the value of a victim” warrants a punitive response to vindicate the victim’s moral worth).

⁴⁰ See Galanter & Luban, *supra* note 3, at 1432-35. For reasons that I explained in the first article in this series, I view Galanter and Luban’s account of punitive damages as primarily (though not exclusively) a victim vindication account, not a retributive justice account. See Markel, *Retributive Damages*, *supra* note 1, at 255 n.62. I note also that a number of other punitive damages scholars, including Professors Colby, Sebok, Zipursky, and Geistfeld, have claimed to be influenced by Professor Hampton’s work and, to varying degrees, have identified themselves as interested in developing the re-

counts of tort law and punitive damages by Zipursky and Sebok, the tort system conventionally empowers victims either to *pursue* punitive damages or to *forbear pursuing* such damages.⁴¹ That is important because it shows that no one forces (punitive) damages on the victim in the common law approach. Rather, the decision to seek legal recourse (or not) permits the victim to exercise her autonomy and seek repair to her dignity interests. The same may be said for allowing victims to have almost unfettered control over settlements with defendants.

As I explain further in Section II.B, however, these two practices reveal an important gap between victim vindication accounts and the interests underlying a retributivist account, properly understood. Retributivists, as I explained in *Retributive Damages*, have strong reasons to give weight to the reduction of both Type I false-positive errors—in which people are mistakenly punished (or excessively punished relative to comparable offenders)—and Type II false-negative errors—in which wrongdoers escape their punishment altogether (or receive too lenient a punishment as compared to other similar offenders in the jurisdiction).⁴² Importantly, the victim vindication accounts say little about the need to build a system that tries to reduce all four categories of Type I and II errors.

Indeed, to the extent that victim vindication supporters invoke retributive justice values to bolster their accounts, this silence is a real weakness.⁴³ After all, the failures to defend procedural safeguards and

lationship between punitive damages and retributive justice. My own view is that their interests and values are better described as consistent with “victim vindication,” and less so with retributive justice, properly understood as a practice of state punishment interested in developing institutions that promote equality and rule-of-law values in the reduction of Type I and Type II punishment errors. For discussion of my view of the proper role of victims in retributive theory, see, for example, Dan Markel, *State, Be Not Proud: A Retributivist Defense of the Commutation of Death Row and the Abolition of the Death Penalty*, 40 HARV. C.R.-C.L. L. REV. 407, 452-57 (2005) [hereinafter Markel, *State, Be Not Proud*].

⁴¹ See Sebok, *supra* note 5, at 1005-06 (“Plaintiffs who may have a valid legal claim for punitive damages are under no obligation to pursue them. In theory, a plaintiff could request a sanction smaller than what justice might otherwise require the wrongdoer to repay.” (footnote omitted)); *id.* at 1028-29 (“[O]ne element of the repair of wrongful losses in tort is the active role of the victim in determining the appropriate remedy for her case of wrongful loss.”); Benjamin C. Zipursky, *Civil Recourse, Not Corrective Justice*, 91 GEO. L.J. 695, 733 (2003) [hereinafter Zipursky, *Civil Recourse*] (“[T]ort cases ultimately require courts to respond to demands by plaintiffs . . .”); Zipursky, *supra* note 5, at 152 (“The state permits the plaintiff to seek and to receive [punitive damages], but the state is not in the driver’s seat.”).

⁴² See Markel, *Retributive Damages*, *supra* note 1, at 247, 266.

⁴³ To its credit, Professor Sebok’s state-sanctioned revenge account is consistent with a desire to reduce “piling on” (or Type I overpunishment) errors that occur

to create meaningful guidelines for cabining jury discretion and judicial review are recipes for Type I error creation. Moreover, giving only victims the right to pursue retributive damages or giving all victim-plaintiffs the unfettered authority to settle a case involving allegations of reckless or malicious misconduct enables more Type II errors.⁴⁴ This should be of concern to nonretributivists as well: certainty of punishment, perhaps more than severity of punishment, has, for the last generation or so, been thought to have an appreciable effect on reducing misconduct.⁴⁵

If we want a retributivist scheme of punitive damages, it has to reflect some concern for reducing all four types of Type I and Type II errors. Of course, a pluralistic scheme of extracompensatory damages should be designed to provide space for the pursuit of both cost internalization and victim vindication as well, and Part II says more about how to do that.⁴⁶ But since these two goals have received sub-

through introducing evidence of harms to strangers to the litigation. See Sebok, *supra* note 5, at 1031-35. But Sebok doesn't address the *public's* interest in reducing Type II errors of either sort, or the procedural safeguards necessary to prevent Type I errors of the mistaken-punishment sort. Similarly, for cases involving fatal risks, the methodology proposed by Professor Geistfeld, *see* Geistfeld, *supra* note 5, is helpful in ensuring some evenhandedness across cases involving certain tort victims. That said, this methodology says little about how to reduce the gamut of Type I and Type II errors outside the relatively narrow but important context of victims facing fatal risks; moreover, even in the context of fatal risks, Geistfeld's account is quiet about the need for reducing Type II errors involving nonpunishment.

⁴⁴ Such Type II errors leading to underenforcement are rife. See Michael J. Saks, *Do We Really Know Anything About the Behavior of the Tort Litigation System—And Why Not?*, 140 U. PA. L. REV. 1147, 1183 (1992) ("One of the most remarkable features of the tort system is how few plaintiffs there are. A great many potential plaintiffs are never heard from by the injurers or their insurers."); *see also* Richard L. Abel, *The Real Torts Crisis—Too Few Claims*, 48 OHIO ST. L.J. 443 (1987); Marc Galanter, *Real World Torts: An Antidote to Anecdote*, 55 MD. L. REV. 1093, 1159 (1996) (stating that "relatively few" tort claims are brought to court and that, even if more claims were filed, the tort system may not have the capacity to handle them).

⁴⁵ See ANDREW VON HIRSCH ET AL., *CRIMINAL DETERRENCE AND SENTENCE SEVERITY: AN ANALYSIS OF RECENT RESEARCH* 45-48 (1999) (examining empirical literature and noting that there exists a correlation between increased certainty of punishment and decreased crime rates but that the evidence showing a correlation between severity of punishment and crime rates is comparatively weaker); Jeffrey Grogger, *Certainty vs. Severity of Punishment*, 29 ECON. INQUIRY 297, 308 (1991) ("The results point to large deterrent effects emanating from increased certainty of punishment, and much smaller, and generally insignificant effects, stemming from increased severity of sanction.").

⁴⁶ Punitive damages might be thought to pursue a mixture of other goals as well. See generally Dorsey D. Ellis, Jr., *Fairness and Efficiency in the Law of Punitive Damages*, 56 S. CAL. L. REV. 1, 3 (1982) ("At least seven purposes for imposing punitive damages can be gleaned from judicial opinions and the writings of commentators: (1) punishing the defendant; (2) deterring the defendant from repeating the offense; (3) deterring

stantial attention already, we must see how they would fit alongside or apart from what a *public* retributive justice theory entails for the implementation of punitive damages.⁴⁷ To that end, let me provide a summary of the basic structure of retributive damages that I proposed earlier.

B. *The Basic Structure of Retributive Damages: A Recap*

While this Section outlines the basic structure of retributive damages, it does not explain in detail the rationale underlying this structure or why this structure is desirable vis-à-vis other remedial or penal options. Those issues are both addressed and defended at length in *Retributive Damages*. As I demonstrated there, retributive justice theory offers not only a reason for reconfiguring punitive damages, but also a set of constraints.⁴⁸ After all, once properly understood, retributive justice is tethered to concerns for equality, modesty, accuracy, proportionality, impartiality, and the rule of law; such notions are largely missing not only from current common law punitive damages practices but also, to varying degrees, from the accounts of those scholars emphasizing punitive damages as vehicles for vindicating a private plaintiff's interest in "poetic justice"⁴⁹ or revenge⁵⁰ or a jury's interest in ventilating its outrage.⁵¹ In some respects, this public retributive interest means ensuring modest and fair sanctions *across the realm of simi-*

others from committing an offense; (4) preserving the peace; (5) inducing private law enforcement; (6) compensating victims for otherwise uncompensable losses; and (7) paying the plaintiff's attorneys' fees." Ellis, however, subjects these possible purposes to close scrutiny and finds some of them lacking. *See id.* at 76.

⁴⁷ While I adopt certain scholars' views on how to conceptualize and implement what I am calling aggravated and deterrence damages, I don't necessarily present the most comprehensive or sophisticated version of those particular approaches. This caveat seems necessary in light of the fact that there are disagreements within the cost internalization school and within the victim vindication camp over various details. In designing a pluralistic structure, I have simply tried to draw on the ideas that seemed most feasible and attractive for a post-*Philip Morris* world of extracompensatory damages.

⁴⁸ *See Markel, Retributive Damages, supra* note 1, at 304-09.

⁴⁹ *See Galanter and Luban, supra* note 3, *passim*.

⁵⁰ *See Colby, supra* note 5, at 433 ("[T]he desire for revenge burns bright in the human heart—too bright for the law to ignore."); Sebok, *supra* note 5, at 1031 ("[S]tate-sanctioned revenge is the best interpretation of [punitive damages].").

⁵¹ *See Wenger & Hoffman, supra* note 37, at 1138-40 (explaining that jurors often reject efficiency rationales for punitive damages and instead base their awards in part on their level of moral outrage).

larly situated defendants; in other respects it means ensuring safeguards to achieve accuracy, impartiality, and proportionality *in a particular case*.

In *Retributive Damages*, I made these claims based largely on the account of punishment that I call the *confrontational conception of retributivism* (CCR).⁵² The CCR seeks to communicate to defendants our seriousness about certain interests by applying some level of coercive condemnatory setback to the defendant's interests on account of her violating the state's law. In the retributive damages context, the statute describing the scope of retributive damages is the dictate of law. Hence, someone who violates that statute stands in a similar position, vis-à-vis the CCR, as someone who, for example, violates a typical criminal prohibition against theft or fraud. The offense warrants a coercive response by the state that adequately and parsimoniously communicates condemnation of that offense to the offender. Assuming that the offender is without further justification or excuse, she ought to be punished through retributive damages because doing so helps instantiate our commitments that we are moral agents capable of conforming our behavior to law and being held responsible; that, under the law, we all are entitled to enjoy equal liberty; and that we will defend our democratic sovereignty regarding that delineation of liberty against usurpations by offenders. By extending punishment

⁵² See Markel, *Retributive Damages*, *supra* note 1, pt. II. My prior works have addressed how this theory applies to other policy issues. See generally Dan Markel, *Against Mercy*, 88 MINN. L. REV. 1421 (2004) [hereinafter Markel, *Against Mercy*] (providing a retributivist critique of sites of unreviewable discretion for grants of mercy by executive figures); Dan Markel, *Are Shaming Punishments Beautifully Retributive? Retributivism and the Implications for the Alternative Sanctions Debate*, 54 VAND. L. REV. 2157 (2001) (considering alternative criminal sanctions through the lens of retributive justice theory); Markel, *State, Be Not Proud*, *supra* note 40, at 457-77 (arguing that retributive justice is incompatible with the death penalty); Dan Markel, *The Justice of Amnesty? Towards a Theory of Retributivism in Recovering States*, 49 U. TORONTO L.J. 389, 392 (1999) (arguing that the particularized amnesty utilized by some recovering states as part of a transitional justice program can be "compatible with justice, even when justice is understood as retributive in nature"). More recently, I have extended this theory to the Supreme Court's Eighth Amendment jurisprudence, see Dan Markel, *Executing Retributivism: Panetti and the Future of the Eighth Amendment*, 103 NW. U. L. REV. (forthcoming Spring 2009), available at <http://ssrn.com/abstract=1263683> (arguing that the Supreme Court's holding in *Panetti v. Quarterman*, 551 U.S. 930 (2007), is predicated on an understanding of retributive punishment as a communicative action directed at the offender and that this conception of punishment diminishes the constitutional justification for the death penalty). I also have extended the theory to the role that a defendant's family status should play in her criminal liability and punishment. See DAN MARKEL, JENNIFER M. COLLINS & ETHAN J. LEIB, *PRIVILEGE OR PUNISH: CRIMINAL JUSTICE AND THE CHALLENGE OF FAMILY TIES* (2009) (examining and often challenging the current use of family status in the criminal justice system).

against violators of this retributive damages statute, we continue to vindicate both the value of persons' rights and interests and our belief in the moral competence of persons to act freely within a zone created by those protected rights and interests.

One virtue of this account, when fully fleshed out, is its ability to explain both the internal intelligibility of retributive justice within a liberal democracy and the limits that may reasonably be placed on that social practice to help distinguish it from naked revenge. Significantly, this account explains the need for reducing Type I errors—in which people are mistakenly punished or excessively punished relative to comparable offenders—and Type II errors—in which offenders escape their punishment altogether or receive too lenient a punishment relative to comparable offenders. Accounts of both retributive justice and retributive damages ought to offer sustained reflection on the reasonable reduction of all of these kinds of error. By contrast, victim vindication and cost internalization lack the conceptual resources to do so effectively.⁵³

To realize these goals, I argued that under the retributive damages framework, when people defy certain legal obligations, the state may either seek to punish them through traditional criminal law or make available the sanction of retributive damages. Such damages would be credited against any further criminal sanctions imposed by the state for the same misconduct. Retributive damages statutes would empower victims—or in some cases, private attorneys general⁵⁴—to act on behalf of the state to seek the imposition of an “intermediate sanction.” These penalties are basically a stripped-down fine; they neither trigger the status of a “conviction” nor do they instigate any collateral consequences or future disabilities as a result of retributive damages liability.

Under this scheme, the amount of the penalty is determined largely by the reprehensibility of the defendant's misconduct. Specifically, the fine's amount is informed by two kinds of measurements.

⁵³ The victim vindication accounts say little about how to achieve consistency and predictability across cases. Furthermore, the dominant cost internalization accounts do not typically require inquiry into and judgment of the reprehensibility of the defendant's actions, so cost internalization proponents are not really interested in communicating condemnation to offenders. See, e.g., Galligan, *supra* note 5; Polinsky & Shavell, *supra* note 5.

⁵⁴ In *Retributive Damages*, I explained why and how nonvictims should have a role in facilitating the punishment of misconduct that involved harmless wrongdoing or wrongs that victims themselves did not seek to vindicate (fully). See Markel, *Retributive Damages*, *supra* note 1, at 279-86.

The first measurement is a number on a reprehensibility scale, while the second translates that reprehensibility score to an amount of damages. As a preliminary matter, the state legislature or sentencing commission would devise a set of guidelines for juries (or judges in bench trials) to help them objectively assess how reprehensible the misconduct is.⁵⁵ The guidelines would calibrate reprehensibility, perhaps on a scale of one to twenty, with twenty being the worst, using many of the factors that courts currently use to evaluate the defendant's reprehensibility.⁵⁶ Some factors, such as a defendant's history of past adjudicated misconduct, might increase reprehensibility, while other factors, such as preexisting compliance programs or remedial actions and restitution measures taken by the defendant upon discovery of the misconduct, might mitigate reprehensibility. In addition, the guidelines would include commentaries with hypothetical examples of misconduct that fell on various places on the scale. The state would assess a percentage of the defendant's financial condition (or net value for entities) that increases with the reprehensibility of the defendant's misconduct.⁵⁷ To use an example, a finding of two on the reprehensibility scale could lead to a retributive damages award of one percent of the defendant's net wealth, and a finding of twenty could lead to ten percent of the defendant's assets being assessed.

To ensure that the defendant does not benefit from the misconduct against the plaintiff, the total retributive damages award should also strip the defendant of gains, if any, in excess of compensatory damages that are owed to the plaintiff and that arose from the misconduct. These payments (the gains and the reprehensibility-based penalty) go to the state.⁵⁸ The defendant should also pay plaintiff's

⁵⁵ See generally Paul H. Robinson & Robert Kurzban, *Concordance and Conflict in Intuitions of Justice*, 91 MINN. L. REV. 1829 (2007) (providing research on widely shared intuitions regarding the scaling of reprehensibility); Sunstein et al., *supra* note 6, at 2077-78 (noting consistency in moral judgments but inconsistency in translating outrage into dollars). Legislatures and courts might also look to Michael Welner's research on the "Depravity Scale." See Michael Welner, *The Depravity Scale*, <http://www.depravityscale.org> (last visited Mar. 15, 2009).

⁵⁶ This scaling approach addresses some of the concerns raised by Cass Sunstein. See, e.g., Cass R. Sunstein, *On the Psychology of Punishment*, 11 SUP. CT. ECON. REV. 171, 179-81 (2004) (recognizing the value of guidelines and benchmarks for improved cognition and fairness in punitive damages awards across cases).

⁵⁷ Cf. Markel, *Retributive Damages*, *supra* note 1, at 290-96 (offering rationales for "scaling fines to the defendant's financial position").

⁵⁸ The gain-stripping aspect of the retributive damages structure makes this approach broadly consistent with the "complete deterrence" approach advocated by economists like Keith Hylton. See Hylton, *supra* note 18, at 464-67 (stressing that an

lawyers' fees (for the marginal labor necessary to prove the defendant's reprehensibility) and a modest and fixed award to the plaintiff—I suggested something in the range of \$10,000—for bringing the case to the public's attention.⁵⁹ These payments (to the state, the plaintiff, and the lawyer) together constitute what I take to be the most sensible, though not the only, way to structure extracompensatory damages designed to advance the goals of retributive justice.⁶⁰ Of course, the plaintiff could also receive an amount based on compensation for aggravated injuries to the person's dignity if compensatory damages in that jurisdiction did not already account for that injury.

Consistent with the notion that retributive damages are supposed to serve as an *intermediate* sanction on the public's behalf, legislatures should authorize courts to order defendants to pay the damages amount as a percentage of profits in coming years in situations where a defendant has reason to doubt its viability if required to pay one lump sum. However, if one is concerned that a defendant committed grave misconduct and then restructured its finances to make it appear that it could not pay the amount owed, the courts might adjust the retributive damages based on the financial condition of the defendant at the time the misconduct (last) occurred.⁶¹

The scheme described above furnishes potential defendants with little basis for complaining that the amount or award of retributive

optimal-penalty system would eliminate the prospect of gain by the offender); *see also* David D. Haddock et al., *An Ordinary Economic Rationale for Extraordinary Legal Sanctions*, 78 CAL. L. REV. 1, 20 (1990) ("Enforcing property rules requires stripping all gain (or more) from a taking."). The retributive damages penalty also includes a wealth- and reprehensibility-informed monetary penalty that puts the defendant in a worse position than she was at the status quo ante. Complete-deterrence models permit but do not require that setback, which is part of how the retributive message of condemnation is communicated. *See* Markel, *Retributive Damages*, *supra* note 1, at 242-43 (contrasting the messages of complete deterrence and retribution).

⁵⁹ Such a flat fee avoids the lottery effects that a plaintiff would enjoy from having the good "fortune" of having a wealthy injurer.

⁶⁰ Other valuation methodologies might also be consistent with retributive justice values. *See* Geistfeld, *supra* note 5, at 286-92, 306 (proposing for torts involving fatal risks a damages valuation that examines government data regarding the monetization of fatal risks); Markel, *Retributive Damages*, *supra* note 1, at 287 n.166, 290 n.181 (explaining why a multiplier of compensatory damages for torts involving purely financial losses might also comply with retributive justice values). Despite my open-mindedness toward these alternative methods of assessing retributive damages, I should clarify that if they were to be used, the amounts imposed would also need to satisfy the retributive goals of stripping the gain and imposing an adequate, proportionate, and parsimonious setback on the defendant.

⁶¹ Indeed, depending on the circumstances, the restructuring to evade payment could arguably be a factor used to raise one's reprehensibility score.

damages is a surprise, since the standards that would be applied to them are no different than the guidelines that have now become familiar in many jurisdictions when assessing criminal liability and sentencing.⁶² Of course, the defendants in criminal cases have more procedural safeguards in place, and thus, if we are deputizing plaintiffs to facilitate the imposition of an intermediate sanction, then we should enhance at least some of the procedural protections available in retributive damages cases—an aspect of the argument I develop in greater detail in Part III. But first, let us turn to some questions about structuring a pluralistic approach to extracompensatory damages.

II. PLURALISM ABOUT PURPOSES

Thus far I have discussed three purposes to guide punitive damages: retributive justice, cost internalization, and victim vindication. As mentioned at the outset, unlike those who endorse a dominant role for one purpose or another,⁶³ I submit that we should be pluralistic about the purposes of punitive damages.⁶⁴ Different cases present

⁶² See generally Richard S. Frase, *State Sentencing Guidelines: Diversity, Consensus, and Unresolved Policy Issues*, 105 COLUM. L. REV. 1190, 1194-1208 (2005) (summarizing states' use of sentencing guidelines).

⁶³ See, e.g., *supra* note 3.

⁶⁴ I am obviously not the first person to note that punitive damages serve different purposes; my hope is that the discussion here advances the ball by explaining in greater detail what a disaggregated scheme of extracompensatory damages would look like if implemented with care. Others who have contemplated disaggregation and pluralism include Professors Galligan, Sharkey, Calabresi, Polinsky & Shavell, Rustad, and Salbu. See, e.g., *Ciraolo v. City of N.Y.*, 216 F.3d 236, 244-46 (2d Cir. 2000) (Calabresi, J., concurring) ("Indeed, it would not be inappropriate to disaggregate the retributive and deterrent functions of extracompensatory damages altogether and allow separate awards to further the two separate goals."); Thomas C. Galligan, Jr., *Disaggregating More-Than-Whole Damages in Personal Injury Law: Deterrence and Punishment*, 71 TENN. L. REV. 117 (2003); Polinsky & Shavell, *supra* note 5; Michael L. Rustad, *Happy No More: Federalism Derailed by the Court that Would Be King of Punitive Damages*, 64 MD. L. REV. 461, 468-93 (2005) (adumbrating the plural goals that punitive damages serve); Steven R. Salbu, *Developing Rational Punitive Damages Policies: Beyond the Constitution*, 49 FLA. L. REV. 247 (1997) (recognizing different rationales for punitive damages and noting that those rationales can be used separately or jointly to calculate damages); Sharkey, *supra* note 5, at 363 ("Notwithstanding the fact that the retributive-based and deterrence-based components of punitive damages are not fully separable, and indeed have potentially synergistic or overlapping effects, there are significant gains to be achieved from treating them as conceptually distinct."). In various respects, my account builds on and departs from these earlier efforts, most significantly in terms of distinguishing between the public interest in retributive justice and the private interest in victim vindication, and also in explaining how cost internalization intersects with these goals, especially after *Philip Morris*.

different problems; not every case requires pursuit of any of these purposes. Pluralism means relabeling punitive damages because the name “punitive” is somewhat misleading in certain contexts. I suggest calling them “extracompensatory damages” and use that umbrella term to encompass what I referred to earlier as retributive, deterrence, and aggravated damages.

Because these three purposes are distinct, jurisdictions need to weigh their attractiveness and develop structures capable of realizing those goals without overlap or confusion. One way to accomplish these goals is to devise special jury forms that ensure that the jury provides a written explanation for what it is awarding and why.⁶⁵ Courts, with legislative input, could provide the jury with instructions modeled on those appearing in this Article’s Appendix, which give clear guidance regarding these various purposes and how to arrive both at appropriate figures for aggravated and deterrence damages and at an appropriate reprehensibility finding for retributive damages. Several other issues of retributive damages require attention, specifically as they interact with deterrence and aggravated damages.

A. *Cost Internalization and Deterrence Damages*

1. Overview

The idea of disentangling the purposes of extracompensatory damages received its most prominent attention in a concurring opinion by Judge Guido Calabresi in *Ciraolo v. City of New York*.⁶⁶ In that

⁶⁵ See, e.g., Galanter & Luban, *supra* note 3, at 1439-40 (urging jurisdictions to require jury explanations). The jury-explanation device undermines the confidentiality of jury deliberations, but it is hard to understand why that confidentiality should be a higher priority than the achievement of the public’s interest in the fair and accurate imposition of justice. The concern of judicial interference with the jury’s role is especially exaggerated given that judges perform various gatekeeping roles.

⁶⁶ 216 F.3d 236, 244-46 (2d Cir. 2000) (Calabresi, J., concurring). Calabresi drew on the work of other scholars in endorsing cost internalization and urging the separation of cost internalization from punishment. See, e.g., Polinsky & Shavell, *supra* note 5, at 906 (“[T]he imposition of damages equal to harm, appropriately multiplied to reflect the probability of escaping liability, achieves proper deterrence.”); see also Galligan, *supra* note 5. Galligan’s account is very instructive regarding the deterrence function, but it does not work through the details of implementing the public’s interest in retributive justice. It also fails to separate the public interest in retributive justice from the victim’s private interest in vindicating the injury to her dignity. Nonetheless, the points that Galligan makes about the proper way to determine optimal deterrence are useful in a context where “total cost internalization” damages would be permissible. See *id.* at 128-34 (proposing for the calculation of punitive damages the use of average

case, the Second Circuit had to decide whether punitive damages were permitted against municipalities engaged in unconstitutional misconduct—specifically, unwarranted strip searches of arrestees in New York City. Bound by Supreme Court precedent,⁶⁷ Judge Calabresi wrote that the award of punitive damages in the court below required reversal. Concurring with his own opinion, however, Judge Calabresi wrote separately to explain why extracompensatory damages should be available to encourage cost internalization—the idea that a defendant should pay for the social harms generated by its activities, regardless of whether all victims decide to bring individual suits.⁶⁸ Under standard accounts, cost internalization is desirable because it will spur optimal amounts of precaution by the defendant, leading to

potential harm—rather than the plaintiff's actual harm—subject to a multiplier so that the plaintiff more accurately serves as a proxy for all who are damaged but do not sue). However, because of the way that *Philip Morris* makes “total cost internalization” constitutionally problematic, *see infra* subsection II.A.2, I do not think that there is a constitutionally permissible basis for structuring deterrence damages in the way that Professor Galligan suggests.

⁶⁷ *See City of Newport v. Fact Concerts, Inc.*, 453 U.S. 247, 259-66 (1981) (affirming the common law absolute immunity of municipalities against punitive damages actions for the bad-faith misdeeds of their officials).

⁶⁸ Judge Calabresi wrote that

[c]osts may not be sufficiently reflected in compensatory damages for several reasons, most of which go to the fact that not all injured parties are in fact compensated by the responsible injurer. For example, a victim may not realize that she has been harmed by a particular actor's conduct, or may not be able to identify the person or entity who has injured her. Where the injurer makes active efforts to conceal the harm, this problem is of course exacerbated. Moreover, even if a victim is aware of her injury and is able to identify its cause, she may not bring suit. A person will be unlikely to sue if the costs of doing so—including the time, effort, and stress associated with bringing a lawsuit—outweigh the compensation she can expect to receive. A victim is especially unlikely to sue, therefore, in cases where the probable compensatory damages are relatively low. As a result, a harm that affects many people, but each only to a limited degree, will generally be given inadequate weight if only compensatory damages are assessed.

In addition, some victims will not sue even if the damages they could expect to receive would exceed the costs of suing. Victims will differ greatly in their knowledge of and access to the legal process, and those who are relatively poor and unsophisticated, as a practical matter, are frequently unable to bring suit to redress their injuries even if those injuries are grave. A harm that disproportionately affects such victims, therefore, is also particularly likely not to be accurately reflected in compensatory damages.

Ciraolo, 216 F.3d at 243-44 (Calabresi, J., concurring) (citations omitted).

appropriate prices and levels of activity while ensuring that defendants do not thrust the costs of their tortious activities onto others.⁶⁹

As a matter of policy prescription, there is much to recommend the cost internalization approach. Broadly speaking, defendants *should* have to pay for the mess that they make; if they can exploit enforcement gaps by private and public parties, there will be an incentive to take insufficient care, which will also arouse the risk of underdeterrence.⁷⁰

But because the cost internalization approach is nonpunitive in nature, we should refer to damages pursuing cost internalization as simply “deterrence” damages, not “punitive” or “retributive” damages.⁷¹ Indeed, as mentioned earlier, this approach does not require inquiry into the mens rea or reprehensibility of the defendant’s conduct to justify the augmentation of damages.⁷² All that matters is that

⁶⁹ According to their supporters, optimal deterrence damages

should be set at a level such that the expected damages of defendants equal the harm they have caused, for then their damage payments will, in an average sense, equal the harm. This implies a simple formula for calculating punitive damages, according to which harm is multiplied by a factor reflecting the likelihood of escaping liability.

Polinsky & Shavell, *supra* note 5, at 954. Some law and economics scholars have registered skepticism with the multiplier principle. See, e.g., Richard Craswell, *Deterrence and Damages: The Multiplier Principle and Its Alternatives*, 97 MICH. L. REV. 2185, 2191-98 (1999) (contending that the use of a static multiplier misapprehends the variable probabilities of punishment, which generally correlate with the seriousness of a crime); see also Sharkey, *supra* note 5, at 368-69 (raising concern that the Polinsky-Shavell approach doesn’t achieve adequate cost internalization). I should note that my proposal largely brackets the skepticism towards the deterrence damages multiplier, but if economists converge on a different method of achieving cost internalization in a way that is compatible with *Philip Morris*, I am certainly open to it.

⁷⁰ See Thomas C. Galligan, Jr., *The Risks of and Reactions to Underdeterrence in Torts*, 70 MO. L. REV. 691, 692 (2005) (arguing that rules developed for individual actions do not deter effectively in the mass tort context).

⁷¹ Again, I note that my use of “deterrence” damages may seem idiosyncratic to those who think of deterrence in terms of complete deterrence. Here I simply mean to describe damages meant to realize cost internalization to the extent that I view permissible in a post-*Philip Morris* world.

⁷² See Polinsky & Shavell, *supra* note 5, at 905-10 (arguing that deterrence damages should be calculated as the amount of the plaintiff’s harm multiplied by the reciprocal of the probability that the defendant will escape liability, and maintaining that the reprehensibility of the defendant’s conduct is irrelevant for deterrence damage purposes); Kenneth W. Simons, *Deontology, Negligence, Tort, and Crime*, 76 B.U. L. REV. 273, 273 (1996) (explaining that, under an optimal deterrence regime, one is “entitled to harm the victim so long as he pays for the harm (with the expectation that this entitlement will induce him to take optimal care)”). But see Sharkey, *supra* note 5, at 405 (requiring a threshold finding of at least recklessness before such nonpunitive extracompensatory

the defendant caused tortious harm that was not otherwise internalized and that such harm is not excused or justified on other grounds.⁷³

2. Deterrence Damages After *Philip Morris*

On account of the Supreme Court's recent decision in *Philip Morris*, proponents of the cost internalization approach now face substantial difficulties. Implicitly repudiating the language of its *TXO* decision,⁷⁴ which permitted punitive damages awards to consider the harm or potential harm from the conduct as a whole,⁷⁵ the Supreme Court reversed course in *Philip Morris*. Writing for a five-to-four majority, Justice Breyer held that the Due Process Clause forbids punishing a defendant for harms to nonparties to the instant litigation because the defendant would not have the ability to "defend against the charge," thus depriving the defendant of notice and imposing a substantial de-

damages can be imposed). For reasons explained elsewhere, Sharkey's threshold requirement of recklessness stands in tension with a better understanding of cost internalization. See Galligan, *supra* note 5, at 62-63 (arguing that augmented awards are intended to deter and not to punish—therefore making focus on the defendant's state of mind irrelevant to calculating these awards); Kelly, *supra* note 32, at 1440 n.36 (explaining that Sharkey's proposal of limiting "societal" damages to instances where defendants engage in conduct that is at least reckless "seems to move the justification back toward retribution, letting state of mind, not undercompensation, differentiate the cases where additional deterrence is needed"); Markel, *Retributive Damages*, *supra* note 1, at 243 & n.9, 244 (explaining that the issue of reprehensibility is separate from the issue of whether a defendant has been forced to bear the costs of her actions).

⁷³ See Polinsky & Shavell, *supra* note 5, at 905-10. To be sure, one might still have residual anxieties about whether deterrence damages would perform an adequate compensatory and norm-projection function—especially in cases where the payment of deterrence damages will not readily be commensurable with the harm suffered. In other words, there is a legitimate concern that "compensation never compensates" in the tort context because the plaintiff never wanted to "earn" that compensation. See JOSEPH WILLIAM SINGER, ENTITLEMENTS: THE PARADOX OF PROPERTY 282 (2000). But that concern indicates only why deterrence damages should not be the exclusive purpose of extracompensatory damages.

⁷⁴ *TXO Prod. Corp. v. Alliance Res. Corp.*, 509 U.S. 443 (1993). For a discussion of the Supreme Court's infidelity to precedent in this context, see Posting of Dan Markel to PrawfsBlawg, *Philip Morris: Up in Smoke?*, http://prawfsblawg.blogs.com/prawfsblawg/2007/02/philip_morris_u.html (Feb. 20, 2007).

⁷⁵ See *TXO Prod. Corp.*, 509 U.S. at 460-62; see also *BMW of N. Am., Inc. v. Gore*, 517 U.S. 559, 582 n.35 (1996) (indicating, implicitly, that punitive damages awards might be permitted to reflect the harm caused to all in-state purchasers of BMWs and not just the harm caused to the individual plaintiff bringing suit); *Kirkland v. Midland Mortgage Co.*, 243 F.3d 1277, 1280 (11th Cir. 2001) ("[P]unitive damages . . . [are] measured to reflect, not the wrong done to a single individual, but the wrongfulness of the conduct as a whole.").

gree of arbitrariness and uncertainty in punishment.⁷⁶ Therefore, in assessing the putative excessiveness of punitive damages in a given case, a court must scrutinize whether a jury's punitive damages award included amounts based on harms suffered by strangers to the litigation.

By restricting the permissible scope of harm for punitive damages, the *Philip Morris* decision now raises questions about whether "total cost internalization" is forbidden.⁷⁷ The ambiguity is subtle and has been introduced in Professor Colby's recent article on the subject.⁷⁸ Colby emphasizes that because damages designed for cost internalization (optimal deterrence) need no finding of reprehensibility to warrant their imposition, they should not logically be thought of as part of the "punitive" damages constitutional analysis, regardless of their scope.⁷⁹ Thus states could, on his view, constitutionally pursue damages meant only to achieve total cost internalization, so long as the state said there was nothing punitive to this cost internalization approach.⁸⁰ Hence "deterrence" damages would essentially serve as a regulatory tool to facilitate total cost internalization, whereas "puni-

⁷⁶ *Philip Morris USA v. Williams*, 549 U.S. 346, 353-54 (2007). The jury had awarded the decedent's wife \$21,000 in economic compensatory damages, \$800,000 in noneconomic compensatory damages, and \$79.5 million in punitive damages. *Id.* at 350. Recently, on remand from the U.S. Supreme Court, the Supreme Court of Oregon upheld the jury verdict, claiming that there was an adequate and independent state ground for the decision. See *Williams v. Philip Morris Inc.*, 176 P.3d 1255, 1260-61 (Or. 2008). Subsequently, the United States Supreme Court agreed, once again, to rehear an appeal brought by Philip Morris—but only on the issue of whether the Oregon Supreme Court failed to abide by the instructions of the remand, not whether the punitive damages in the case were constitutionally excessive. *Philip Morris USA Inc. v. Williams*, 128 S. Ct. 2904 (2008). The Supreme Court, however, recently announced that its grant of certiorari on this issue was improvidently granted, leaving the decision by the Oregon courts to stand. *Philip Morris USA Inc. v. Williams*, No. 07-1216, slip op. (U.S. Mar. 31, 2009).

⁷⁷ By "total cost internalization," I mean the full scope of harm caused by a defendant's wrongdoing, and not just the harm caused to the plaintiff in the instant litigation.

⁷⁸ See Colby, *supra* note 5, *passim*.

⁷⁹ *Id.* at 467-79. Because cost internalization refers to a different kind of deterrence (optimal) than the one probably intended under the Court's pronouncements (complete deterrence), Colby's conclusion is plausible as a matter of theory and logic. But because I think that this conclusion is realistically at odds with the gravamen of the *Philip Morris* decision, I doubt that the *Philip Morris* Court meant that one could pursue statewide cost internalization with little to no constitutional oversight, for reasons I explain shortly.

⁸⁰ See *id.* at 476 ("Williams does not stand in the way of implementing an extracompensatory remedy that seeks optimal deterrence."). My sense is that Professor Sharkey's social-damages proposal would similarly seek to restrain the reach of *Philip Morris* so that it would not apply to nonpunitive extracompensatory damages.

tive” damages, on his analysis, would serve to advance the private interest in victim vindication, not the public interest in retributive justice.⁸¹

My view is that Colby’s argument about the future of total cost internalization is wrong because of its cramped interpretation of *Philip Morris*. His interpretation reads the case so narrowly as to render its holding an effective nullity, giving with one hand a right to a jury instruction that defendants have desperately sought and then stripping away the value of that right with the other hand.⁸² It is hard to believe that the Court and various litigants would be divided so sharply on an issue with such apparent lack of material consequence. And unlike the notorious decision in *United States v. Booker*,⁸³ which recognized an important right for the defendant only to demolish its significance, there are no separate opinions in *Philip Morris* regarding the “merits” and the “remedy.”⁸⁴ As a single opinion, *Philip Morris* is therefore less susceptible to plausible charges of schizophrenic reasoning.

To my mind, the better reading of *Philip Morris* finds that cost internalization’s future is impeded but not destroyed. Cost internalization remains feasible to the extent that jurisdictions make available class actions or other aggregative-litigation strategies that protect the rights of defendants; once a class is certified, the people who were previously nonparties become parties to the litigation. Additionally, the prospect of deterrence damages outside of the class action survives *Philip Morris* to the extent that the defendant may have escaped having to compensate the instant plaintiff(s). So whereas cost internalization prior to *Philip Morris* should plausibly have included all the harms caused by the defendant (and therefore something closer to what Judge Calabresi and Professor Sharkey advocated), it now should be-

⁸¹ Professor Colby posits that a public interest in retributive justice cannot constitutionally be pursued outside the criminal law because of the absence of constitutional criminal procedural safeguards in civil suits. *Id.* at 440-57. For reasons I will explain in Part III, I find this view to be mistaken.

⁸² Unsurprisingly, other scholars supportive of cost internalization damages have also suggested that the Supreme Court’s Due Process jurisprudence may not apply in its entirety to nonpunitive damages designed simply to compensate society for various harms to nonparties within state lines. *See, e.g.,* Sharkey, *supra* note 5, at 428-33 (rejecting the argument that due process extraterritoriality and multiple-punishment concerns would prevent the implementation of a societal-damages scheme); *Ciraolo v. City of N.Y.*, 216 F.3d 236, 246 n.8 (2d Cir. 2000) (Calabresi, J., concurring) (suggesting by implication that, in contrast to truly punitive damages, socially compensatory damages do not require additional procedural protections).

⁸³ 543 U.S. 220 (2005).

⁸⁴ *See* Michael W. McConnell, *The Booker Mess*, 83 DENV. U. L. REV. 665 (2006) (discussing the doctrinal chaos embedded in *Booker*’s two separate opinions).

come, I think, a more focused inquiry. In other words, the operative questions are what harm did the defendant cause this case's plaintiff(s), and what is the likelihood that the defendant would escape having to pay for that harm to this case's plaintiff(s)? Thus, using the Polinsky and Shavell multiplier method, if a jury believed a plaintiff had only a one-third chance of discovering that it was this defendant who caused the harm to the plaintiff, then that should permit the jury to impose deterrence damages that are double the compensatory damages. On this view, *Philip Morris* permits a state to apply a multiplier, for example, but it must be based solely on the likelihood that the harm to the plaintiff would not be compensated by the defendant.⁸⁵

Regardless of whose reading of *Philip Morris* is correct, there can be little question that the Court would almost certainly still impose constitutional limits on deterrence damages to ensure that the state's regulatory mechanism was not being used to regulate extraterritorial conduct that might have been lawful in other jurisdictions.⁸⁶ Needless to say, jury instructions and a jury verdict form can reflect these various considerations. The Appendix to this Article makes an effort to craft such jury instructions based on my reading of *Philip Morris*'s implications for cost internalization.⁸⁷

3. Standard of Review for Deterrence Damages

Isolating the cost internalization function of extracompensatory damages suggests that the standard of review on appeal might require revision. In *Cooper Industries, Inc. v. Leatherman Tool Group, Inc.*, the Supreme Court announced that in federal cases the amount of punitive damages should be reviewed de novo—a decision partially predicated on the claim that determinations of punitive damages involve a morally evaluative component that has to be weighed against a defen-

⁸⁵ Cf. *BMW of N. Am., Inc. v. Gore*, 517 U.S. 559, 592-94 (1996) (Breyer, J., concurring) (discussing the use of a multiplier as a potentially constitutionally plausible limit on punitive damages).

⁸⁶ See, e.g., *State Farm Mut. Auto. Ins. Co. v. Campbell*, 538 U.S. 408, 422 (2003) ("A basic principle of federalism is that each State may make its own reasoned judgment about what conduct is permitted or proscribed within its borders, and each State alone can determine what measure of punishment, if any, to impose on a defendant who acts within its jurisdiction.").

⁸⁷ I reiterate that I don't have a dog in this particular fight; if it turns out that Professor Colby's reading of cost internalization after *Philip Morris* is correct, a view which seems similar to Professor Sharkey's view articulated pre-*Philip Morris*, then I would be happy to have the jury instructions and potentially the procedural safeguards dealing with cost internalization adjusted accordingly.

dant's due process rights.⁸⁸ That holding seems applicable only with respect to the condemnatory part of extracompensatory damages. By their nature, deterrence damages will involve matters of empirical estimation (i.e., the likelihood that the defendant would escape having to compensate this plaintiff) and thus are no different than compensatory damages. If so, then determinations of deterrence damages deserve greater deference when reviewed on appeal.⁸⁹

4. Who Receives Deterrence Damages

We must also consider to whom deterrence damages should be paid. Prior to *Philip Morris*, the calculation of deterrence damages for cost internalization might require looking at all the tortious harm that the defendant caused and would thus counsel that deterrence damages be paid to a specific (state) fund to compensate harmed nonlitigants.⁹⁰ That scheme would keep funds available to future claimants who were harmed by the defendant's misconduct. If I read *Philip Morris* correctly on the issue of cost internalization, however, the need for a public fund to distribute deterrence damages will diminish correspondingly. Since the post-*Philip Morris* inquiry for cost internalization is likely to be restricted to a specific case's plaintiff(s), the plaintiff(s) have a stronger claim to all deterrence damages than they did prior to *Philip Morris*, when harms to other nonparties could be considered by the jury in determining the amount of punitive damages.

⁸⁸ *Cooper Indus., Inc. v. Leatherman Tool Group, Inc.*, 532 U.S. 424, 432-36 (2001).

⁸⁹ Cf. *Gore*, 517 U.S. at 592-94 (Breyer, J., concurring) (alluding to the possibility of "more deferential review" of damages designed to perform a "constraining" role of cost internalization); Sharkey, *supra* note 5, at 446 (suggesting deferential review for compensatory societal damages but stringent de novo review for "'morally'-based retributive 'punitive' damages"). I agree with Sharkey's endorsement of deferential review for cost internalization, but I think that it is important to further decouple review of victim vindication from review of the public retributive function. I say more about that *infra* subsection III.B.2. Also, since Sharkey's proposal is for "total cost internalization," whereas mine permits only what I might call *Philip Morris*-compatible cost internalization, I think that deferring to the jury for these damages is a bit easier to justify because there are fewer challenges to jury competence when juries are limited to the inquiry of the likelihood that the defendant would have evaded compensation to the plaintiff(s)/victim(s) only. See *id.* at 447-50 (raising and deftly addressing objections to juries' competence to calculate compensatory societal damages).

⁹⁰ See Galligan, *supra* note 5, at 140-41 (discussing the potential use of placing "societal compensatory damages" in a public fund); see also Sharkey, *supra* note 5, at 402 (suggesting the possibility of using "augmented damages" to establish a fund, for the benefit of nonparties to the case, "to offset the type of harm at issue in the case"); *id.* at 392 ("Societal damages, as envisioned in this Article, would redress the harms inflicted by the defendant *upon parties not before the court.*" (emphasis added)).

In short, plaintiffs (who are victims) should receive deterrence damages.⁹¹

5. Reconciling Deterrence Damages with Retributive Damages

If a state adopted a pluralistic structure for extracompensatory damages, it would have to determine whether deterrence damages should offset retributive damages or vice versa. Professors Polinsky and Shavell argue that the amount of deterrence damages awarded should *offset* the amount of damages awarded for the purpose of retribution because, from their perspective, deterrence damages also serve to punish the defendant to some extent.⁹² This conclusion seems mistaken. Deterrence damages are a purely nonstigmatic, “cool” cost internalization device.⁹³ They do not “punish” a defendant any more than state incorporation fees “punish” a defendant.

If defendant (*D*) causes harm to victim (*V*) without justification, a *prima facie* reason for *V*’s having an avenue of recourse for compensation against *D* exists, but it is not a reason, in the absence of some culpable form of *mens rea*, to condemn *D* in the language of retribution. So if cost internalization is the goal of deterrence damages, and there is no necessary threshold of a finding of a culpable *mens rea*, then accordingly, no offset for deterrence damages should be available against one’s retributive damages tab.⁹⁴

⁹¹ Theoretically, cost internalization advocates might encourage a cost internalization strategy that avoids conferring upon a plaintiff a monopoly to seek recourse or settle. In other words, where underdeterrence is a concern because of victims who fail to bring sufficient cases, one might want to permit a private attorney general (or a public agency) to ensure that the defendant fully internalizes the costs of tortious conduct that a victim decides not to pursue (to the full extent). The same desire to decouple deterrence from compensation might be true where one is concerned about overcompensation to plaintiffs or insufficient precautions taken by victim-plaintiffs. For examples of arguments in favor of decoupling compensatory damages awarded to plaintiffs from deterrence damages paid by defendants, see A. Mitchell Polinsky & Yeon-Koo Che, *Decoupling Liability: Optimal Incentives for Care and Litigation*, 22 RAND J. ECON. 562 (1991), and David Rosenberg, *Decoupling Deterrence and Compensation Functions in Mass Tort Class Actions for Future Loss*, 88 VA. L. REV. 1871, 1873-74 (2002).

⁹² See Polinsky & Shavell, *supra* note 5, app. (detailing model jury instructions that contend that the amount of cost internalization damages should offset the amount of damages for the purposes of retribution).

⁹³ See *id.* at 906 (“That a defendant’s conduct can be described as reprehensible is in itself irrelevant. Rather, the focus in determining punitive damages should be on the injurer’s chance of escaping liability.”).

⁹⁴ Cf. *Ciraolo v. City of N.Y.*, 216 F.3d 236, 246 (2d Cir. 2000) (Calabresi, J., concurring) (“But a separate award of punitive damages would be allowed only in cases where the defendant’s conduct was sufficiently reprehensible to deserve punishment

Conversely, a defendant should not be able to credit the amount of retributive damages paid against a deterrence damages bill. Even after a defendant pays her deterrence damages, she may in fact be in a better position than she was before the tort was committed. In other words, the tortious activity might still be profitable if only the amount necessary for cost internalization has to be paid. But if retributive damages are also warranted because the defendant's activity was reprehensible, then that residual gain needs to be stripped under the retributive approach and the defendant must endure an additional setback. Thus there is no sense in trying to credit retributive damages against deterrence damages. It is as if a defendant is saying, "I would like to credit my fines against my tax bill." Allowing a defendant to reduce its deterrence damages award through a retributive damages award confuses the defendant (and society) about the nature of the reason for the extracompensatory damages.⁹⁵

In sum, the prospect of deterrence damages should in principle be considered in any civil tort action. One can imagine situations where either deterrence or retributive damages, or both, would be appropriate in light of the different purposes served by them and the different facts of the cases. A jury determining extracompensatory damages should be instructed to separate its decisions about which damages are necessary.⁹⁶ A special verdict form reflecting the various differences could achieve this goal in much the way that this Article's Appendix suggests. Such a separation of categories of damages on a jury verdict form would also facilitate proper appellate review of the extracompensatory damages. Of course, such deterrence damages should be subject to the type of constraints articulated by various scholars of cost internalization,⁹⁷ and, if applicable, the federalism concerns of the Supreme Court.

apart from whatever assessment was required to compensate the individual victim or society as a whole." (emphasis omitted)).

⁹⁵ By respecting the distinctive values associated with retributive justice and cost internalization, the proposal here departs from those of both Galanter & Luban and Polinsky & Shavell. See Galanter & Luban, *supra* note 3, at 1451 (arguing that decoupling punitive awards and moral judgments "simply misses the point"); Polinsky & Shavell, *supra* note 5, app. (presenting model jury instructions that advocate the subtraction of cost internalization damages from the amount of damages meant to achieve retributive punishment).

⁹⁶ See *infra* Section II.B for further discussion on the intersection of deterrence damages and aggravated damages.

⁹⁷ Economists disagree about the precise architecture for punitive damages. Compare Robert D. Cooter, *Economic Analysis of Punitive Damages*, 56 S. CAL. L. REV. 79 (1982) (arguing that punitive damages should be used only in cases of gross, inten-

B. *Aggravated Damages for Victim Vindication*

1. Overview

Aside from the public's interests in cost internalization and retributive justice,⁹⁸ a third purpose of extracompensatory damages warrants special attention. This third form of extracompensatory damages provides a remedy through which plaintiffs can seek recourse against defendants for the special dignitary harms caused by the defendants' misconduct. I will call awards for these kinds of harms "aggravated damages."⁹⁹

Interestingly, and contrary to the Supreme Court's historical account in *Cooper Industries*, Professor Sebok argues that

punitive damages have never served the compensatory function attributed to them by the [Supreme] Court in *Cooper*. . . . If punitive damages served a compensatory function, it would have been for a *category of injury that is still not considered compensable by contemporary tort law, namely the injury of insult that wounds or dishonors*.¹⁰⁰

If Professor Sebok is correct, then not all noneconomic awards reflected in compensatory damages today encompass the special nature of personal dignity harms. This would suggest a reason for addressing these special injuries through aggravated damages meant to achieve what I am calling "victim vindication."

The victim vindication rationale has a number of supporters.¹⁰¹ For example, Professor Sebok interprets the availability of punitive damages as a way for victims to pursue "state-sanctioned revenge" by creating an avenue of recourse against the defendant.¹⁰² Recently,

tional fault and should be determined based on the costs of compliance or benefits of noncompliance that motivated the violation), *with* Polinsky and Shavell, *supra* note 5 (advocating the multiplier approach).

⁹⁸ See Markel, *Retributive Damages*, *supra* note 1, at 257-66.

⁹⁹ My inspiration for this term came from two commonwealth scholars. See Chapman & Trebilcock, *supra* note 5, at 746-47 (discussing an English case, *Rookes v. Barnard*, [1964] All E.R. 367, 407 (Eng.), that distinguished aggravated from punitive damages). I am not claiming that my proposed use of the term "aggravated damages" captures historical, extant, or emerging practices in Canadian and English cases. However, for illuminating descriptions of aggravated damages in comparative perspective, see Anthony J. Sebok & Vanessa Wilcox, *Aggravated Damages*, in PUNITIVE DAMAGES (Helmut Koziol & Vanessa Wilcox eds., forthcoming 2009).

¹⁰⁰ Sebok, *supra* note 15, at 204-05 (emphasis added).

¹⁰¹ See *supra* subsection I.A.2.

¹⁰² Sebok, *supra* note 5, at 1023-29.

Professor Colby has embraced this rationale as well.¹⁰³ These victim vindication models empower plaintiffs to vindicate particular harms to their dignity, but they also give plaintiffs the power to waive such recourse, satisfying themselves with “mere” compensatory damages or no damages at all.

There are merits to these victim vindication accounts as interpretive reconstructions of the constitutional architecture of tort law and particularly of punitive damages, but the reconstructions are not flawless.¹⁰⁴ That said, the bigger problem is that the victim vindication accounts fail to explain from a normative perspective why they represent the best structure, all things considered. For one thing, if embraced, these accounts would fail to meaningfully realize the *public’s* interest in retributive justice even if they plausibly serve to vindicate the autonomy or dignity of the particular victim by conferring upon her some control over her legal rights. Victims who, because of mercy or frustration with the legal process, forego recourse are creating Type II errors about which retributivists and the public should be concerned. The same can be said of settlements that satisfy plaintiffs but leave grave misconduct concealed and unpunished by some form of public rebuke.

Of course, merely because a victim vindication account fails to realize retributive justice does not mean that the goal of empowering victims cannot be accommodated within a pluralistic account of extra-compensatory damages. One potential benefit of a victim vindication model that goes beyond facilitating a right to state-sanctioned revenge is that it serves as a way to achieve compensation for plaintiffs (or their lawyers) who bring suits against the perpetrators of antisocial conduct for which compensatory damages alone would not suffice.¹⁰⁵ That

¹⁰³ See Colby, *supra* note 5, at 423-40 (arguing that punitive damages are best understood as a form of punishment for “private wrongs”).

¹⁰⁴ For example, as a descriptive matter, Sebok’s account acknowledges the difficulty of incorporating the *State Farm* Court’s presumptive single-digit ratio into his theory. See Sebok, *supra* note 5, at 1029-36 (criticizing Justice Kennedy’s opinion that “[s]ingle-digit multipliers are more likely to comport with due process,” *State Farm Mut. Auto. Ins. Co. v. Campbell*, 538 U.S. 408, 425 (2003), as lacking a “principled foundation,” as justified merely by a desire for simplicity, and as “impossible to square with a post-*Lochner* theory of due process”).

¹⁰⁵ Of course, to the extent that victim vindication models do supplement insufficient compensatory damages, punitive damages are not the appropriate way to remedy these problems; instead, we should address the problematic rules of lawyer compensation. See Kelly, *supra* note 32, at 1441 (arguing that punitive damages merely “dodge” problems with class actions and other difficulties of litigation, while also “undermin[ing] efforts to correct these problems”).

function, however, is largely incidental to the underlying purpose: to give a plaintiff access to a mechanism that promises to offer some repair to the plaintiff's injured dignity.

One final issue bears mention here. In what follows, I sometimes address the damages sought under the victim vindication model as aggravated damages that compensate the plaintiff for the injury to the victim's dignity. These are terms that some of the model's adherents might reject because labeling the vindication of the victim's dignity in *compensatory* terms may distort the underlying spirit of revenge or vindication animating this conceptual approach.¹⁰⁶ I accept the point being made here, and I do not want to crudely mischaracterize these other views.

From an external perspective, however, the victim vindication model and what I call aggravated damages are interested in the same thing: giving the plaintiff unfettered control over the choice to seek a remedy, usually in the form of money that would go directly to the plaintiff, designed to repair the injury to her dignity. Moreover, describing the victim vindication damages in compensatory terms may broaden their appeal beyond enthusiasts of revenge.¹⁰⁷ Saying that aggravated damages will give the victim control over a remedy designed to repair a harm not otherwise captured by compensatory damages makes intuitive sense. By contrast, saying that victims should have a form of revenge when we already have a criminal justice system (or a retributive damages scheme) might not make sense to many people. Finally, from a pragmatic perspective, the more the victim vindication proponents characterize aggravated damages as revenge or punishment, rather than as a form of compensation, the more they will invite the kind of constitutional scrutiny that they probably hope to avoid.

¹⁰⁶ Indeed, scholars like Arthur Ripstein, Sebok, and now Colby specifically reject the equation between "vindicating" the plaintiff's interest in dignity repair and "compensating" the plaintiff for the injury to her dignity, claiming that dignity is not something that is "compensable." See, e.g., Colby, *supra* note 5, at 435-36, 436 n.187 (citing Ripstein and Sebok with approval). But see Ernest J. Weinrib, *Punishment and Disgorgement as Contract Remedies*, 78 CHI-KENT L. REV. 55, 91 (2003) ("The courts properly regard such [aggravated] damages as compensatory rather than punitive, since they repair a loss, albeit an intangible one.").

¹⁰⁷ See Markel, *Retributive Damages*, *supra* note 1, at 270-71 (explaining the differences between retributive justice and revenge).

2. Possible Limits on Aggravated Damages

To be clear, I urge the use of an aggravated damages scheme *only if dignitary harms are not already covered under traditional compensatory damages in the particular jurisdiction*. Moreover, if the purpose of aggravated damages is to *repair* the harm to an individual's dignity, one might plausibly entertain whether juries (or victims) could instead require the defendant to apologize to the plaintiff if juries (or victims) believe that an apology or some other remedial sanction—such as an injunction or a restorative justice process involving a victim-defendant mediation or circle of concern and care¹⁰⁸—sufficiently substitutes for (or ought to supplement) money damages. The discussion below will assume *arguendo* that some gap exists to warrant the need for aggravated damages in the form of money.

Some might fear, however, that every injury to a victim will seem to be an invitation to a jury to award aggravated damages. Insofar as this fear is justified, jurisdictions might consider limiting aggravated damages only to cases satisfying two conditions. First, jurisdictions might decide that corporate plaintiffs or other entities should not be able to collect aggravated damages because they lack the kind of dignity or autonomy interest that we find necessary to vindicate through the use of aggravated damages.¹⁰⁹ A second limit, about which I am somewhat ambivalent, would restrict such damages only to those cases where the plaintiff is suing a defendant on the basis of a tort focused on denigrating the status of the individual *qua* *that* individual. Thus, mass-products-based disputes, such as those in *BMW v. Gore* or *Philip Morris*, would exclude aggravated damages on these grounds because the purpose of the conduct in those situations was not to make false

¹⁰⁸ See, e.g., Circles—Restorative Justice, <http://www.restorativejustice.org/intro/tutorial/processes/circles> (last visited Mar. 15, 2009) (explaining the use of “healing circles” as a way to address harm to a victim caused by an offender); see also Dan Markel, *Wrong Turns on the Road to Alternative Sanctions: Reflections on the Future of Shaming Punishments and Restorative Justice*, 85 TEX. L. REV. 1385 (2007).

¹⁰⁹ In the successor article to this one, *Punitive Damages and Complex Litigation*, *supra* note 9, I discuss the intelligibility of punishing a corporation for its misconduct. As a result, it might seem inconsistent to deny corporations access to dignity-based aggravated damages while at the same time thinking that they have sufficient independent identity to warrant retributive condemnation. I think that one can reduce this inconsistency by looking at the nature of the misconduct: for example, it might be that the insult to dignity is defamation of the corporate plaintiff (e.g., “the managers of X Corporation intentionally sell spoiled food”). There, the misconduct injures the dignity of the entity as a team, in which case the team arguably should be able to collect aggravated damages on top of other losses.

claims about the worth of the victim as such, but rather to undertake profit-seeking conduct that was insufficiently mindful of the legal norms protecting citizens. While this limit is controversial because some people might feel insulted even by mass-market misconduct—"the airline's general policy treated *all of us* with so little regard!"—that kind of misconduct is perhaps better addressed through the use of retributive damages, as that kind of misconduct should fall within that scheme's reprehensibility determination. To be sure, there will be hard cases at times, but judges and juries could decide that those are situations where the defendant must face both aggravated *and* retributive damages.

3. Reconciling Aggravated Damages and Retributive Damages

If jurisdictions permitted both aggravated and retributive damages, how would they work alongside each other? Plaintiffs would have control over aggravated damages. They could settle or not seek aggravated damages—the same as with compensatory damages. If the plaintiff sought aggravated damages and retributive damages, the jury would decide what amount (or other remedy) should go to the plaintiff for the sake of vindicating the victim's interest in her dignity. The jury would also decide whether the defendant's misconduct warranted retributive condemnation of the sort targeted by the reprehensibility scale. Having both kinds of damages available facilitates sharpened jury decision making about which damages are achieving which purposes. Without such a pluralistic structure, juries cannot ensure that the appropriate remedy is imposed, and they therefore risk providing an unnecessary windfall to the plaintiff (or to the state).¹¹⁰

¹¹⁰ I therefore find myself likely to disagree with those jurisdictions that have prohibited juries from finding out what portion of the extracompensatory damages goes to the state on account that such information will invariably and unduly inflate the amount of damages awarded. See Sharkey, *supra* note 5, at 438-39 & nn.358-60 (citing cases and statutes establishing that jurors are not informed of the division of extracompensatory damages). At least under my scheme, however, there are a number of robust safeguards and review mechanisms to ensure that this undue inflation of damages is less likely to occur. I share Professor Sharkey's skepticism toward the rationales for keeping juries in the dark about allocation. See *id.* at 439-40, 440 n.366 (citing Michelle Chernikoff Anderson & Robert J. MacCoun, *Goal Conflict in Juror Assessments of Compensatory and Punitive Damages*, 23 LAW & HUM. BEHAV. 313, 320-22 (1999) (finding both that jurors were actually more likely to award punitive damages when those damages went directly to plaintiffs, rather than to the government, and that the amount of damages was unaffected by the recipient of the award)).

4. Reconciling Aggravated Damages and Deterrence Damages

Aggravated damages and deterrence damages both assist in ensuring cost internalization. In theory, when determining deterrence damages, we should consider having juries use a multiple of the aggregated amount of compensatory and aggravated damages. In practice, however, it is unlikely that there will be many cases involving both aggravated and deterrence damages. That's because when there is a vivid injury to one's personal dignity, it is less likely that the defendant will escape having to compensate the plaintiff for the harm. For example, defamation might warrant aggravated damages, but, because of its public features, it is unlikely that deterrence damages would be appropriate, unless the plaintiff could show a strong likelihood that the defendant would not have to pay in full for the harm that the defendant caused.

5. Standard of Review for Aggravated Damages

To the extent that a jurisdiction characterizes aggravated damages as serving a truly compensatory or reparative role, these damages, like deterrence damages, should not be viewed as subject to the heightened review in federal cases called for by the Court in *Cooper Industries*.¹¹¹ Rather, aggravated damages should be subject to the same kind of review applied to other damages designed to compensate a plaintiff's injury. That standard of review would not mean that due process review is unavailable.¹¹² But it does suggest that judicial review of aggravated damages that go exclusively to the plaintiff should be somewhat deferential and thus similar to the judicial review extended to compensation for injuries generally. And, of course, to the extent that aggravated damages are excessive because they seem to figure in amounts to noninjured parties, they ought to be subject to some of the same limits appearing in *State Farm* and *Philip Morris*: namely, evidence of harms to nonparties, especially outside the state, is only permitted if it helps prove and assess the nature and scope of the in-

¹¹¹ See *Cooper Indus., Inc. v. Leatherman Tool Group, Inc.*, 532 U.S. 424, 436 (2001) (requiring de novo review of the amount of punitive damages awarded in federal cases).

¹¹² And, of course, legislatures may decide to impose caps or multiplier limits on aggravated damages in the way that some jurisdictions do for other damages. See Rustad, *supra* note 29, at 1300 (listing different types of "tort reform"). But see Markel, *Retributive Damages*, *supra* note 1, at 330-32 (providing a critique of the use of caps and multipliers in various cases).

jury to the dignity of the plaintiff, and not if the evidence is used to punish the defendant based on harms to others or to regulate a defendant's out-of-state conduct.¹¹³

C. *A Pluralistic Extracompensatory Damages Scheme: Applications*

As mentioned above, this Article endorses a pluralistic approach to extracompensatory damages. In some situations, the retributive function of extracompensatory damages is appropriate, even when aggravated or deterrence damages are not. Likewise, it may be the case that deterrence damages are warranted, even when aggravated or retributive damages are not. (It is less likely that aggravated damages would be warranted in the absence of some amount of retributive damages since that which would insult the dignity of a plaintiff would also tend to register on the reprehensibility scale.) And, of course, it might be that all three kinds of damages are warranted in a given case.

How might such a trifurcated structure work? It might be helpful to look at some of the fact patterns of major punitive damages cases to get a sense of how this structure would operate. If we use the facts of the Exxon Valdez oil spill as an example, aggravated damages against the company would be unavailable since the conduct of the company was reckless regarding the harm caused¹¹⁴ and was not directed at any one individual or collective victim. Exxon's direct misconduct (independent of vicarious liability for its drunk captain's misconduct) grew out of its management's knowing retention of an alcohol-abusing ship captain in a high-risk endeavor.¹¹⁵ This is not the kind of harm that would, in my view at least, trigger a need to compensate the injury to a particular plaintiff's dignity. As to deterrence damages, the operative question would be whether, and how likely, it is that Exxon would have escaped having to pay for the harm it caused to these plaintiffs. Because of the vividness and magnitude of the harm, as well as the ca-

¹¹³ This limited use of the evidence of wrongs to other persons makes sense in situations where the plaintiff needs to show that the misconduct in question was not an accident or mistake, but rather a malicious denigration of the plaintiff's dignity interests.

¹¹⁴ See *Exxon Shipping Co. v. Baker*, 128 S. Ct. 2605, 2614 (2008).

¹¹⁵ See *In re Exxon Valdez*, 296 F. Supp. 2d 1071, 1077 (D. Alaska 2004) ("Exxon officials knew that it was dangerous to have a captain with an alcohol problem commanding a supertanker . . . [and] knew that carrying huge volumes of crude oil through Prince William Sound was a dangerous business, yet they knowingly permitted a relapsed alcoholic to direct the operation of the *Exxon Valdez* . . ."), *vacated and remanded*, 490 F.3d 1066 (2007), *vacated and remanded sub nom.*, *Exxon Shipping Co. v. Baker*, 128 S. Ct. 2605 (2008).

capacity of the defendant to pay for the harm, it was unlikely that Exxon would not face suit for its misconduct, and thus deterrence damages would be inappropriate. On the other hand, the recklessness of Exxon's managerial misconduct would warrant some finding of reprehensibility and therefore an award of retributive damages toward the low end of the scale. If we assume, for sake of argument, that Exxon's reprehensibility was a two on a scale of one to twenty (with twenty being the most reprehensible), and that a finding of two entailed a one percent of net value penalty, then a multibillion dollar award in retributive damages against Exxon would not be inappropriate under the retributive damages scheme.¹¹⁶ (Some might suggest that Exxon's size is being taxed unfairly under this model. I address this criticism at length in *Retributive Damages*.¹¹⁷)

By contrast, consider the facts in *BMW v. Gore*. Dr. Gore sued BMW because BMW failed to disclose that it had repainted the car Dr. Gore purchased, in an effort to obscure the fact that the car had suffered minor damage in the manufacturing or transportation process.¹¹⁸ This nondisclosure was pursuant to a policy that was lawful in many states but not in Alabama, where Dr. Gore lived.¹¹⁹ Applying the pluralistic framework, we note first that because the economic harm that Dr. Gore endured was associated with the mass production of goods, there was no real individualized dignitary harm worth vindicating (assuming that the second limit on aggravated damages that I suggested is applied¹²⁰). But because of the nondisclosure and the relatively obscure nature of the harm, there was a relatively low likelihood that Dr. Gore—or others similarly positioned—would have discovered

¹¹⁶ There are some interesting questions relating to time and punishment. The penalty could be based on the net financial condition of the individual/entity *at the time of the misconduct* (adjusted for inflation) in order to deter the defendant from trying to reengineer its finances to its apparent detriment at adjudication. The goal behind benchmarking the defendant's value or financial condition this way is to ensure that the defendant is not penalized for growth or wealth independent of the tort. On the other hand, if there is a reason to suspect that the defendant's growth or current wealth is driven by the results of the underlying misconduct—say a person profits from wrongfully accessing another's trade secrets—that would be a good reason for instead using the defendant's financial condition at the time of adjudication, assuming it is higher than at the time of the commission of the tort.

¹¹⁷ See Markel, *Retributive Damages*, *supra* note 1, at 290-96.

¹¹⁸ *BMW of N. Am., Inc. v. Gore*, 517 U.S. 559, 563-64 (1996).

¹¹⁹ *See id.* at 570-73.

¹²⁰ *See supra* subsection II.B.2.

the harm, suggesting a need for deterrence damages.¹²¹ As to retributive damages, the appropriate reprehensibility score would be quite low in *BMW v. Gore* because the nondisclosure of the policy regarding retouched cars was lawful in some jurisdictions and because BMW's failure to disclose a material, but not life-threatening, fact about its cars should be viewed with disfavor but not serious outrage, moral or otherwise.¹²²

Having explored the basic structure of a pluralistic approach to extracompensatory damages, let's now turn to one of the most important aspects of implementation: whether and which procedural safeguards should attach and on what rationale.

III. PROCEDURAL PROTECTIONS, PLURALISM, AND PUNITIVE DAMAGES

As described earlier in Section I.B, the retributive damages structure already entails a number of limits on the amount of retributive damages that a defendant would face in any given case: the guidelines and commentary would provide much more notice than currently exists in common law jurisdictions,¹²³ while the use of a scale that tracks reprehensibility would reduce the arbitrariness of the punishments doled out to similarly situated defendants who commit the same misconduct. This Part examines, albeit in a relatively preliminary manner, which additional safeguards may be warranted for "simple" civil

¹²¹ Cf. *State Farm Mut. Auto. Ins. Co. v. Campbell*, 538 U.S. 408, 433-43 (2003) (Ginsburg, J., dissenting) (discussing how the defendant targeted the "weakest of the herd"—"the elderly, the poor, and other consumers who are . . . most vulnerable to trickery and deceit," including persons like the Campbells and those "unlikely to defend themselves" and therefore unlikely to take action that would force the defendant to compensate them).

¹²² In *Pacific Mutual Life Insurance Co. v. Haslip*, 499 U.S. 1, 12-14 (1991), the defendant was vicariously liable for the embezzlement of plaintiffs' insurance payments, but the company also was on notice that it had a rogue agent, much like in *Exxon*. The defendants in *Haslip* and *Exxon* could plausibly be said to have exercised manifestly insufficient care to ferret out misconduct by their agents. However, those cases would warrant lower reprehensibility scores than the defendants in *TXO Production Corp. v. Alliance Resources Corp.*, 509 U.S. 443 (1993), and *State Farm*, where the wrongs by the tortfeasors were purposeful and deceitful. Those four cases involved economic torts, however, and should be considered less reprehensible than the misconduct perpetrated by the defendant in *Philip Morris*, where the defendant's outrageous misconduct years ago effectively amounted to mass manslaughter.

¹²³ See *supra* text accompanying notes 55-57 (proposing that states create guidelines that would allow juries to measure reprehensibility more objectively).

litigation involving the potential for retributive and nonretributive extracompensatory damages.¹²⁴

Section A begins with some background to the debate over procedural safeguards. Some scholars think that punitive damages require virtually no procedural safeguards, while others argue that the full panoply of constitutional criminal procedural safeguards should apply. Contrary to both views, I argue that the determination of whether and which procedural safeguards should apply depends on which purposes are being advanced. I first detail the procedures that would be warranted for aggravated and deterrence damages, which are largely compensatory in nature and therefore require a less searching review for abuse. Then, for cases involving retributive damages, I argue that an intermediate level of procedural safeguards is warranted.

Section B applies this logic to a cluster of specific issues, including standards of proof, jury trial rights, access to counsel, the privilege against self-incrimination, bifurcated proceedings, standards of appellate review, and double jeopardy matters associated with exposure to multiple punishments for the same misconduct. The double jeopardy discussion also examines the interactive effects that retributive damages should have vis-à-vis the criminal justice system.

A. *The Debate over Punitive Damages and Procedural Safeguards*

1. Some Background

Few things have divided scholars of punitive damages law more than the issue of procedural safeguards and whether they are constitutionally and normatively required. Some scholars, such as Professors Galanter and Luban, argue that defendants in punitive damages cases warrant no additional procedural safeguards because the safeguards that the accused enjoy in criminal cases are to protect against *governmental*, not private-party, overreach.¹²⁵ On their view, criminal defendants need additional procedural safeguards because of “two concerns—about the centralized power of the state and about state abuse of prisons and physical violence.”¹²⁶ Because they believe that neither concern is implicated by punitive damages—actions for punitive damages are brought by private parties who cannot obtain relief in the

¹²⁴ For a discussion of safeguards primarily related to wrongs involving torts and complex litigation, see Markel, *Punitive Damages*, *supra* note 9.

¹²⁵ Galanter & Luban, *supra* note 3, at 1457-58.

¹²⁶ *Id.* at 1457.

form of physical punishment—there is no need for increased procedural safeguards with respect to punitive damages claims.¹²⁷

By contrast, notwithstanding the Supreme Court's efforts to regulate punitive damages under the Constitution over the last fifteen years, some scholars and courts maintain that punitive damages are improper either *because* or *to the extent that* they are being used to advance public goals such as retributive justice.¹²⁸ Joining this chorus more recently, Professor Colby argues that using punitive damages to punish public wrongs violates procedural due process in the absence of the criminal procedural safeguards afforded under the Constitution.¹²⁹ Colby's view instead is that punitive damages are permitted under the Constitution only to vindicate the private wrong against the plaintiff and thus operate as a "form of legalized private revenge."¹³⁰ Additionally, as discussed earlier, Colby argues that extracompensatory damages for the sake of cost internalization are constitutionally permissible, even after *Philip Morris*, so long as they are not deemed "punitive."¹³¹ Using the nomenclature of this Article, Colby endorses both aggravated and deterrence damages as constitutionally permissible but views retributive damages as constitutionally problematic because they serve the public interest in retributive justice without all of the procedural protections that one receives in criminal cases.

2. Pluralism and Procedural Safeguards

To my mind, both sets of views on the safeguards issue are misguided because the answer to which procedural safeguards are consti-

¹²⁷ *Id.* at 1457-58.

¹²⁸ See Grass, *supra* note 7 (concluding that punitive damages are penal in nature and therefore are essentially criminal law sanctions); Redish & Mathews, *supra* note 12 (arguing that the seeking of punitive damages is, in effect, a state function that demands constitutional protections); Wheeler, *supra* note 7 (contending that where the government authorizes private individuals to seek damages for the purposes of punishment—punitive damages—the Constitution requires that criminal law constitutional protections apply); see also Colby, *supra* note 7, at 606 n.73 (providing citations to courts and commentators noting that defendants subject to punitive damages lack the protections provided in criminal trials).

¹²⁹ Colby, *supra* note 5, at 415-16.

¹³⁰ *Id.* at 396. In this respect, Professor Colby is basically embracing the approach endorsed by Professor Sebok, *supra* note 5.

¹³¹ See Colby, *supra* note 5, at 396 ("Williams thus allows the states to address the underdeterrence concern by implementing the recommendation of law and economics scholars to create a category of extracompensatory damages designed to ensure optimal deterrence.").

tutionally necessary and normatively desirable depends on which purpose of punitive damages is being pursued.

a. *Aggravated Damages and Deterrence Damages*

Earlier in Part II, I discussed the standard of review associated with both aggravated and deterrence damages. That discussion foreshadows the proper approach to thinking about procedural safeguards. Let's start with aggravated damages, which, like compensatory damages, the plaintiff should be able to seek or forbear from seeking in trial, or settle privately without the regulation of the state. The same power of the victim should apply to deterrence damages in a post-*Philip Morris* world because the focus for such damages must train on the likelihood of the defendant's evading compensating the plaintiff for the harm caused. In both situations, the plaintiff has the right to seek payment from the defendant for the full harm that she caused to the plaintiff. And because aggravated and deterrence damages are not assigned for the purpose of punishment, there would be little need for them to have procedural safeguards vastly different from those that should attach to cases in which only compensatory damages are sought.¹³²

That does not mean that courts should not review the amount of aggravated or deterrence damages. Cases like *State Farm* and *Philip Morris* should still apply insofar as they provide assistance in limiting potential federalism concerns in cases involving the resolution of complex wrongdoing across jurisdictions. In other words, those cases might still provide guidance on how to reduce the incidence of states' using a tort plaintiff to regulate a defendant's conduct in another state, especially when that conduct may be lawful.¹³³ The Court would almost certainly want to ensure that suits involving aggravated and deterrence damages did not run afoul of this federalism principle.

However, whether other constitutional restrictions apply to aggravated damages is hard to say. Professor Colby suggests that the Court might wish to apply all current punitive damages due process rules to

¹³² Some scholars have questioned whether we in fact need more due process protection for the review of compensatory damages awards. See, e.g., Mark A. Geistfeld, *Due Process and the Determination of Pain and Suffering Tort Damages*, 55 DEPAUL L. REV. 331 (2006) (arguing that constitutional due process constraints on punitive damages should also apply to pain and suffering damages).

¹³³ See Michael P. Allen, *The Supreme Court, Punitive Damages and State Sovereignty*, 13 GEO. MASON L. REV. 1, 21-25 (2004) (reading *State Farm* through a federalism lens and finding that the holding rests upon respect for state sovereignty).

aggravated damages.¹³⁴ Professor Sebok, by contrast, is adamantly opposed to the use of the presumptive single-digit multiplier established in *State Farm*.¹³⁵ My view is that review of aggravated damages should be relatively deferential, especially when the jury is informed about the nature and recipient of aggravated damages as compared to the nature and recipient of retributive damages. That said, if aggravated damages were required (by courts or legislatures) to be set at a fixed ratio of compensatory damages, this requirement would logically trigger the same inegalitarian and troubling effect of stating, in cases tied to variable economic losses, that an encroachment against a poor person's dignity is worth less than an encroachment upon a richer person's dignity.¹³⁶ Using a multiplier approach would, in many cases involving risk of injury, just exacerbate much of the extant arbitrariness in the context of compensatory damages.¹³⁷

As to deterrence damages, the procedural safeguards that one might advocate would depend on one's reading of *Philip Morris*. Under my view of *Philip Morris*, the permissible inquiry is limited to only the likelihood that the defendant won't have paid compensation to the plaintiff in the case.¹³⁸ That limit substantially reduces the scope of potential abuse by juries, especially in the simple litigation context and, accordingly, creates substantially less need to impose a very stringent set of regulations or safeguards. On the other hand, if Professor Colby is correct in his reading of *Philip Morris*, I would imagine that there would be a number of other procedural protections that a defendant would plausibly desire and that the Court would apply.¹³⁹ The

¹³⁴ See Colby, *supra* note 5, at 455-56 n.281. Somewhat relatedly, Professor Romero argues that legislatures should impose limits on all punitive damages. Leo M. Romero, *Punitive Damages, Criminal Punishment, and Proportionality: The Importance of Legislative Limits*, 41 CONN. L. REV. 109 (2008).

¹³⁵ See Sebok, *supra* note 5, at 1029 (describing the ratio rule as regrettable and lacking in principle).

¹³⁶ See Markel, *Retributive Damages*, *supra* note 1, at 290-91 (explaining that such income-based practices undermine commitments to human equality).

¹³⁷ See Geistfeld, *supra* note 132, at 342 (noting that plaintiffs "with similar pain-and-suffering injuries often are awarded significantly different amounts of damages").

¹³⁸ For references to other scholars who share my view on *Philip Morris*'s implications, see Colby, *supra* note 5, at 469 n.337, and sources cited therein.

¹³⁹ Specifically I'd imagine that such a reading of *Philip Morris* would warrant applying the kind of searching appellate review that the Court employed in *BMW v. Gore* and *State Farm*, though with less attention to concerns about reprehensibility and more focus on fidelity to the federalism and fair-notice concerns. Colby himself is expressly noncommittal on what constitutional limits would apply to deterrence damages under this scenario. *Id.* at 476 n.371. If Professor Sharkey's proposal for nonpunitive com-

scope of such protections might draw inspiration from the discussion below regarding procedural safeguards for retributive damages.

b. *Retributive Damages*

As to retributive damages, the questions about procedural safeguards are more complex. Recall Colby's view that "it is unconstitutional to punish for public wrongs without criminal procedural safeguards."¹⁴⁰ By contrast, others like Galanter and Luban have argued that there really is no need to apply the procedural safeguards commonly used in criminal cases because, in punitive damages cases, it is a private party who is instigating the punishment against the defendant.¹⁴¹

How would these views apply to the proposed retributive damages framework? Colby initially appears to dismiss my normative proposal as unconstitutional, based on the idea that retributive damages serve the public interest in retributive justice and yet are not criminal sanctions.¹⁴² As I argue below, the mistake that Colby—along with others¹⁴³—makes is in concluding that punitive damages for public retributive purposes cannot pass constitutional muster if they are designed to serve as an intermediate sanction that would be accompanied by a proportionate level of procedural protections. For reasons that follow from the "intermediate" nature of the sanction, Galanter and Luban are also mistaken to shield punitive damages from increased scrutiny.

pensatory societal damages is compatible with *Philip Morris*, then it is also a useful place to look for guidance on some of the constitutional issues regarding what I have called deterrence damages. See Sharkey, *supra* note 5, at 428-40.

¹⁴⁰ Colby, *supra* note 5, at 469.

¹⁴¹ Galanter & Luban, *supra* note 3, at 1454-60.

¹⁴² See Colby, *supra* note 5, at 445-46 (raising the concern that my view of retributive damages basically creates an end-run around the Constitution).

¹⁴³ See, e.g., *Murphy v. Hobbs*, 5 P. 119, 121 (Colo. 1884) (criticizing a judge's punitive damages instructions to a jury on the grounds that, inter alia, "punishment by fine is inflicted" without the defendant being afforded the protections of criminal procedure); *Spokane Truck & Dray Co. v. Hoefer*, 25 P. 1072, 1074 (Wash. 1891) ("It seems to us that there are many valid objections to interjecting into a purely civil action the elements of a criminal trial, intermingling into a sort of a medley or legal jumble two distinct systems of judicial procedure."); Grass, *supra* note 7 (arguing that punitive damages perform the same function as criminal law and therefore demand the same constitutional protections); Wheeler, *supra* note 7, at 322-51 (outlining the constitutional provisions suggesting that punitive damages should trigger the protections of criminal procedure).

i. Colby's Error

The conclusion that Colby draws about my proposal is mistaken for two reasons. First, it is wrong because retributive damages would do exactly what the Supreme Court thinks that punitive damages may lawfully do now: serve as “quasi-criminal” sanctions to advance the public interest in retributive justice.¹⁴⁴ Colby’s argument, notwithstanding its ingenuity, is by today’s standards a radically external critique of the Supreme Court’s understanding of the goals achieved by punitive damages. Indeed, from a constitutional perspective, my own account—proposing retributive damages as an intermediate sanction—demonstrates a far better “fit” with the Court’s own statements on punitive damages. For decades now, the Supreme Court has described punitive damages as “private fines levied by civil juries”¹⁴⁵ that “advance governmental objectives” of retribution.¹⁴⁶ This point of view found expression again in the Supreme Court’s recent decisions in *Exxon* and *Philip Morris*.¹⁴⁷ And, notwithstanding the quasi-criminal-punishment nature of the punitive damages sanction and its furtherance of retribution, the Court has repeatedly declined the invitation by defense counsel to insist upon any, let alone all, of the Constitution’s criminal procedural safeguards. That seems an important, if not decisive, point as a matter of constitutional interpretation.

Of course, Colby’s argument is not predicated on precedent so much as logic. The basic idea shared by Colby and like-minded critics of punitive damages is that punitive damages levied in the public’s interest are no different than criminal penalties and thus deserve the same safeguards that the Constitution affords in that context. But a

¹⁴⁴ See *Cooper Indus., Inc. v. Leatherman Tool Group, Inc.*, 532 U.S. 424, 432 (2001) (acknowledging that punitive damages have been labeled “quasi-criminal” (citing *Pac. Mut. Life Ins. Co. v. Haslip*, 499 U.S. 1, 19 (1991))); *Gertz v. Robert Welch, Inc.* 418 U.S. 323, 350 (1974) (stating that punitive damages “punish reprehensible conduct” and “deter its future occurrence”).

¹⁴⁵ *Gertz*, 418 U.S. at 350.

¹⁴⁶ *Haslip*, 499 U.S. at 47-48 (O’Connor, J., dissenting) (citing *Gertz*, 418 U.S. at 350); see also *Browning-Ferris Indus. of Vt., Inc. v. Kelco Disposal, Inc.*, 492 U.S. 257, 275 (1989) (“[P]unitive damages are imposed through the aegis of courts and serve to advance governmental interests . . .”).

¹⁴⁷ See *Exxon Shipping Co. v. Baker*, 128 S. Ct. 2605, 2621 (2008) (“The consensus today is that punitives are aimed not at compensation but principally at retribution and deterring harmful conduct.”); *Philip Morris USA v. Williams*, 549 U.S. 346, 352 (2007) (“This Court has long made clear that ‘[p]unitive damages may properly be imposed to further a State’s legitimate interests in punishing unlawful conduct and deterring its repetition.’” (quoting *BMW of N. Am. v. Gore*, 517 U.S. 559, 568 (1996))).

problem with this line of argument is that it seems to assume that there *is* an identifiable and fixed basket of criminal procedural safeguards that applies to all cases of criminal punishment.¹⁴⁸ The truth of the matter is more complex. The Supreme Court and state legislatures and courts have extensively scaled the amount and intensity of both constitutional and statutory procedural safeguards to the severity of the punishment.¹⁴⁹ Importantly, under the Federal Constitution, a defendant facing the death penalty receives more protections than one facing incarceration.¹⁵⁰ Similarly, a defendant facing six months

¹⁴⁸ After I showed Colby a draft of this paper, he revised his paper, *supra* note 5, to clarify that punitive damages for public wrongs (i.e., retributive damages) should receive only the amount of procedural protections that criminal fines receive, which is a more nuanced position than the claim that public-minded punitive damages require all procedural safeguards from the criminal context. Importantly, Colby's final position is still different from mine. I see retributive damages as a civil sanction that is qualitatively different from criminal penalties because such damages entail no resulting conviction, less of a stigma, and no intended collateral consequences. Nonetheless, Colby sees such "retributive damages" as constitutionally equivalent to criminal fines.

Significantly, Colby is not alone in claiming that such retributive damages (or publicly minded punitive damages) would likely equate to criminal penalties warranting criminal procedural safeguards. See, e.g., Redish & Mathews, *supra* note 12, at 20 ("Financial penalties imposed for no purpose other than to punish are appropriately categorized as coercive, just as imprisonment is. This conclusion is reinforced by the simple fact that the same special constitutional protections apply in criminal cases seeking only the imposition of financial penalties as apply in cases in which imprisonment is at issue."); Wheeler, *supra* note 7, at 337 ("Either the purpose of the sanction is punitive, in which case all of the procedural safeguards apply, or the purpose is not punitive, in which case none applies."). But these views are, to my mind, incorrect. As illuminated in the text, there is no right to counsel for criminal fines, nor is there a right to a jury trial in that context. See generally *Scott v. Illinois*, 440 U.S. 367 (1979) (holding that counsel is only required where actual imprisonment results); *Duncan v. Louisiana*, 391 U.S. 145, 159 (1968) (holding that a jury trial is not constitutionally required for cases where the maximum penalty is incarceration of less than six months). Moreover, these accounts similarly fail to contemplate the possibility, developed in the text, that retributive damages could serve as a designated civil and intermediate sanction, thus warranting a smaller or otherwise different bundle of safeguards than would exist if criminal fines were applied.

¹⁴⁹ Some constitutional safeguards may also be scaled to the severity of the misconduct involved. See Eugene Volokh, *Crime Severity and Constitutional Line-Drawing*, 90 VA. L. REV. 1957 (2004) (outlining potential approaches to balancing the strength of constitutional protections against the crime that the government actor is seeking to prevent).

¹⁵⁰ See generally 1 WAYNE R. LAFAVE ET AL. CRIMINAL PROCEDURE § 1.8(e) (3d ed. 2008) (discussing special constitutional rules in the capital punishment context); Stuart P. Green, *Rationing Criminal Procedure: A Comment on Ashworth and Zedner*, 2 CRIM. L. & PHIL. 53, 54 (2008) ("[D]efendants charged with capital offenses are entitled to additional protections not generally available to defendants in non-capital cases, including a bifurcated trial at which guilt and sentencing are decided separately, special jury se-

or more in prison enjoys more constitutional protections than a defendant facing less than six months.¹⁵¹ And a defendant receiving any incarceration receives more constitutional protections than one receiving a noncarceral sentence.¹⁵² This last point is particularly relevant to retributive damages because criminal fines can be levied on someone without a right to appointed counsel or to a jury trial.¹⁵³ Moreover, the Supreme Court has contracted the scope of a number of procedural safeguards in cases involving relatively minor crimes.¹⁵⁴

Thus, if criminal fines can be levied without jury trials and counsel, it should follow that the intentionally less severe penalty of retributive damages could also have a different and perhaps smaller basket of procedural safeguards. Remember that retributive damages are situated as an intermediate sanction falling on the civil side between criminal fines and compensatory damages. Consequently, determinations of retributive damages against someone would not serve as predicates for impeaching that person's testimony in a future trial or as a basis for enhancing one's punishment in a subsequent criminal trial for the same or different conduct. Nor would they trigger any disqualifications professionally (e.g., debarment) or civically (e.g., jury service or voting). With retributive damages cases, the defendant faces no collateral consequences.¹⁵⁵ Moreover, the condemnatory stigma associated with retributive damages is designed to be interme-

lection procedures, special rules regarding the introduction of aggravating and mitigating evidence, and an automatic right of appeal.”).

¹⁵¹ See generally 1 LAFAYE ET AL., *supra* note 150, § 1.8(c) (discussing special constitutional rules applying only to crimes involving a certain period of time in prison).

¹⁵² See *Scott*, 440 U.S. 367 (holding that defendants may not be imprisoned absent the enjoyment of a right to counsel but that no right to counsel exists for less severe punishments); *Callan v. Wilson*, 127 U.S. 540 (1888) (guaranteeing a jury trial only for nonpetty criminal offenses); *Mackin v. United States*, 117 U.S. 348, 351 (1885) (requiring a grand jury indictment for crimes punishable by imprisonment); 1 LAFAYE ET AL., *supra* note 150 (discussing special constitutional rules applying only to crimes involving some time in prison).

¹⁵³ *Scott*, 440 U.S. at 373-74 (appointed counsel); *Codispoti v. Pennsylvania*, 418 U.S. 506, 512 (1974) (jury trial); *Duncan*, 391 U.S. at 159 (jury trial).

¹⁵⁴ See *Green*, *supra* note 150, at 57 n.13 (citing cases where, inter alia, the Court either has given the government more leeway in construing what counts as a search or seizure in cases involving relatively insignificant offenses or has limited the obligation to read a defendant her *Miranda* rights in minor cases).

¹⁵⁵ See generally Gabriel J. Chin, *Race, the War on Drugs, and the Collateral Consequences of Criminal Conviction*, 6 J. GENDER RACE & JUST. 253 (2002) (providing an overview of the range of collateral consequences).

diate in nature.¹⁵⁶ The defendant will not be branded “a criminal” with a conviction that trails forever after. These aspects of retributive damages distinguish them from criminal fines and justify their status as a civil penalty, notwithstanding the fact that they are intended as a sociolegal rebuke for wrongful misconduct and carry with them a penalty meant to advance a public interest.¹⁵⁷

¹⁵⁶ Cf. Comment, *Criminal Safeguards and the Punitive Damages Defendant*, 34 U. CHI. L. REV. 408, 411-12 (1967) (arguing that the stigma from a criminal fine is more apparent than the reputational loss from punitive damages even though the amount of a punitive damages award might exceed that of a fine for comparable conduct). Admittedly, some criminal offenses are “strict liability” in nature and thus might not be viewed as conveying as much condemnation as a retributive damages finding that says that one acted with malice or recklessness, but the broader point still holds that criminal sanctions are generally viewed as more condemnatory than those currently associated with punitive damages (or retributive damages, per this proposal).

If retributive damages are truly going to be intermediate in nature, then jurisdictions adopting retributive damages schemes may also have to adjust fines in the criminal context to always be some amount greater than the amount awarded in retributive damages for similar misconduct. But it is not clear that a conventional fine needs to be higher than a retributive damages award in order to signal that it is a steeper penalty; the mere fact of a criminal conviction (and the process leading to and consequences flowing from that conviction) may do the work of ensuring that the social meaning of a fine remains more distinctively condemnatory than the intermediate sanction associated with a retributive damages award.

¹⁵⁷ The Supreme Court extends great deference to a legislative determination that a penalty is civil or criminal, but occasionally, a multifactor test is used to illuminate (but not decide dispositively) whether a civil law intends and effects a criminal punishment instead:

Whether the sanction involves an affirmative disability or restraint, whether it has historically been regarded as a punishment, whether it comes into play only on a finding of *scienter*, whether its operation will promote the traditional aims of punishment—retribution and deterrence, whether the behavior to which it applies is already a crime, whether an alternative purpose to which it may rationally be connected is assignable for it, and whether it appears excessive in relation to the alternative purpose assigned are all relevant to the inquiry

Kennedy v. Mendoza-Martinez, 372 U.S. 144, 168-69 (1963) (footnotes omitted). Under this test, retributive damages would likely qualify as some form of punishment, but they do not involve the “affirmative disability or restraint” typically associated with criminal punishment through incarceration. See *Hudson v. United States*, 522 U.S. 93, 104 (1997). Nonetheless, retributive damages do share many characteristics with criminal sanctions: they are only assessed upon a finding of malice or recklessness; they promote the distinct end of retribution; they involve conduct that is frequently the trigger of criminal sanction; and they have no alternative purpose as compensation, first, because the structure for extracompensatory damages specifically contemplates the plaintiff’s receiving compensation for traditional and aggravated damages, and second, because the only money that a defendant pays in retributive damages to the plaintiff, per my proposal, is a small “reward” (not compensation) for channeling awareness of undetected misconduct into the public eye. With all of that said, in *Hud-*

In short, under the Constitution and by widespread legislative and state judicial practice, procedural protections are largely scaled according to the severity of the penalty.¹⁵⁸ Once this pattern is recognized, its implications for a *noncriminal* publicly minded penalty must be appreciated. As indicated earlier, the amount and nature of the procedural safeguards should reflect the intermediate nature of the sanction. Saying the word “intermediate” does not tell us, of course, where exactly on the continuum between compensatory damages and criminal fines to draw the line, but the safeguards should be enough to ensure additional confidence in the verdict than what would typically result from attempts to force the defendant to internalize her costs and to compensate her victims while still not equaling the amount required for a traditional criminal prosecution involving a fine. Moreover, if we want to remain faithful to a retributive vision of fair, accurate, modest, and humane punishment, it is important to show not only why Colby (and his like-minded predecessors) are wrong, but also why Galanter and Luban are mistaken—for taking the opposite view that effectively no additional procedural safeguards are necessary or desirable for punitive damages imposed for public retributive ends.¹⁵⁹

son the Court nevertheless found that the civil penalties of both fines and debarment were not sufficiently “criminal” in nature to trigger the double jeopardy protection against multiple criminal punishments for the same conduct. *Id.* at 105. If *Hudson* is still good law, retributive damages would permissibly operate as only a moderate form of civil sanction for several reasons: first, the legislature’s explicit characterization of the sanction as civil will carry great weight; second, the liability for retributive damages will not be equivalent to the status of a conviction; and third, per my proposal, they would not be intended to trigger any collateral consequences. For an earlier instructive overview of punitive-civil-sanctions jurisprudence focusing primarily on agency-initiated civil punishment, see Kenneth Mann, *Punitive Civil Sanctions: The Middle-ground Between Criminal and Civil Law*, 101 YALE L.J. 1795, 1800 (1992). See also Jonathan I. Charney, *The Need for Constitutional Protections for Defendants in Civil Penalty Cases*, 59 CORNELL L. REV. 478 (1974); Mary M. Cheh, *Constitutional Limits on Using Civil Remedies to Achieve Criminal Law Objectives: Understanding and Transcending the Criminal-Civil Law Distinction*, 42 HASTINGS L.J. 1325 (1991).

¹⁵⁸ See 1 LAFAYETTE ET AL., *supra* note 150, § 1.8(c); see also Mann, *supra* note 157, at 1870 (noting the constitutional idea that “the more severe the sanction, the more the procedure must protect against the sanctioning of the innocent”).

¹⁵⁹ See Galanter & Luban, *supra* note 3, at 1461 (“The appropriate measure to control punitive damages consists of a requirement that juries provide a plausible rationale for the size of punitive awards, coupled with a large dollop of judicial deference to the retributive sentiments jurors express in those awards.”).

ii. Galanter and Luban's Errors

Recall that Galanter and Luban argue that virtually no safeguards are warranted because punitive damages involve privately instigated lawsuits and no incarceration or state violence.¹⁶⁰ To the extent that their second point about prisons and violence is true,¹⁶¹ the point stems from a larger concern that requires articulation. With due respect to Galanter and Luban, the reason to consider procedural safeguards is not simply that we're concerned with limiting the centralized power of the state, but also that we're largely concerned with what the state may do with that power. Indeed, framed this way, Galanter and Luban's concern about state abuse of prisons and physical violence makes more sense—it is not that we're unconcerned with violence or prisons as such, but rather that we're properly worried about the state's use of coercive measures against individuals' autonomous choices more generally.

And that fear of state coercion (and condemnation) makes retributive damages eligible for concerns about due process, just as any other sanction that works a coercive condemnatory deprivation. To be sure, the intensity of the concern will vary, but the fact of concern that the penalty might be abusively or arbitrarily imposed remains. Nonetheless, Galanter and Luban state that “[t]he most important point is that punitive damages are sought by individual plaintiffs: they involve a totally decentralized use of the legal system to impose punishment, and they raise none of the classical liberal worries about aggrandizing state power.”¹⁶²

To my mind, the problem with Galanter and Luban's argument is that they fail to separate sufficiently the functions of detection, prosecution, adjudication, and punishment. In any punitive damages claim, it is still the court that adjudicates and the state that enforces the judgment against the defendant, regardless of who detects and

¹⁶⁰ *Id.* at 1457-58; *cf.* Darren Weirnick, *Punitive Damages Against Corporations: Functionalist Retributivism* (2001) (unpublished manuscript, on file with author) (“Because punitive damages exact only monetary penalties and are filed by private plaintiffs, it is not at all obvious that a higher standard of proof, or constitutional rights . . . should apply even when individuals are the defendants in punitive damages cases. . . . When corporations are the defendants, the parallels are even less obvious.”).

¹⁶¹ The violence might also apply to criminal fines or forfeitures if the defendant's noncompliance triggers a court's contempt—the same outcome that might occur in the context of civil punitive damages.

¹⁶² Galanter & Luban, *supra* note 3, at 1457-58.

prosecutes the claim.¹⁶³ And from the defendant's perspective, she will surely observe that abusive awards of retributive damages deplete her bank account much the way that abusive criminal fines do, making the purported "decentralization" of punishment through punitive damages of dubious significance,¹⁶⁴ and of even more dubious significance when the government sues for punitive damages in its own right.¹⁶⁵ A defendant reasonably wants procedural safeguards to avoid the mistaken assignation of censure and accompanying harsh treatment, regardless of whether the person who detects the underlying wrong is a public prosecutor or a private plaintiff.¹⁶⁶ Thus, the scope of the protections we want does not really turn on who is the source of information regarding the misconduct so much as on the nature of the sanction itself. Galanter and Luban make the contrary claim but never convincingly explain why.

¹⁶³ A judge presides over and rules on matters of law (and, in some cases, fact), and appellate courts subject damages awards to searching review. Moreover, it is the state that enforces the judgment against the losing party. So the state's role is by no means insignificant or radically different from the role that the state plays in a criminal law case; this overlap was even more pronounced during the hundreds of years in which private litigants prosecuted criminal cases and would even capture the fines imposed against a defendant. *See, e.g.,* *Steel Co. v. Citizens for a Better Env't*, 523 U.S. 83, 127-28 (1998) (Stevens, J., concurring) (explaining the historical role of private parties in the criminal justice system).

¹⁶⁴ *See* Zipursky, *supra* note 5, at 146 ("If the private plaintiffs are really delegates of state power, then the awesome power of the state is being used, albeit in a decentralized way.").

¹⁶⁵ *E.g.,* *United States v. Ely*, 142 F.3d 1113, 1121-22 (9th Cir. 1997) (holding that the government could still criminally prosecute the defendants, even though the Federal Deposit Insurance Commission (FDIC), acting as receiver for a failed bank, had already sued for punitive damages).

¹⁶⁶ An analogy might be found by reference to privacy intrusions. If a private investigator for a plaintiff comes rummaging (with no notice or temporary authorization) through my garage, I am reasonably upset or resentful. I can imagine I may reasonably be even more upset when the government is the intruder instead of a private investigator for a third party—after all, the government purports to act in my name. But to the extent that my choice is to shield something from exposure to public view, the legitimacy of the interest that I have in keeping that information private doesn't hinge simply on whether the intruder wears a government badge. Galanter and Luban seem to think that the Bill of Rights is normatively relevant only in actions where the state is also the prosecutor (and not just the punisher) and they rely for support on the expressly political fears that classical liberals had about the government's capacity to use its prosecutorial force to marginalize or attack political dissidents. Galanter & Luban, *supra* note 3, at 1457. That might be a basis for limiting the government's investigative powers through something like the Fourth Amendment, but it is not a basis for limiting the reach of procedural safeguards during trial or punishment where the fear that a person has about abuse is predicated on what the government does as judge or enforcer of punishment.

iii. Why an Intermediate Compromise?

Contra Galanter and Luban, we can see why some procedural safeguards are necessary: because in the context of retributive damages, the state has the power, through its courts and enforcement agencies, to impose a coercive condemnatory sanction. In other words, many of the same concerns of error and abuse that motivate procedural safeguards in the criminal context also arise in the retributive damages context, though to a lesser extent because the penalties and consequences are less condemnatory and severe. On the other hand, contra Colby, retributive damages awards (unlike criminal punishments) bring no designation of the defendant as a criminal who will subsequently endure a welter of collateral consequences.¹⁶⁷ Correspondingly, under current law, a defendant may be found liable for punitive damages on weaker standards of proof than if convicted for a crime; moreover, defendants in a civil case enjoy no privilege against self-incrimination vis-à-vis punitive damages liability, no constitutional right to appointed counsel, and no constitutional right against double jeopardy.

These differences give us a compelling reason to accept the designation of the retributive damages sanction as civil and intermediate. Indeed, it makes further sense to disallow certain safeguards precisely because the defendant enjoys various benefits in the civil context that she does not in the criminal context. In the criminal context, for example, the government can hold a defendant for pre-trial detention or make her pay bail; civil plaintiffs lack this power. Moreover, in criminal cases, the government can also gain pre-trial discovery against a defendant through grand jury investigations that seek testimony and documents from others without the defendant's enjoying reciprocity against the government. Because the defendant enjoys these advantages in the *civil* case, it makes sense to distinguish, from a constitutional and normative perspective, those scenarios where society seeks to punish a defendant through a moderate civil penalty and those where society wants to punish through a criminal penalty by

¹⁶⁷ I should note that collateral consequences are sometimes products of regulatory decisions as opposed to "punitive" ones, and in some contexts, they appear as a result of proof of misconduct, not of conviction. However, the dominant recent approach to the imposition of collateral consequences is to look at the defendant's convictions, rather than other sources of evidence regarding the defendant's misconduct. See Gabriel J. Chin, *Are Collateral Sanctions Premised on Conduct or Conviction?: The Case of Abortion Doctors*, 30 FORDHAM URB. L.J. 1685 (2003).

bringing the full force of social condemnation associated with those punishments.

In what follows, I address (albeit in an admittedly preliminary fashion) which additional procedural safeguards are appropriate in a retributive damages case (involving simple litigation).¹⁶⁸ Although “the full panoply” of criminal procedural safeguards does not constitutionally apply,¹⁶⁹ I explain why some heightened level of safeguards is normatively desirable. In short, for retributive damages we should strive to exhibit more concern for Type I error reduction than is warranted in suits involving mere compensatory damages but less concern for Type I error reduction than is warranted in criminal prosecutions of defendants in cases involving fines.

B. *Retributive Damages as Intermediate Sanction: Applying the Logic to Procedural Safeguards*

This Section contends that certain legislative safeguards should be applied to retributive damages understood as an intermediate sanction. Before proceeding further, it bears emphasis that *Retributive Damages* already provides a number of limits regarding the amount of retributive damages.¹⁷⁰ The limits that follow would stand in addition to the ones already discussed.

1. Standard of Proof

Under the traditional rule, punitive damages in a civil suit can be awarded if the plaintiff is able to establish a defendant's liability for punitive damages (specifically the heightened mens rea of recklessness or malice) by a mere preponderance of evidence.¹⁷¹ This stands in contrast to the criminal cases that require evidence proving the de-

¹⁶⁸ See also Markel, Punitive Damages, *supra* note 9 (addressing the appropriate safeguards for complex litigation).

¹⁶⁹ Cf. Sebok, *supra* note 5, at 1002 (discussing the claim that the full panoply of constitutional safeguards should apply to punitive damages).

¹⁷⁰ See Markel, *Retributive Damages*, *supra* note 1, at 287-89 (calling for sentencing guidelines and commentaries to inform the finding of a defendant's reprehensibility on a scale, which would ensure that determinations of the amount of retributive damages are not made on an ad hoc basis, jury by jury).

¹⁷¹ See *id.* at 250-51 (noting this traditional rule but also explaining that most states now require proof of mens rea by clear and convincing evidence). The Supreme Court has not yet required a heightened standard of proof for punitive damages. See *Pac. Mut. Life Ins. Co. v. Haslip*, 499 U.S. 1, 22 n.11 (1991).

defendant's guilt beyond a reasonable doubt.¹⁷² Over the last two decades, many states have adopted the intermediate standard of "clear and convincing" evidence to govern the awards of punitive damages.¹⁷³

The embrace of a heightened, yet still intermediate, standard of proof (something like clear and convincing evidence) is more consistent with the view that retributive damages are an intermediate sanction that warrant an intermediate evidentiary standard of proof. More importantly, it is also consistent with the retributivist concern for striking the correct balance between Type I and Type II errors. As a sanction's severity increases, we worry more about its misapplication. The clear and convincing standard will theoretically create more Type I errors than "beyond a reasonable doubt" but fewer Type I errors than "preponderance of the evidence."¹⁷⁴

Notwithstanding the views of those who, like Galanter and Luban, believe that "preponderance of the evidence" is sufficient,¹⁷⁵ this recent trend toward "clear and convincing" makes good sense once retributive damages are understood as an intermediate sanction to advance the public interest in retributive justice. However, as mentioned in Part II, where the awards to the plaintiff are simply designed to ensure full cost internalization by the defendant in the form of aggravated and deterrence damages, a preponderance of the evidence should be sufficient even though jurisdictions do no grievous wrong in asking for heightened review as well.¹⁷⁶

2. Availability and Standard of Appellate Review

In order to facilitate notice to citizens and evenhandedness across cases, my retributive damages proposal involves a rebuke by the public

¹⁷² See *In re Winship*, 397 U.S. 358, 364 (1970) (confirming that due process mandates the beyond a reasonable doubt standard).

¹⁷³ 1 LINDA L. SCHLUETER & KENNETH R. REDDEN, PUNITIVE DAMAGES, § 5.3 (H) (2) (5th ed. 2005). One state, Colorado, requires juries to find that the proof establishes culpability beyond a reasonable doubt for punitive damages to attach. COLO. REV. STAT. § 13-25-127(2) (2008).

¹⁷⁴ Cf. *Santosky v. Kramer*, 455 U.S. 745, 756 (1982) (discussing the use of the clear and convincing evidence standard for civil cases involving the potential imposition of stigma or deprivation of liberty). Whether all jurors differentiate among the different standards is hard to know, but inasmuch as it is a problem, clear instructions about the different standards along with examples may help.

¹⁷⁵ See Galanter & Luban, *supra* note 3, at 1459-60; Weirnicks, *supra* note 160.

¹⁷⁶ Those who view aggravated damages as "private punishment," rather than as compensation for uncompensated injuries to dignity, may welcome a higher standard of proof, though they have not all said as much.

to be made consistent with the scale of punishment provided by the sentencing guidelines and commentaries. Because of the public nature and the significance of the interest being vindicated, appellate courts should apply a more searching standard of review (de novo or “hard look”) to the defendant’s score on the reprehensibility scale discussed in Section I.B. Since the determination is chiefly a legal one—whether the finding is consistent with the scale established by the legislature—appellate courts are institutionally competent to assert their interpretive authority. On the other hand, appellate court deference should extend to the findings of fact that serve as a predicate for the reprehensibility determination. Those underlying facts—e.g., the amount of money involved in the fraud or whether the defendant lied to the plaintiff—should be reviewed for clear error. Appellate court deference would also properly extend to the factual determination, if necessary, of the defendant’s financial condition or value, as well as determinations of reasonable lawyers’ fees or the defendant’s net gains exceeding compensatory damages. These are all fact-bound matters that invite deference to the initial factfinder’s relatively “local” knowledge.¹⁷⁷ Thus, a differentiated standard of review is required for the component parts of a retributive damages award.

By contrast, as mentioned earlier, aggravated and deterrence damages are really wrinkles associated with ensuring that the defendant internalizes the full costs of its activities to the extent consistent with *Philip Morris*. Those damage awards should be reviewed deferentially by appellate courts since they involve basically factual matters that juries or trial court judges are in an institutionally superior place to make vis-à-vis appellate courts. Here the standard might be something like the following: could a rational factfinder reasonably determine the amount of (aggravated/deterrence) damages that the lower court (or jury) found here?

3. Is a Unanimous Jury Required?

a. *Judges or Juries*

In most jurisdictions, juries are charged with determining the amount of punitive damages.¹⁷⁸ In a few states, such as Connecticut

¹⁷⁷ See generally Paul Horwitz, *Three Faces of Deference*, 83 NOTRE DAME L. REV. 1061, 1085-90 (2008) (analyzing epistemic superiority as the basis for deference).

¹⁷⁸ See Rustad, *supra* note 29, at 1305 (labeling as “radical” the idea of allowing judges to assess punitive damages).

and Kansas, juries decide whether to award punitive damages and judges decide the quantum of such damages in separate proceedings considering various aggravating and mitigating factors.¹⁷⁹ Some tort scholars have supported this division of labor as a rationalizing force, while others have disagreed because they view the jury's assessment of punitive damages as a historical legal right that facilitates a popular check on powerful and wealthy individuals or entities.¹⁸⁰

Truthfully, the account of retributive damages that I defend is less concerned with the identity of the decision maker than with the accuracy, modesty, and fairness of such decisions. My proposal comports with either juries or judges, or both, playing a role. From a Type I or Type II error-reduction perspective, there seems to be relatively little added value in having a jury decide the defendant's reprehensibility score, since that score is, by hypothesis, determined in large measure by reference to judgments that the polity has made regarding reprehensibility.

Having a jury involved, however, may have other benefits. First, the judgment of one's "peers" rendered on behalf of the state may enhance the probability that the defendant will internalize the condemnation issued through retributive damages. Second, in a legal culture where the jury's role is viewed as significantly involved in the meaning of punishment,¹⁸¹ a *defendant* may decide that making a case before a jury of one's peers is vitally important. These two considerations, of course, stand in some tension with each other. The first point of view suggests *requiring* that the jury determine punitive damages in order to "enrich" the social norm that retributive damages in-

¹⁷⁹ See *id.* at 1305-06.

¹⁸⁰ Compare David G. Owen, *Punitive Damages in Products Liability Litigation*, 74 MICH. L. REV. 1257, 1320-22 (1976) (suggesting that the trial judge, rather than the jury, should measure awards of punitive damages, in order to protect against excessive awards), with Partlett, *supra* note 38, at 1411-12 (defending a robust role for juries in punitive damages cases on the basis of republican theory), and Wenger & Hoffman, *supra* note 37, at 1148 (arguing that juries should continue to measure punitive damages awards).

¹⁸¹ See Michael T. Cahill, *Punishment Decisions at Conviction: Recognizing the Jury as Fault-Finder*, 2005 U. CHI. LEGAL F. 91, 96-108 (arguing that criminal juries already play a significant role, through their determinations of guilt or innocence, in the punishment process and that this role in punishment should be made more explicit); Richard E. Myers II, *Requiring a Jury Vote of Censure to Convict*, 88 N.C. L. REV. (forthcoming 2009) (proposing, as a means of reinforcing criminal law's condemnatory role, a requirement that juries vote to "censure" a defendant before convicting her).

volve a condemnation by one's peers on behalf of the state.¹⁸² The second point, however, entails giving the defendant a waivable right to a jury trial.

As noted in the previous Section, however, defendants do not enjoy an unfettered constitutional right to a jury trial in criminal cases. Though state constitutions or statutes might afford more rights, the federal constitutional floor is established when the offense is punishable by more than six months in prison or when the offense bears other indicia of being deemed constitutionally "serious."¹⁸³ In civil cases, a defendant in federal court has a right under the Seventh Amendment to a jury trial for cases involving issues that were triable under common law in 1791.¹⁸⁴ But the federal right is not incorporated against the states except in cases involving a federally created right.

That rule should leave states a good bit of discretion, constitutionally speaking. From a normative point of view, the retributive damages aspect of the extracompensatory damages is an intermediate civil punishment, so it is not unreasonable to err on the side of the defendant's preferences. This vector is also consistent with, though not entailed by, the normative justifications for the Supreme Court's holdings in *Apprendi v. New Jersey*¹⁸⁵ and *Blakely v. Washington*.¹⁸⁶ The retributive damages scheme contemplates a jury's making the decisions—with respect to the reprehensibility of the conduct and the financial condition of the defendant—that are critical to the maximum amount of penalty that the defendant will face. But those defendants who fear "grabby" juries with redistributive inclinations

¹⁸² This view might underlie the rule adopted in federal criminal cases that requires both the prosecution and the court to consent to a defendant's waiver of a jury trial. FED. R. CRIM. P. 23(a).

¹⁸³ *Blanton v. City of N. Las Vegas*, 489 U.S. 538, 541-42 (1989); *see also* *United States v. Nachtigal*, 507 U.S. 1, 4-5 (1993) (per curiam) (holding that an offense was not constitutionally "serious" even though it carried a penalty of up to six months' incarceration, a \$5,000 fine, a five-year term of probation as an alternative to incarceration, and other penalties).

¹⁸⁴ *See* *Chauffeurs, Teamsters and Helpers Local No. 391 v. Terry*, 494 U.S. 558, 573-74 (1990) (holding that the monetary damages sought were "the type of relief traditionally awarded by courts of law" and that the Seventh Amendment therefore created an entitlement to a jury trial).

¹⁸⁵ 530 U.S. 466, 490 (2000). The *Apprendi* Court held that, outside of a prior conviction, any factual finding that increases the statutory-maximum penalty for an offense must be proven beyond a reasonable doubt.

¹⁸⁶ 542 U.S. 296, 313 (2004). *Blakely* affirmed and extended *Apprendi*'s holding.

should be able to rely on judges for a (possibly more) impartial determination of their liability.

It goes without saying that it would be somewhat odd if a state extended a jury right to a defendant in a retributive damages case, but not to one in a criminal case with comparatively low stakes; nonetheless, that would be a permissible choice under the Supreme Court's jurisprudence.

b. *Unanimity*

According to some scholars, one of the safeguards that a publicly minded retributive damages sanction should entail is a right to a unanimous jury.¹⁸⁷ But this is an odd assertion. As is true of the right to counsel, or even a right to a jury at all, there is no unwavering constitutional commitment to a unanimous verdict in a criminal trial. In *Apodaca v. Oregon*¹⁸⁸ and *Johnson v. Louisiana*,¹⁸⁹ the Supreme Court permitted states to punish someone on the basis of 10-2 and 9-3 verdicts, respectively, in noncapital cases. The Court has also allowed smaller juries to decide other important criminal cases involving incarceration as a punishment.¹⁹⁰ Thus, there is no constitutional basis for assuming that impositions of retributive damages require a right to a jury verdict, let alone a right to a unanimous jury verdict.

From a normative perspective, rules requiring unanimity for acquittals are problematic for a variety of reasons and therefore should not be adopted.¹⁹¹ Unanimity rules for determinations of criminal guilt make more sense, much like the rationales for the beyond a reasonable doubt rule for the standard of proof. But for the intermediate sanction of retributive damages, a supermajority of the sort permitted in criminal cases by the Supreme Court would be an appropriate compromise between reducing Type I and Type II errors.

¹⁸⁷ See, e.g., Grass, *supra* note 7, at 243 n.13.

¹⁸⁸ 406 U.S. 404 (1972).

¹⁸⁹ 406 U.S. 356 (1972).

¹⁹⁰ See *Williams v. Florida*, 399 U.S. 78, 103 (1970) (holding that a six-person jury is sufficient). The Court did later clarify that, in cases involving nonpetty crimes, if a six-person jury were used, it would have to reach a unanimous verdict to convict. *Burch v. Louisiana*, 441 U.S. 130, 138 (1979).

¹⁹¹ See Ethan J. Leib, *Supermajoritarianism and the American Criminal Jury*, 33 HASTINGS CONST. L.Q. 141, 142-45 (2006) (arguing against "the obsession with unanimity" in the criminal justice system's decision rules).

4. Procedural Bifurcation

With respect to retributive damages, defendants in jury trials should have a waivable right to bifurcate the evidence associated with their financial condition from the evidence associated with liability. This rule would preserve the factfinder's ability to make a determination of the defendant's liability and reprehensibility without the "polluting" effects associated with evidence of one's financial condition. Conversely, during the liability phase, plaintiffs should be able to thwart the introduction by the defense of any evidence of their alleged penury to try to generate sympathy from the jury. These rules would facilitate the reduction of Type I errors without the concomitant cost of raising Type II errors.

Moreover, in some cases, defendants might wish to trifurcate the proceedings among (1) basic liability and compensatory damages, (2) liability for retributive damages, and (3) the defendant's financial condition. To my mind, there is nothing wrong with this division, but it may be unnecessary if the proof that shows the need for liability and compensatory damages also shows with clear and convincing evidence that the defendant acted maliciously or recklessly.

Courts might decide, upon a defendant's motion, to have the jury hear evidence for aggravated or deterrence damages at the same time as the proof offered for retributive damages, since, in many cases, evidence of malice is relevant to both retributive and aggravated damages and evidence of concealment is relevant to all three kinds of extracompensatory damages. And, of course, defendants can always stipulate to some matters and litigate others if that would resolve disputes more expeditiously.

5. Confrontation of Adverse Witnesses and Compulsory Process

As a general matter, one doesn't see too much litigation in punitive damages cases over the defendant's rights to confront adverse witnesses or to compel favorable witnesses to appear before the court. One reason is the language of the Sixth Amendment, which restricts these rights to "criminal prosecutions."¹⁹² But the other reason for this litigation void is that the defendant already has access to these witnesses in civil tort actions. With respect to Confrontation Clause issues, the plaintiff is usually the adverse witness against a punitive dam-

¹⁹² U.S. CONST. amend. VI.

ages defendant, and the plaintiff is typically quite keen to testify at trial. If the plaintiff testifies, the defendant has the ability to cross-examine.

Though it needs no further argument, the ability to defend against a sanction imposed by the state requires that a defendant have the right to present a defense on her behalf, both by examining the weaknesses of hostile testimony and bringing her own favorable witnesses and evidence to light. To the extent that this rule requires special articulation, defendants should be given the requisite assurances.

6. Access to Counsel

Civil defendants enjoy no constitutional right to appointed counsel, as the Supreme Court has bestowed that right only upon criminal defendants who have prison time imposed on them.¹⁹³ Because a defendant is not constitutionally entitled to counsel if she faces only a criminal fine, there is no precedential need to extend counsel to defendants facing the less severe sanction of retributive damages. But states and Congress are free to go above the floor provided by the U.S. Constitution. As with jury trials, however, it would be odd for a state to provide appointed counsel for retributive damages while not doing so for criminal cases involving only fines.

As a normative matter, whether to have a right to appointed counsel for retributive damages raises interesting questions. After all, there is a long-running dispute over whether defense lawyers protect innocent persons from mistaken punishment or simply make it easier for more offenders to escape condign punishment.¹⁹⁴ But the answer to this virtually irresolvable empirical question is not readily ascertained, in part out of a commitment to the idea that if a person was legally found not to have committed any illegal wrongdoing, then no adverse inferences ought to be drawn against her. The Supreme Court has long grappled with this trade-off in this and other contexts, though

¹⁹³ See *Scott v. Illinois*, 440 U.S. 367, 373-74 (1979) (holding that counsel is only required under the Constitution if the defendant had "actual imprisonment" imposed). For the most part, there is no constitutional right to appointed counsel in the civil context, but some states do provide appointed counsel in some civil contexts, such as where indigent parents face termination of their parental rights. *E.g.*, *In re K.L.J.*, 813 P.2d 276 (Alaska 1991).

¹⁹⁴ Compare, *e.g.*, *Gideon v. Wainwright*, 372 U.S. 335, 344 (1963) (describing contributions of lawyers to truth seeking), with WILLIAM TWining, *THEORIES OF EVIDENCE: BENTHAM AND WIGMORE* 75-79 (1985) (describing Bentham's view that lawyers impeded the factfinding mission of courts).

not altogether satisfactorily, often simply emphasizing the Blackstone-inspired preference for reducing errors involving false convictions over false acquittals.¹⁹⁵

Access to appointed counsel is now regarded primarily as an unalloyed good in our legal culture instead of the more mixed assessment extended to lawyers a couple centuries back, when their presence was viewed with greater suspicion and as an intrusion into the achievement of a just outcome.¹⁹⁶ If it turns out that the modern view is correct, and that lawyers are in fact a boon to truth, then the failure of the criminal justice system to provide counsel to *all* indigent defendants is one that should be *repaired*, not reproduced, in the context of retributive damages.

A more modest measure might involve legislatures authorizing and funding systems of indigent defense for defendants in retributive damages cases, but only after a court's staff attorney or judicial officer performs some screening of the plaintiff's complaint. This measure might be more consistent with the view that when there are trade-offs to be made, society may wish to consider the Laplace-Nozick principle that seeks to minimize the aggregate of the risk of false liability determinations (or in the criminal context, false determinations of guilt) and the risk of being a victim of serious wrongdoing.¹⁹⁷

A decidedly worse mode of reasoning would be to deny all defendants publicly funded counsel simply on the grounds that judgment-proof defendants can always be prosecuted in a criminal court. This argument problematically raises concerns that the poverty of a defendant provides a reason to bring the power of the state and the con-

¹⁹⁵ See 4 WILLIAM BLACKSTONE, COMMENTARIES *352 ("[I]t is better that ten guilty persons escape, than that one innocent suffer."); cf. *In re Winship*, 397 U.S. 358, 372 (1970) (Harlan, J., concurring) (relying on the "fundamental value determination of our society that it is far worse to convict an innocent man than to let a guilty man go free").

¹⁹⁶ See, e.g., TWINING, *supra* note 194, at 75 (identifying, as one of Bentham's primary concerns, "the sinister interests of the legal fraternity").

¹⁹⁷ See Larry Laudan, The Social Contract and the Rules of Trial: Re-Thinking Procedural Rules 31-43 (2008) (unpublished manuscript), available at <http://ssrn.com/abstract=1075403> (noting that the social contract requires the state to reduce aggregate risk under the Laplace-Nozick thesis). But see Posting of Dan Markel to PrawfsBlawg, "Legal Epistemology Is Ninety Per Cent Quantitative. The Other Half Is Qualitative," <http://prawfsblawg.blogspot.com/2008/08/legal-epistemol.html> (Aug. 10, 2008) (urging greater granularity in the comparisons that Laudan makes and more mindfulness toward the potentially unjust distributive patterns associated with who bears the consequences of tradeoffs between Type I and Type II error-reduction strategies).

comitant collateral consequences of convictions against that defendant, whereas a wealthier defendant might be able to persuade a prosecutor to forbear from subsequent prosecution by pointing to a prior intermediate sanction through retributive damages.

As a practical matter, failure to provide counsel to indigent defendants facing retributive damages may not arouse any sense of urgency. After all, a civil defendant who cannot afford counsel often lacks the deep pockets that motivate a plaintiff to seek substantial compensatory damages. Moreover, if a defendant is poor and uninsured, the plaintiff will often have difficulty finding counsel to bring the case, which makes the problem even less significant.¹⁹⁸ That said, counsel may decide to take the case against poor defendants if there are provisions for reasonable attorneys' fees provided by the state to subsidize access to judgments of retributive damages against defendants of all sizes. Alternatively, as described in my earlier companion article, the legislature may decide to limit access to retributive damages by insisting on a minimum amount in controversy or by making the loser pay the costs and fees of the opposing side's counsel. Both requirements would screen out many weak or low-value cases brought by pro se or vexatious plaintiffs.¹⁹⁹

7. Privilege Against Self-Incrimination

The Fifth Amendment's privilege against self-incrimination states that no one "shall be compelled in any criminal case to be a witness against himself."²⁰⁰ The privilege permits an individual (but not a corporation) to refuse to answer any questions put to her during any proceeding if she in good faith believes that the testimony will either "support a conviction" against her or "furnish a link in the chain of evidence" against her.²⁰¹ A person can invoke the privilege during any kind of proceeding—including in a civil trial for punitive damages

¹⁹⁸ See generally Tom Baker, *Blood Money, New Money, and the Moral Economy of Tort Law in Action*, 35 LAW & SOC'Y REV. 275 (2001) (analyzing the incentives of plaintiffs' lawyers in tort cases).

¹⁹⁹ See Markel, *Retributive Damages*, *supra* note 1, at 297-300 (suggesting various permutations for institutional design that would reduce Type I and Type II errors).

²⁰⁰ U.S. CONST. amend. V.

²⁰¹ *Hoffman v. United States*, 341 U.S. 479, 486 (1951); see also *Carlson v. United States*, 209 F.2d 209, 214 (1st Cir. 1954) (stating that a person invoking the privilege in bad faith may be found guilty of perjury).

and in the discovery process of such a trial.²⁰² Moreover, the Supreme Court has held that if persons feared that their statements would trigger some substantial sanctions aside from fines or imprisonment, then that too would suffice to permit the privilege to be invoked.²⁰³

Nonetheless, the Court has also limited the reach of what counts as “coerced” self-incrimination by stating that not all civil penalties can trigger the privilege—in other words, merely facing sanction has not been deemed a sufficient basis to refuse to answer questions.²⁰⁴ Moreover, in civil cases where the privilege is invoked, the factfinder is permitted to draw an adverse inference against the person.²⁰⁵ Additionally, a person invoking the privilege during any civil proceeding who is later granted immunity from the sanction in question could be required to testify.²⁰⁶

The scope of the privilege has long been a source of controversy and puzzlement.²⁰⁷ The text of the Fifth Amendment indicates that the privilege is “expressly limited to ‘any criminal case.’”²⁰⁸ Furthermore, the Supreme Court has never held that fear of punitive damages alone, despite their “quasi-criminal” nature, permits an invocation of the privilege.

Because of the lack of consensus regarding the privilege’s rationale, it is hard to understand what its proper scope should be as a constitutional matter. And unlike the standard-of-proof issue, there is no

²⁰² See *Lefkowitz v. Turley*, 414 U.S. 70, 77 (1973) (“The [Fifth] Amendment . . . privileges [the individual] not to answer official questions put to him in any other proceeding, civil or criminal, formal or informal . . .”).

²⁰³ See *Spevack v. Klein*, 385 U.S. 511, 515-16 (1967) (finding a threat of disbarment sufficient to invoke privilege).

²⁰⁴ See *United States v. Ward*, 448 U.S. 242, 251-55 (1980) (finding that in establishing a civil penalty, Congress did not impose a scheme that was so punitive as to allow invocation of the privilege).

²⁰⁵ See *Baxter v. Palmigiano*, 425 U.S. 308, 318 (1976) (declaring “the prevailing rule” to be that “the Fifth Amendment does not forbid adverse inferences against parties to civil actions when they refuse to testify in response to probative evidence offered against them”).

²⁰⁶ See, e.g., *Hale v. Henkel*, 201 U.S. 43, 67 (1906); see also 18 U.S.C. § 6003 (2006) (providing the procedure for granting immunity).

²⁰⁷ See Ronald J. Allen, *Theorizing About Self-Incrimination*, 30 CARDOZO L. REV. 729, 730-39, 734 n.25 (2008) (canvassing much of the literature and demonstrating why most efforts to explain or justify the privilege against self-incrimination are “uniformly unconvincing”); William J. Stuntz, *Self-Incrimination and Excuse*, 88 COLUM. L. REV. 1227, 1228 (1988) (“It is probably fair to say that most people familiar with the doctrine surrounding the privilege against self-incrimination believe that it cannot be squared with any rational theory.”).

²⁰⁸ *Ward*, 448 U.S. at 248 (quoting U.S. CONST. amend. V).

obvious middle ground appropriate for an intermediate sanction, so we must muddle through a different way. From the normative perspective of whether a jurisdiction should extend the privilege to defendants fearing retributive damages liability, we have to consider whether the privilege advances or hinders retributive justice interests. If the privilege is anti-retributive, then that is a good basis for limiting the safeguard generally and permitting, but not requiring, an adverse inference from any defendant's silence. If the privilege is pro-retributive, then that is a good basis for extending the safeguard to a defendant facing retributive damages and not allowing the factfinder to make an adverse inference. Turning to that precise issue, we need a sense of how the privilege affects the incidence of Type I and Type II errors.²⁰⁹

Unfortunately, it is hard to reach a conclusive determination. The primary argument against the privilege is that it deprives the court and the jury of the accused party's testimony when that party is often in the best position to verify or credibly deny the accusations made against her. As Jeremy Bentham wrote, "Evidence is the basis of justice: exclude evidence, you exclude justice."²¹⁰ Thus, it is often said that having the privilege works to the advantage of the guilty who are not subjected to the perils of cross-examination.²¹¹

In contrast, there are two arguments that the privilege works to reduce Type I errors of mistaken punishment. First, as the Supreme Court stated in *Wilson v. United States*, not every innocent person, "however honest," can avoid "nervousness when facing others and attempting to explain transactions of a suspicious character," and thus, even a person who is innocent may nonetheless be worried that his testimony will "increase rather than remove prejudices against him."²¹²

A second argument looks at how the privilege works through the signaling effects associated with its use. According to Professors Seidmann and Stein, the privilege actually works in favor of the inno-

²⁰⁹ To be sure, there are other values important to retributivism, but I will limit my discussion to these core issues.

²¹⁰ 5 JEREMY BENTHAM, *RATIONALE OF JUDICIAL EVIDENCE* 1 (Littleton, Colo., Fred B. Rothman & Co. 1995) (1827).

²¹¹ For a particularly incisive analysis of this point, see Allen, *supra* note 207, at 734-36.

²¹² 149 U.S. 60, 66 (1893).

cent and against the self-interest of the guilty.²¹³ This counterintuitive argument rests on the dynamic effects associated with the privilege. The basic idea is that innocent people will not invoke the privilege because they have nothing to hide and are willing to share their story and be subject to cross-examination. The factually guilty will not testify because they will be subject to cross-examination and will be less able to corroborate their stories. Because of this common intuition, juries will tend to credit the stories of people who testify more than the stories of those who do not. Defendants invoking the privilege will provoke the suspicion of the juries and will likely be found guilty. In this respect, one version of the argument is that the privilege will work as a sorting mechanism between the guilty and the innocent.

But the dynamic-effects story needs revision for two separate reasons. First, if guilty persons start testifying to exploit to their advantage the received wisdom that only the innocent testify, then the norm that is supposed to work to the benefit of the innocent becomes corrupted. More iterations of the game, in other words, will lead to a pooling of guilty and innocent people testifying and a situation in which the only people invoking the privilege are the guilty “suckers” or the innocent persons worried that they will be wrongfully lumped with the guilty “exploiters” who testify. This quickly turns into a guessing game of how large the number of “exploiters” is within any pool of testifying defendants. Of course, there are two checks on this pooling problem. First, the guilty persons who subject themselves to cross-examination will have a harder time in individual cases proving to the jury that they are credible because there is likely to be some doubt sown in their minds by an effective lawyer’s cross-examination. Second, the jurors are not repeat players and thus are less likely to view the game associated with the privilege as a game involving repeat players, the dynamic effects of which are to be scrutinized.

But there are also empirical reasons to be skeptical of the Seidmann-Stein theory. As Professor Bibas points out, in real life, most suspects will, and do, talk to the police upon arrest because they have very good reasons to worry more about the police drawing adverse inferences from their silence than about the adverse inferences that ju-

²¹³ See Daniel J. Seidmann & Alex Stein, *The Right to Silence Helps the Innocent: A Game-Theoretic Analysis of the Fifth Amendment Privilege*, 114 HARV. L. REV. 430, 430 (2000) (“Because the right to silence is available, innocent defendants still tell the truth while guilty defendants may rationally exercise the right. Thus, guilty defendants do not pool with innocent defendants by lying, and as a result, triers of fact do not wrongfully convict innocent defendants.”).

ries will draw at trial.²¹⁴ As Bibas concludes, notwithstanding the apparent elegance of the Seidmann-Stein theory, the privilege only helps the guilty, not the innocent.²¹⁵

If Professor Bibas is correct, then we cannot rely on the privilege to work in favor of securing the reduction of Type I errors; this seems especially important when the privilege appears to increase Type II errors of false acquittals or other nonpunishments of those who are factually guilty. As many have argued, there is good reason to believe that an orderly inquiry into a defendant's actions does not itself offend basic norms of dignity.²¹⁶ At the same time, such an inquiry permits the factfinder to draw inferences of truth or falsehood. Of course, that does not mean that there aren't good reasons for separately prohibiting statements when the surrounding circumstances cast doubt on their voluntariness. But if the privilege does serve a truth-impeding role, then that alone is a good reason for not extending it to allow the fear of retributive damages to serve as a basis for its invocation. Moreover, even if extended, the privilege should permit, but not require, factfinders to draw an adverse inference against the defendant who, in a civil torts proceeding, invokes the privilege.²¹⁷

Of course, the overlap between conduct that renders one eligible for retributive damages and conduct that renders one eligible for criminal punishment is substantial. Consequently, we also need to have an approach for defendants who do not want to testify in civil proceedings lest they incriminate themselves for criminal prosecution purposes. This problem is not unique to the retributive damages context, so we can easily reference the current strategy: namely, that defendants can seek a stay of civil proceedings pending the resolution of a criminal prosecution and force the government to prosecute them first, or, alternatively, defendants can seek immunity from

²¹⁴ See Stephanos Bibas, *The Right to Remain Silent Helps Only the Guilty*, 88 IOWA L. REV. 421, 421 (2003) ("Though [the Seidmann-Stein] theory predicts that rational suspects remain silent, roughly *eighty to ninety percent of suspects talk to the police*").

²¹⁵ See *id.* at 432 (casting doubt on the Seidmann-Stein theory by exposing as flawed the assumptions on which the theory is based and by identifying the factors that it overlooks). But see Alex Stein, *The Right to Silence Helps the Innocent: A Response to Critics*, 30 CARDOZO L. REV. 1115 (2008).

²¹⁶ See *Palko v. Connecticut*, 302 U.S. 319, 326 (1937) (Cardozo, J.) ("Justice . . . would not perish if the accused were subject to a duty to respond to orderly inquiry.").

²¹⁷ There is obviously much more to be said on the policies underlying the privilege embedded in the Fifth Amendment. A good starting point would be the recent symposium in the *Cardozo Law Review*. See Symposium, *The Future of Self-Incrimination: Fifth Amendment, Confessions, & Guilty Pleas*, 30 CARDOZO L. REV. 717 (2008).

criminal prosecutions that rely on the testimony as links in the chain of evidence.

8. Duplicative Punishment Concerns in Simple Litigation Contexts

This subsection considers the relationship that retributive damages awards should have relative to prior or subsequent sanctions through the criminal justice system. In effect, this discussion largely addresses many of the basic questions associated with double jeopardy law and policy in the *simple litigation* context—where *X* commits *Y* misconduct and in so doing injures *Z* and only *Z* in a manner that would be eligible for both retributive damages and criminal prosecution in and only in jurisdiction *Q*.²¹⁸ The following discussion explains why (1) a defendant can face retributive damages and then criminal proceedings for the same misconduct; (2) a defendant who is acquitted in criminal court should under some circumstances be subject to a subsequent retributive damages proceeding; and (3) a defendant who is convicted in criminal court should be able to block subsequent retributive damages predicated on the same conduct toward the same victim. I also address parallel proceedings in civil and criminal cases.

a. *Double Jeopardy Doctrine and Punitive Damages*

The Double Jeopardy Clause of the Fifth Amendment works to prevent three separate events: “a second prosecution for the same offense after acquittal; a second prosecution for the same offense after conviction; and multiple punishments for the same offense.”²¹⁹ As a matter of current law, courts have held that earlier punitive damages determinations do not serve as “prior punishment” such that a subsequent criminal prosecution for the same misconduct to the same victim would be precluded under the double jeopardy provision of the Fifth Amendment.²²⁰ Various courts have similarly held that criminal

²¹⁸ In the successor article, *supra* note 9, I will address, among other things, concerns of duplicative punishment arising from multiple injuries inflicted by the same course of conduct.

²¹⁹ *United States v. Halper*, 490 U.S. 435, 440 (1989), *overruled on other grounds by Hudson v. United States*, 522 U.S. 93, 95-96 (1997). The Court has also made clear that notwithstanding the text of the Fifth Amendment’s reference to “life or limb,” it also covers monetary penalties and incarceration. *See Dep’t of Revenue v. Kurth Ranch*, 511 U.S. 767, 769 n.1 (1994) (“[I]t is well settled that the [Fifth] Amendment covers imprisonment and monetary penalties.”).

²²⁰ *See, e.g., Helvering v. Mitchell*, 303 U.S. 391, 399 (1938) (holding that a legislature can “impose both a criminal and a civil sanction in respect to the same act or

finer or other criminal punishments for certain misconduct would not thwart subsequent recoveries of punitive damages based on the same conduct.²²¹ Moreover, a defendant who is acquitted in criminal court may still face subsequent punitive damages or civil penalties in civil court.²²² That is because the courts have focused on the fact that the Double Jeopardy Clause does not apply to litigation between private parties.²²³

That rationale, however, only goes so far. In some cases, the government's prosecution of a defendant has been permitted even after the government sued initially on its own (quasi-private) behalf and recovered punitive damages for the underlying misconduct.²²⁴ By contrast, when the government convicted someone on drug charges and then later sought to levy a very heavy punitive "tax" on the defendant for the drug conduct, the government's subsequent tax was invalidated.²²⁵ Critics of punitive damages without procedural safeguards could reasonably point to these double jeopardy distinctions as evidence of arid formalism—in both situations, the government is going to capture the gains associated with the civil penalties and criminal

omission; for the double jeopardy clause prohibits merely punishing twice, or attempting a second time to punish criminally, for the same offense"); *United States v. Ely*, 142 F.3d 1113, 1121 (9th Cir. 1997) (finding that a prior recovery of punitive damages by the FDIC against bank directors for federal bank fraud did not preclude subsequent criminal prosecution); *Hansen v. Johns-Manville Prods. Corp.*, 734 F.2d 1036, 1042 (5th Cir. 1984) (holding that because punitive damages awards are not criminal sanctions, multiple awards of punitive damages are consistent with the Double Jeopardy Clause of the Fifth Amendment); *cf. Hudson*, 522 U.S. at 99 (holding that double jeopardy only protects against "multiple *criminal* punishments for the same offense").

²²¹ See, e.g., *Shore v. Gurnett*, 18 Cal. Rptr. 3d 583, 586-87 (Ct. App. 2004) (finding that double jeopardy did not apply to punitive damages awarded in a wrongful-death suit following the conviction of the defendant for vehicular manslaughter).

²²² See *State v. McDowell*, 699 A.2d 987, 989 (Conn. 1997) ("Jeopardy attaches only to proceedings which are 'essentially criminal.'" (quoting *Breed v. Jones*, 421 U.S. 519, 528 (1975))).

²²³ See *Halper*, 490 U.S. at 451 ("The protections of the Double Jeopardy Clause are not triggered by litigation between private parties."); see also *Hudson*, 522 U.S. at 110-11 (Stevens, J., concurring) (observing that the purpose of the Double Jeopardy Clause is to prevent the State from "mak[ing] repeated attempts to convict an individual").

²²⁴ See *Ely*, 142 F.3d at 1121; *United States v. Beszborn*, 21 F.3d 62, 67-68 (5th Cir. 1994) (allowing a federal agency first to pursue punitive damages against defendants and then to pursue criminal indictments).

²²⁵ See *Dep't of Revenue v. Kurth Ranch*, 511 U.S. 767, 784 (1994) ("This drug tax is not the kind of remedial sanction that may follow the first punishment of a criminal offense. Instead, it is a second punishment within the contemplation of a constitutional protection . . .").

finer. More troubling, the current doctrine creates the conditions for substantial overpunishment from a retributive perspective.

b. *The Duplicative Punishment Problem*

Because of the lack of double jeopardy protection afforded to the status of punitive damages, scholars, including Colby and Wheeler, have endorsed the idea that a publicly minded punitive damages sanction should trigger double jeopardy protection.²²⁶ Their concern, like that of numerous courts and legislatures around the country, is that multiple punishments or successive prosecutions could lead to overkill.²²⁷

From a retributive damages perspective, such overkill is a matter of substantial and legitimate concern. But even if defendants enjoyed full “double jeopardy” protection, that would hardly serve as substantial relief to the overkill concern. As explained immediately below, American double jeopardy law does little to protect criminal defendants from duplicative prosecutions and punishments. As a result, extending to civil defendants the same protections available to criminal defendants might entail disappointing results. Let me explain quickly why that is the case and then suggest what should be done.

The federal constitutional floor for double jeopardy law is set very low. Specifically, under the Supreme Court’s *Blockburger* test, a criminal defendant can be tried for a crime that has elements 1, 2, and 3 and can then be subsequently tried for a crime that has elements 2, 3, and 4, even if the underlying criminal transaction was exactly the same event in time.²²⁸ Moreover, although some states offer greater protection through statute or state constitutions, the majority of states provide no more protection than that offered by *Blockburger*.²²⁹ And because of the often sprawling and redundant features of many

²²⁶ See Colby, *supra* note 5, at 452-53 (noting that punitive damages only call for criminal procedural protection when they are understood as punishment for public, not private, wrongs); Wheeler, *supra* note 7, at 272 (concluding that the procedural safeguards in the Fourth, Fifth, and Sixth Amendments could apply to punitive damages where the action is criminal in substance).

²²⁷ See, e.g., *Roginsky v. Richardson-Merrell, Inc.*, 378 F.2d 832, 841 (2d Cir. 1967) (Friendly, J.) (discussing the danger of overkill from multiple punishments).

²²⁸ See *Blockburger v. United States*, 284 U.S. 299, 304 (1932) (“[T]he test to be applied to determine whether there are two offenses or only one, is whether each provision requires proof of a fact which the other does not.”).

²²⁹ See MARC L. MILLER & RONALD F. WRIGHT, CRIMINAL PROCEDURES 1002-05 (3d ed. 2007) (providing a survey of double jeopardy doctrines around the states).

American criminal codes, the Double Jeopardy Clause itself does not provide a powerful shield against the prospect of overkill.²³⁰

The problem of overkill is potentially compounded further by the federal constitutional standard because of the dual sovereigns doctrine, under which a defendant who is acquitted or convicted of misconduct in criminal court in one jurisdiction might still be subsequently pursued and convicted for the same misconduct in criminal court if the misconduct occurred in overlapping jurisdictions.²³¹ In other words, absent legislative intervention to the contrary, a punitive damages suit in one jurisdiction will be legally irrelevant to whether an alternative jurisdiction may host a prosecution or civil suit against a defendant for the same misconduct. This is true even when the misconduct at issue affects only one victim.²³²

Because double jeopardy law provides very little real protection to criminal defendants, there is, in reality, very little disparity between the civil and the criminal defendant in this regard: both civil and criminal defendants are actually exposed to risks of overkill. This suggests that either the double jeopardy provisions in the criminal context need strengthening via statute or judicial rule, as some states have done,²³³ or that civil defendants facing retributive damages in multiple jurisdictions are, in many respects, in the same helpless position as criminal defendants. To the extent that one is comfortable with current double jeopardy jurisprudence in criminal cases, one should not fear that civil defendants are being radically mistreated as compared to criminal defendants.²³⁴

²³⁰ Cf. Paul H. Robinson & Michael T. Cahill, *The Accelerating Degradation of American Criminal Codes*, 56 HASTINGS L.J. 633, 635 (2005) (lamenting "legislative hyperactivity" culminating in "unnecessary and often inconsistent" statutes).

²³¹ See *Bartkus v. Illinois*, 359 U.S. 121, 132-33 (1959) (holding that state and federal prosecutions based on the same actions do not violate the Fifth Amendment); *United States v. Lanza*, 260 U.S. 377, 382 (1922) ("[A]n act denounced as a crime by both national and state sovereignties is an offense against the peace and dignity of both and may be punished by each.").

²³² See *Heath v. Alabama*, 474 U.S. 82, 88 (1985) (applying the dual sovereigns doctrine to permit prosecution in two states for the murder of a victim who was kidnapped in one state and killed in another).

²³³ See MILLER & WRIGHT, *supra* note 229, at 1002-05 (discussing the various ways that states define double jeopardy provisions).

²³⁴ Galanter and Luban recognize that multiple punitive damages awards are "extremely troublesome" from their perspective because "if each award has been appropriately scaled to the heinousness of the deed, multiple awards amount to overpunishment." Galanter & Luban, *supra* note 3, at 1455 n.302. Nonetheless, Galanter and

The preceding discussion should inform but not constrain the normative vision of what is an appropriate level of protection—from a retributive justice perspective—for defendants facing accusations stemming from the same misconduct in both civil and criminal courts. I turn to that issue next.

c. Solving the Overkill Problem in the Simple Context

Because of the dynamic and complex interaction between the civil and criminal justice systems, the problem of overkill requires legislative attention, even in the simple litigation context where there is just one wrongful injury caused by one person to another. This subsection explains how, under the retributive damages regime, defendants would get more protection than they do under current law. Of course, this discussion assumes that the misconduct in question is subject to both criminal sanction and retributive damages; legislatures obviously can choose to restrict some conduct to separate spheres.

i. Should Retributive Damages Follow Criminal Prosecution?

For the reasons explained below, retributive damages generally should not be allowed to follow criminal convictions imposed in the simple litigation context for the same misconduct against a particular victim. Whether they ought to be permitted subsequent to acquittals in criminal trials should depend on the availability of new evidence in the new suit. However, if a plaintiff in a separate sovereign sought retributive damages against a defendant following a criminal prosecution in a separate jurisdiction, there would be no difficulty on my view with that separate proceeding. The same conclusion of permitting subsequent retributive damages suits would apply even within the same jurisdiction if the defendant perpetrated multiple counts of the same misconduct.

Under the constitutional rule described earlier, defendants enjoy no double jeopardy protection for punitive damages vis-à-vis earlier criminal prosecutions. In other words, an offender does not generally deduct the criminal fine paid from the amount of subsequent punitive damages. This makes sense when the subsequent tort suit achieves a compensation function not already achieved by restitution.

Luban fail to address how punitive damages might also raise similar concerns of unfairness vis-à-vis the criminal justice system.

But when the government establishes an extracompensatory damages scheme in part to effectuate retributive justice, the fact of the prior criminal penalty and where it was imposed together warrant greater attention when fashioning a retributive damages award. On the view that retributive damages awards are intermediate sanctions *less* severe than a criminal sanction, the criminal penalty in a particular jurisdiction must presumptively be viewed both as expressing to the public sufficient condemnation by the state of the offender's misconduct and as communicating that condemnation to the offender. Thus, allowing a retributive damages award against *X* to supplement that criminal penalty would be duplicative where it is the same wrong *Y* and the same victim *Z* involved in the same jurisdiction *Q*. While that postconviction retributive damages award should be forbidden, the victim should still be able to pursue compensatory damages (and aggravated or deterrence damages if applicable).

The questions are more normatively complex when a defendant is acquitted in the criminal prosecution and then faces a retributive damages lawsuit in the same jurisdiction. On the one hand, the fact that the government already sought and lost an opportunity for retributive punishment might suggest that it is time to let sleeping dogs lie rather than risk another Type I error. On the other hand, retributive damages are meant to be an intermediate sanction with a lower standard of proof, and removing the availability of retributive damages might mean more Type II errors that society wants to avoid, in which case it would be permissible to allow a plaintiff (whether a victim or a private attorney general) to pursue retributive damages.²³⁵ My sense is that if the subsequent retributive damages action is predicated on the exact same evidence that the government adduced in its prosecution of the defendant, then retributive damages should not be permitted. But if there is a reason to think the private litigation efforts add new and important evidence not earlier adduced during the criminal proceedings, then that would be a good reason for permitting subsequent retributive damages claims.

By contrast, where the retributive damages award is levied by a different sovereign from the one that imposed the earlier criminal punishment, then the defendant has offended several sovereigns and she

²³⁵ See Markel, *Retributive Damages*, *supra* note 1, at 280-86 (discussing the private attorney general (PAG) structure). One might fairly question whether the PAG benefits are as strong in the context of a retributive damages action following a defendant who was acquitted in a criminal case.

is liable for having wronged two (or more) sovereigns separately through the same misconduct. To be sure, the defendant may wonder why she is being penalized two (or more) times by different sovereigns for the same misconduct. But on the assumption that punishment only occurs for avoidable and culpable conduct, the defendant is on notice that each jurisdiction may want to vindicate its interest in punishing the particular breach against its own legal order.²³⁶ Moreover, if only one jurisdiction were permitted to prosecute the misconduct, there would have to be some rather arbitrary mechanism by which it is decided who gets to vindicate that particular legal interest. In light of the facts that the dual sovereigns doctrine (1) promotes comity (by avoiding conflicts) in both horizontal (state-state) and vertical (federal-state) federalism contexts, and (2) reflects the status quo in the criminal and civil contexts today, there is not much reason to think that the same notion should not apply with respect to the intermediate sanction of retributive damages.²³⁷ Thus, a criminal conviction in another jurisdiction would not preclude the imposition of retributive damages in a second jurisdiction that could also criminally punish the conduct but had not done so through criminal proceedings.

Moreover, even in a jurisdiction that had prosecuted a person for particular conduct, if that conduct injured more than one person, then the same jurisdiction could impose retributive damages for wrongs to those individuals not earlier treated as victims in the criminal trial. In other words, retributive damages would also be available subsequent to a criminal conviction to punish a defendant for misconduct affecting a victim who was not the focus of the initial criminal penalty. That additional penalty of retributive damages is reasonably imposed because, from the criminal justice system's perspective, the misconduct against a second victim is simply another "count" of the same underlying misconduct or is separate misconduct altogether. And when multiple counts of the same crime or different crimes af-

²³⁶ I leave aside for now concerns that persons might reasonably raise about being punished in polities with universal jurisdiction over certain forms of extreme wrongdoing.

²³⁷ This position is consistent with, and expressive of, the institutional account of retributivism discussed in Part II of Markel, *Retributive Damages*, *supra* note 1, which acknowledges that the same conduct can be subject to sanction in several jurisdictions, even consecutively. A pure "moral" retributivist might think that punishments for the same conduct that separately are proportional would in the aggregate be an unfair "piling on" because they offend a prepolitical conception of desert. For some discussion of the shortcomings of a purely moral, as opposed to political, account of retributivism, see Markel, *Against Mercy*, *supra* note 52, at 1451-53.

fecting different victims occur, jurisdictions commit no wrong when they set punishment (whether fines or incarceration) for those counts cumulatively or consecutively. Indeed, a state that groups penalties for all coterminous crimes might appear to be punishing crimes at a “buy one, get one/two/three free” discount. How bad that message is must be weighed against other competing values; and, of course, consecutive punishments for different counts are also consistent with lower sentences across the board for each count.²³⁸

ii. Should Criminal Prosecution Follow Retributive Damages?

In light of the preceding discussion, we should also consider whether retributive damages awards, if reached first, should have a preclusive effect on subsequent criminal prosecution and punishment. Would it be justifiable for the government to prosecute an offender after evidence in a retributive damages case indicated that there was a sufficient basis for prosecution under a beyond a reasonable doubt standard? In my view, such subsequent criminal punishment is justifiable, assuming that the tortious misconduct was separately subject to a criminal law sanction.

As I explained in *Retributive Damages*, the retributive damages sanction is meant to be an *intermediate* sanction, such that if proven, it *invites* prosecutors to *consider* filing subsequent criminal prosecutions.²³⁹ Indeed, one reason to cross-fertilize retributive justice with the civil tort system is to harness the power of private parties to ferret out and bring to public attention the defendant’s misconduct.²⁴⁰ Thus, under this proposal, there would be no protection against subsequent prosecution afforded to a defendant first found liable for retributive dam-

²³⁸ As I explain in Markel, *Punitive Damages*, *supra* note 9, defendants worried about the dangers of seriatim punishment by different plaintiffs should be able to obtain a defensive class action that would aggregate all plaintiffs for retributive damages purposes arising from specific misconduct. *See generally* Exxon Shipping Co. v. Baker, 128 S. Ct. 2605, 2611 (2008) (dealing with the aggregate claims for punitive damages of the plaintiff and “others, including commercial fisherman and Native Alaskans”). These aggregative strategies would ultimately offer more protection than criminal defendants receive now because if a defendant commits a crime with a jurisdictional nexus available to twenty states, she can be punished separately and consecutively twenty times.

²³⁹ Markel, *Retributive Damages*, *supra* note 1, at 320.

²⁴⁰ Cf. Peter J. Boyer, *The Bribe*, NEW YORKER, May 19, 2008, at 44, 54 (describing how employees of State Farm Insurance, suspicious of the company’s fraudulent practices, shared inculpatory information with private plaintiffs’ lawyers who then passed that information on to law enforcement officers).

ages. If the government sees damning evidence during the retributive damages proceeding and decides to prosecute criminal actions based on that misconduct, it should be able to do so. After all, if the government wants more severe sanctions and more severe condemnation, it can only seek those ends through affording the defendant more stringent procedural safeguards. Under this proposal, however, the amounts paid in retributive damages in the civil proceeding should be deducted from the total of any subsequent monetary penalties imposed on the defendant for proven criminal misconduct that is identical to the misconduct established in the earlier civil proceedings.

But just as the state may seek separate punishments for separate wrongs against separate persons stemming from the same misconduct,²⁴¹ so too should prior retributive damages awards not be credited against subsequent damage awards for the same misconduct when that misconduct has affected people who were not plaintiffs in the earlier retributive damages suit. Indeed, this is the flip side of the Supreme Court's decision in *Philip Morris* and is also a free-standing principle.

To be sure, there are certain disadvantages that attach to making a defendant go through two trials for the same misconduct. A defendant convinced of her innocence will reveal many, if not all, of her cards at the initial retributive damages tribunal. The government effectively gets a preview of this defense. From a retributivist perspective, however, this may have the effect of reducing both Type I and Type II errors over time. When the government sees the defendant's case in a retributive damages trial, the government may conclude that the defendant's case is strong enough that a jury will not convict under the higher standard of proof (and greater number of procedural safeguards) that attaches to criminal defendants.

One might be concerned about the expense that a defendant must incur in trying to defend her misconduct a second time in a criminal trial. But since we cannot reasonably ban civil suits just because the government might stake a criminal claim against a defendant, this problem of civil and criminal trials is bound to persist regardless of whether a retributive damages scheme exists. So the additional expense is only the *marginal difference* between litigating a

²⁴¹ See, e.g., Ann M. Simmons, *Metrolink Killer Is Sentenced to 11 Life Terms in Prison*, L.A. TIMES, Aug. 21, 2008, at B5 (reporting on a defendant who faced eleven consecutive counts of incarceration for the murders of eleven people caused by one action).

case for compensatory damages and litigating one that also permits retributive damages.

Moreover, if there were no records of civil trials available from which the prosecutor could examine the case against the defendant, there likely would be more crippling criminal cases brought against defendants. Under the retributive damages scheme described here, a defendant may be more likely to persuade a prosecutor to forbear from pursuing criminal proceedings. Alternatively, the defendant can ask for a stay of the civil proceedings involving retributive damages, pending resolution of criminal proceedings, an issue that I take up next.

iii. Can Retributive Damages Coincide with Criminal Prosecution?

What if the government has indicted a defendant, thus signaling its intent to prosecute *X* defendant for *Y* misconduct against *Z* victim, but a private action seeking retributive damages is subsequently filed prior to conviction in the criminal case? In other words, what happens if there are parallel civil and criminal proceedings? Typically, if the government has already publicly committed its resources to the prosecution of *X* for the same *Y* misconduct against *Z*, the defendant can seek to stay the civil proceedings pending the outcome of the criminal trial.²⁴² That rule makes sense here, in the context of the retributive damages scheme. In that case, if the government wins, there would be no subsequent retributive damages brought for *Y* misconduct against *Z*.

The parallel-proceedings problem can also arise if *Z* files a retributive damages action against *X* before the government's indictment. If the government can just file an indictment midway through the retributive damages case, then the retributive damages "payments" to the plaintiff and her lawyers will not be a very satisfactory incentive to develop and pursue certain kinds of claims or information. What should happen then is that the government should either wait until after the retributive damages action is over, or, alternatively, the government should be able to buy the retributive damages claim from the plaintiff by paying the plaintiff her reward and fairly compensating the lawyers for their time and investment. The defendant can then ask to stay the civil proceeding while the criminal prosecution proceeds. Thereafter, *Z* and her lawyers should be able to vindicate their

²⁴² The advantage to the state gained by making *X* litigate in two fora simultaneously is an unfair one, achieved only through the chance of having the wrongdoer distracted by two simultaneous proceedings.

interests in compensatory, aggravated, or deterrence damages at a point that suits them after the criminal proceedings. And in cases involving a defendant with limited resources, the state will need to decide whether the compensation function of tort law should trump the social interest in retributive justice. My sense is that the precise dimensions of such a trade-off between competing moral imperatives are best left to politics to decide.

Finally, for all these scenarios, one must also decide which test to adopt to determine whether the subsequent litigation is in fact based on the same misconduct. Because of the large number of crimes and torts, the *Blockburger* test described earlier will not do much in reducing the unfairness if all we must do is compare the elements of the crime and the elements of the tort; too often we might find the action passing muster under *Blockburger*'s minimal requirements. That would be good reason for states to reconsider whether *Blockburger* satisfies the fairness concerns of avoiding overpunishment.²⁴³ A transactional approach or evidentiary approach that limited what evidence could be adduced might provide better strategies.²⁴⁴

To summarize, under my proposal, retributive damages actions would be permitted and encouraged prior to criminal prosecution. Such amounts would be credited against criminal penalties that the government would assess. However, if the government has already secured a criminal conviction for the relevant conduct, that would preclude claims for retributive damages based on the underlying misconduct. Retributive damages based on the same misconduct to be proven in criminal court should not be permitted if the government has already filed an indictment (or similar declaration of accusation) against the defendant in a pending criminal proceeding, but if the criminal proceeding results in an acquittal, the retributive damages proceeding should then be permitted if the plaintiff is able to adduce evidence different from that which the government adduced in its case.

²⁴³ It will be the defendant's responsibility to bring information of prior adjudications to the court's attention, since the defendant will be in the best position to inform a civil plaintiff that she has already been indicted. If there is a "block" because of coincidental criminal prosecution, the courts should allow the relevant statute of limitations to equitably toll in the event that the prosecution fails.

²⁴⁴ See MILLER & WRIGHT, *supra* note 229, at 1002-05.

9. Excessiveness Review Under the Eighth Amendment

One might wonder, last, if the allocation of retributive damages to the state would cause the Supreme Court to revisit its earlier decision in *BFI v. Kelco* regarding the application of the Eighth Amendment to punitive damages.²⁴⁵ In *Kelco*, a majority of the Court found that the Eighth Amendment's prohibition on excessive fines did not cover punitive damages because they were directed to the private party, not to the state.²⁴⁶ However, the Court reserved judgment on the issue of whether punitive damages going to the state would in fact trigger Eighth Amendment review.²⁴⁷ Although not squarely presented by the parties, the issue could have been addressed in both *State Farm* and *Philip Morris*, where the statutes in Utah and Oregon, respectively, involved split-recovery schemes under which a significant portion of the punitive damages would go to the state.²⁴⁸ Nonetheless, the Court did not address the possible application of the Eighth Amendment, instead focusing its analysis on the Due Process Clause of the Fourteenth Amendment.

Assuming *arguendo* that retributive damages would be deemed a fine falling under the scope of the Eighth Amendment, my sense is that the proposal here would survive Excessive Fines Clause review for reasons I mentioned in the earlier companion article.²⁴⁹ Put briefly, in an era where the Court has extended extraordinary deference to legislative schemes of punishment allowing, for example, the incarceration of a defendant for at least twenty-five years for the theft of a few golf

²⁴⁵ *Browning-Ferris Indus. of Vt., Inc. v. Kelco Disposal, Inc.*, 492 U.S. 257 (1989).

²⁴⁶ *Id.* at 275-76.

²⁴⁷ *Id.* at 275 n.21. If there were such review, it would likely follow the framework established under *United States v. Bajakajian*, 524 U.S. 321 (1998). That doctrinal framework, by emphasizing whether the punishment is "grossly disproportionate" to the underlying offense, *id.* at 334, is not so strikingly different from the "guideposts" framework that the Court has articulated for the review of punitive damages under *BMW v. Gore* and *State Farm*, under which the primary determinant of the reasonableness of a punitive damages award is the reprehensibility of the defendant's misconduct.

²⁴⁸ See UTAH CODE ANN. § 78B-8-201(3)(a) (2008) (granting to the plaintiff the first \$50,000 of a punitive damages award and one-half of the excess of punitive damages over \$50,000); OR. REV. STAT. § 31.735(a) (2007) (granting forty percent of the punitive damages award to the plaintiff).

²⁴⁹ See Markel, *Retributive Damages*, *supra* note 1, at 327-32. That said, when the state collects damages from private entities, other practical problems worthy of attention can arise. See, e.g., Hanoeh Dagan & James J. White, *Governments, Citizens, and Injurious Industries*, 75 N.Y.U. L. REV. 354 (2000).

clubs,²⁵⁰ there could be little plausible basis for an objection to a civil penalty that stripped unlawful gains and removed no more than ten percent of one's prior wealth.²⁵¹ Significantly, this retributive damages penalty would occur only after proceedings in which the defendant enjoyed a cluster of important procedural safeguards, as described above.

IV. IMPLEMENTING EXTRACOMPENSATORY DAMAGES

This Part addresses two additional key issues associated with the implementation of extracompensatory damages: insurance and settlement.

A. *Should Insurance Be Available—If So, When?*

Assuming a contract existed between an insured and insurer that created liability insurance to cover punitive damages,²⁵² there is still the question of whether such contracts ought to be enforced on public policy grounds. Currently, nine states prohibit the availability of insurance for punitive damages, but the majority of jurisdictions permit such insurance and about a dozen states have not decided conclusively through courts or statutes what the rule is.²⁵³ Even in those states taking a hostile view of insurance for punitive damages, most will make an exception when the punitive damages liability is imposed vicariously, as opposed to directly.²⁵⁴ The rationales in support of in-

²⁵⁰ See *Ewing v. California*, 538 U.S. 11, 29-31 (2003) (finding that the defendant's "history of felony recidivism" helped to justify his sentence under a three-strikes law).

²⁵¹ Cf. *Bajakajian*, 524 U.S. at 336 (observing the need for deference to the legislative penalty scheme and a willingness to strike down only those fines that are grossly disproportionate to the severity of the offense).

²⁵² See Tom Baker, *Reconsidering Insurance for Punitive Damages*, 1998 WIS. L. REV. 101, 115 ("[T]here is little dispute that, on their face, most primary general and automobile policies provide coverage for punitive damages."). Nonetheless, insurers employ a variety of strategies to limit their exposure to paying punitive damages claims. See *id.* at 116-25 (describing how insurance companies use underwriting practices, policy limits, contract provisions, and refusals to pay based on public policy claims to control such payments). The analysis in this Section applies principally to those risks that insurance companies want to insure *ex ante*; their claims about why their contracts should not be enforced *ex post* in particular cases is a problem primarily of contract interpretation that I leave aside.

²⁵³ See Catherine M. Sharkey, *Revisiting the Noninsurable Costs of Accidents*, 64 MD. L. REV. 409, 430 (2005) ("[T]he legislative and judicial trend in the past several decades has been squarely in the direction of expanded insurability."); *id.* at 456-60 (presenting results from a recent fifty-state survey).

²⁵⁴ *Id.* at 428-29.

insurance for punitive damages have in recent years largely been deterrence oriented.²⁵⁵ Unfortunately, a number of courts and commentators have defended the opposition to insurance for punitive damages on retributive justice grounds.²⁵⁶

My goal here is to weaken the opposition, at least in some contexts where that opposition is articulated in the discourse of retributive justice. As I will argue, this hostile view of insurance is largely mistaken, at least when insurance markets for retributive damages are permitted to operate in situations when the harm arises from insurable risks—i.e., those risks that are probabilistic in nature and not “highly susceptible to [an insured’s] moral hazard.”²⁵⁷ The policy that I defend, however, explains why the insurance policy for retributive damages must also require some form of coinsurance on the part of the defendant, beyond the premium paid *ex ante*, in order to ensure that the defendant faces a direct and coercive setback to her interests through the imposition of an intermediate sanction. The logic of this discussion also applies to considerations of insurance for aggravated and deterrence damages.

1. Insurance for Insurable Reckless Risks

Say a law firm partner announced to an insurer, “I’d like to buy insurance in case I get sued for when I kill one of my associates next week—I’m really frustrated with his performance.” Practically speaking, no company would insure for this risk of retributive damages because insurers will not cover damages that are “expected or intended”

²⁵⁵ See *id.* at 430-31 (collecting cases referring to this deterrence rationale); see also George L. Priest, *Insurability and Punitive Damages*, 40 ALA. L. REV. 1009, 1031 (1989).

²⁵⁶ E.g., *Nw. Nat’l Cas. Co. v. McNulty*, 307 F.2d 432, 440 (5th Cir. 1962) (“The policy considerations in a state where . . . punitive damages are awarded for punishment and deterrence, would seem to require that the damages rest ultimately as well [as] nominally on the party actually responsible for the wrong.”), *superseded by statute*, VA. CODE ANN. § 38.2-227 (2007) (providing that it is not against Virginia public policy to obtain insurance providing coverage for punitive damages owed due to “negligence, including willful and wanton negligence,” but that it is contrary to public policy to obtain insurance covering punitive damages owed because of intentional misconduct), as recognized in *United Servs. Auto. Ass’n v. Webb*, 369 S.E.2d 196, 197 (Va. 1988).

²⁵⁷ Priest, *supra* note 255, at 1029; see also Baker, *supra* note 252, at 114-25 (detailing insurance company mechanisms for coping with the moral hazard problem associated with punitive damages). Moral hazard has been defined as the phenomenon by which injury and activity rates increase as a response to a decrease in the expected costs of injury. See, e.g., Priest, *supra* note 255, at 1023 n.55.

from the insured's perspective.²⁵⁸ In the context of intentional conduct—that is, where the defendant is purposefully or knowingly causing an unjustified harm to someone—the availability of insurance for that misconduct might make it appear as if the insurers of punitive damages are bankrolling or emboldening the conduct and are thus perhaps complicit in it. After all, the insurers have before them a person who wants to create gains from trade on the basis of planned misconduct that is expected, with practical certainty, to unjustifiably injure someone. Unsurprisingly, many states allowing insurance for punitive damages still maintain an exclusion for intentional conduct.²⁵⁹ But aside from the legal prohibitions, insurers won't take on risks that are not insurable in the way defined at the outset.

By contrast, in the context of reckless conduct, where the defendant disregards a substantial and unjustifiable risk of causing harm through indifference to the rights and interests of others, the issue of insurance for retributive damages is more complex because some risks arising from reckless conduct will satisfy an insurer's test for "insurability."²⁶⁰ Some states have proscribed the availability of all insurance for punitive damages on grounds that it is repugnant to public policy to enforce contracts that work to ratify malicious or reckless conduct.²⁶¹

²⁵⁸ See, e.g., Priest, *supra* note 255, at 1015 (noting that in every insurance and punitive damages case studied by the author, the insurance policy excluded intentional acts and covered only harms "neither expected nor intended from the standpoint of the insured" (quoting *Am. Home Assurance Co. v. Safeway Steel Prods. Co.*, 743 S.W.2d 693, 695 (Tex. App. 1987))). Priest also explains how the exclusion of intentional conduct from insurance policies works to the benefit of lowering premiums for those "insureds not intentionally engaging in acts causing harm" and prevents losses to those who would suffer from intentional harms if insurance for such harms were permitted. *Id.* at 1026.

²⁵⁹ Sharkey, *supra* note 253, at 432 n.118 (listing these states).

²⁶⁰ The same complexity attaches when a state predicates retributive damages liability on a mens rea formula like "wanton disregard."

²⁶¹ See, e.g., *Am. Sur. Co. of N.Y. v. Gold*, 375 F.2d 523, 525 (10th Cir. 1966) ("[W]e are convinced from the weight and logic of the case law that Kansas would hold a policy insuring against punitive damage awards to be violative of the public policy of that state . . ."); *McNulty*, 307 F.2d at 433-34 (holding that an insurance policy providing coverage for punitive damages would "contravene public policy"); *Crull v. Gleb*, 382 S.W.2d 17, 23 (Mo. Ct. App. 1964) ("We hold that to allow a motorist to insure himself against judgments imposed against him for punitive damages, which were assessed against him for his wanton, reckless or willful acts, would be contrary to public policy.").

My concern is that failing to enforce contracts for insurance based on reckless misconduct may be shortsighted.²⁶²

Why would permitting insurance for retributive damages for reckless conduct involving insurable risks be attractive? To understand the answer we have to look at the nature of insurance and the incentives of the parties created *ex ante* by the insurance policy for retributive damages. Insurance for retributive damages not only ameliorates the judgment-proof defendant problem while also promoting the distributive goal of loss spreading; it also, through the law of large numbers, reduces the amount of harm by the very aggregation of probabilistic and independent events.²⁶³ Additionally, and perhaps most importantly, the insurer will often have strong financial incentives to monitor the insured's behavior and to ensure that the insured does not undertake conduct that can instigate retributive damages.²⁶⁴ Moreover, at the time of the initial purchase of insurance, the insurer is motivated to engage in an optimal level of "risk-pool definition," which basically involves separation of clients into pools of varying levels of risky activity so one can underwrite insurance policies at the right prices: for example, charging smokers more for life insurance.²⁶⁵

²⁶² To be sure, the line between reckless misconduct and malicious misconduct can be hard to draw at times. Drunk driving—the conduct that motivates many punitive damages cases involving insurance coverage—might be thought of as a hard case. In most cases, the defendant's state of mind is understood to be reckless because there is a substantial and unjustifiable risk of serious injury to both the defendant (which thereby lessens the moral hazard effect) and to others. The defendant usually does not intend to or know that she will crash and hurt someone, and the risk of harm caused might not result in actual harm. Still, the conduct, even when harmless, is itself condemnable as a serious wrong because of the manifestly insufficient regard for the well-being of others that such conduct evinces. And if someone died as a result of the defendant's drunk driving, prosecutors might plausibly say that the defendant acted with such depraved-heart recklessness that the *mens rea* for common law murder would be satisfied.

²⁶³ See Priest, *supra* note 255, at 1021-22 (explaining the math and reasoning of the law of large numbers).

²⁶⁴ But see Tom Baker & Sean J. Griffith, *The Missing Monitor in Corporate Governance: The Directors' & Officers' Liability Insurer*, 95 GEO. L.J. 1795, 1807-17 (2007) (providing some empirical qualitative evidence of instances in which insurers did not seek to invest in monitoring during the life of the insurance policy).

²⁶⁵ *Id.* at 1022. Contrary to some views informed by modern portfolio theory, which would counsel insurers simply to acquire a diversified portfolio of risk, recent empirical evidence confirms that, in the context of entities, insurance companies do in fact analyze a prospective insured's corporate culture to assess risks of misconduct. See Tom Baker & Sean J. Griffith, *Predicting Corporate Governance Risk: Evidence from the Directors' & Officers' Liability Insurance Market*, 74 U. CHI. L. REV. 487, 517 (2007) (noting that directors' and officers' liability insurance underwriters repeatedly told the authors

The risk sorting is often dynamic or experience rated.²⁶⁶ The incentives for risk sorting begin when screening prospective insureds, and they often continue to exist over the course of the policy. The insurer will typically have an incentive to engage in monitoring and the insured will want to keep premiums down; market forces will spur the insured to invest in compliance mechanisms and the insurer will help ensure that those investments are effective.²⁶⁷ Moreover, to reduce moral hazard temptations, insurance companies typically insist on some form of coinsurance (whether through deductibles or copayments based on flat or variable percentage amounts) when it comes to insurable risks.²⁶⁸

that corporate culture and character are important considerations in assessing risk). Perhaps the insurers are making an economic mistake, but they would not be doing this risk sorting unless they thought that it was to their advantage to do so; in making the risk-selection choices, they are pricing the expected cost of punitive damages vis-à-vis a particular client into the premium they charge. That said, Baker and Griffith argue that, at least in the context of directors' and officers' insurance, the levels of insurance purchased by management for the benefit of the shareholders will be inefficient and strongly influenced by agency costs. See Baker & Griffith, *supra* note 264, at 1800 (arguing that managers buy certain forms of directors' and officers' insurance for "self-serving" reasons).

²⁶⁶ As Professor Miriam Baer rightly pointed out in a comment on an earlier draft, the possibility of experience-rating the insured should not obscure the fact that other factors (whether there is a hard or soft market for insurance, for example) can drown out the experience aspect of the premium. In other words, many market entrants might lower the premiums even for those who are prone to driving recklessly. Conversely, it is possible that the reduction in premium that one gets from driving extremely carefully might be overcome by the overall increase in car insurance when a major insurer goes belly-up and the insurance markets contract suddenly.

²⁶⁷ As discussed *supra* note 265, however, insurers often have incentives to reduce information costs associated with specific investigations and simply follow the portfolio-theory views of insurance or actuarial tables associated with particular kinds of risk borne by particular buyers of insurance. See also Lawrence A. Cunningham, *Too Big to Fail: Moral Hazard in Auditing and the Need to Restructure the Industry Before It Unravels*, 106 COLUM. L. REV. 1698, 1743 (2006) ("Most insurance underwriting exercises involve classifying risks using general actuarial tools rather than specific investigation."). But see *id.* at 1743-44 (discussing the types of insurance products that typically involve specific investigations). The fact that insurance companies choose to conduct their business this way is relevant; under my proposal, one is not mandating the availability of an insurance policy for retributive damages, but rather is permitting it. And if it is permitted, the insurance company should be able, within reason, to set the terms of the agreement and how it thinks it should proceed. Thus, I am not especially concerned that insurance companies might lose money by failing to monitor insureds closely during the life of the policy. See Baker & Griffith, *supra* note 264, at 1813.

²⁶⁸ Should a legislature require that insurers only be able to write insurance policies that have a coinsurance scheme? It would be a practical concession to the negative appearance that some judges and scholars have said is created by permitting insurance for punitive damages. But it would also ensure that there is a direct, rather than

How do these principles apply to a generic lawsuit where a defendant is sued for harms caused by reckless drunk driving and the jury finds a basis for retributive damages liability based on the recklessness? At the outset, the insurance company might engage in loss-prevention techniques by requiring anyone (of certain age or other characteristics) interested in purchasing car insurance to install a breathalyzer-ignition system. Indeed, an insurance market provides insurers with incentives to work with technology innovators that can ensure that the breath of the person driving is the breath of the person blowing the air. As for the defendant, to the extent that she processes costs and benefits at the time of the misconduct in question,²⁶⁹ or at a point when she can take steps to avoid the risk of future misconduct, she will fear that her conduct will be excluded as intentional, or, if not deemed intentional, that it will trigger increases in her insurance premiums or even subsequent policy cancellation. She will also fear that her insurance policy might not cover the entire amount of retributive damages perhaps because of contractual caps on insurance payouts or because of contractual coinsurance obligations. She also has reason to fear that she will have to pay litigation costs and endure the intermediate condemnatory signal associated with retributive damages and the risk that such damages may trigger subsequent criminal liability.

Thus, contra Judge Calabresi,²⁷⁰ there are good instrumental reasons that retributive damages associated with drunk driving should be insurable if the insurance companies are willing to assume those risks. The logic of having insurance for retributive damages even makes sense when no actual harm erupts, as long as the insurer *ex ante* believes that the risk of harm is by nature insurable and that the risk is capable of being reduced through loss-prevention strategies.²⁷¹ Hav-

mented, setback to the defendant's interests that would be independently justifiable and warranted.

²⁶⁹ See generally Paul H. Robinson & John Darley, *Does Criminal Law Deter? A Behavioral Science Investigation*, 24 OXFORD J. LEGAL STUD. 173 (2004) (providing reasons for skepticism about defendants' amenability to deterrence signals).

²⁷⁰ See CALABRESI, *supra* note 18, at 269-70 (arguing that in cases such as careless driving, when "normal individuals can choose whether or not to engage in wrongful conduct before an accident, an appropriate noninsurable penalty is necessarily a more effective deterrent than an already paid insurance premium").

²⁷¹ Say a law student sees her drunk professor leaving the bar, stumbling into a car, and starting that car; the student then follows the professor (with a camcorder) for a mile, watching the professor swerve across the road. The professor ultimately drives off the road into some grass, hurting no one, and falls asleep. My scheme would have

ing insurance available for retributive damages brought by private attorneys general facilitates achieving the intrinsic benefits of retribution and the norm-projection benefits associated with having a private attorney general seek retributive damages when, without such insurance, there might not be any incentive to bother vindicating these wrongs.²⁷²

The same logic would apply even more strongly to those areas—products liability, unconscious bias that creates employment discrimination, and mass torts—where the resulting harm is something that the entity's leaders strongly want to avoid but may not be in a position to know about.²⁷³ By drawing on insurers' objectivity, expertise, and experience, insureds can use insurers' involvement to structure operations to achieve greater compliance and safety.²⁷⁴ Indeed, in certain contexts involving sophisticated parties, one can imagine how a defendant's failure to seek out best (safety) practices from insurers

legislatures authorize persons to bring actions for retributive damages in instances like this. See Markel, *Retributive Damages*, *supra* note 1, at 279-86.

²⁷² Cf. Baker, *supra* note 252, at 129 (concluding that insurance availability will increase the likelihood that plaintiffs will bring suit).

²⁷³ In the context of purchasers of insurance, we also need to think carefully about the differences in treatment of defendants who are individuals versus those that are partnerships, and between public- versus private-firm defendants. First, with corporate entities, there is an agency-cost problem worth spotting: managers may underpurchase—or overpurchase—the amount of insurance needed to protect the interests of the owners. Second, the determination of how much insurance to purchase and who makes that decision may not relate well to who can nimbly respond to the appropriate signals established through insurance markets and who deserves the attribution of blame for failing to make good decisions. Thanks to Dave Hoffman for helping me recognize this. Additionally, the idea of moral hazard in the publicly held corporation has a different valence than when applied to an individual, for example. In the public corporation, the *shareholders* are indirectly paying for the insurance, while some *risk manager* is arranging for the corporation's yearly insurance policy, and some *manager* or *director* violates the law. In that case of disaggregation, insurance for retributive damages may be creating a hazard, but it is not necessarily a “moral” one, in the sense that the purchaser of the insurance is trying to benefit unduly through its lack of adequate care. For more on this, see Miriam Hechler Baer, *Insuring Corporate Crime*, 83 IND L.J. 1035, 1083-84 (2008). Of course, *ex ante*, it is unclear why shareholders should be benefiting from structures that would insulate them from punishment for the wrongs that these structures of ownership and management perpetrate. See Markel, *Punitive Damages*, *supra* note 9.

²⁷⁴ See RICHARD V. ERICSON ET AL., *INSURANCE AS GOVERNANCE* ch. 8 (2003) (detailing ways in which insurers actively encourage those that they insure to adopt measures to limit risk and prevent loss); Sharkey, *supra* note 253, at 413 (“Insurance companies, as private regulators, are well positioned to achieve deterrence through experience rating of firms and other actors, as well as by providing risk management services.”).

might be regarded as a culpable omission if and when suits for retributive damages are filed. In short, foreclosing the market for insurance categorically may lead to more encroachments against the rights and interests of workers or customers, and since retributive justice is properly sensitive to how its structures prevent or cause such encroachments,²⁷⁵ insurers have good, publicly justifiable reasons to cover even insurable risks from reckless misconduct.²⁷⁶

2. Does Access to Insurance Blunt the Retributive Condemnation?

But what of the claim that such insurance undermines the condemnatory signal associated with retributive damages? With the structure of insurance described above in place, retributivists cannot reasonably complain that the retributive damages have lost their condemnatory communicative value. As long as the party suffering the damages award can plausibly be said to experience a coercive condemnatory setback to their interests—which they will face through increased premiums, a coinsurance payment, and some reputational damage—there has been punishment sufficient to satisfy the demands of retributive damages as an *intermediate* sanction. And to the extent that the objectivity and expertise brought by the insurer facilitate a culture of compliance at the insured's workplace or home, the policy can be said to be facilitating the respectful treatment of the rights of the individuals in those spaces.²⁷⁷ As Professor Baker rightly argues,

[W]hether or not a punitive damages award is insured, that award makes a public statement about the value of the victim that contradicts the private assessment of the perpetrator at the time the wrong was committed. Moreover, insured or not, a punitive damages award represents a consequence for extreme culpable behavior that goes beyond that for less culpable behavior. Although an insured punitive damages award clearly will hurt less than an uninsured one, it nevertheless does hurt: insur-

²⁷⁵ See Baker, *supra* note 252, at 110-11 (discussing how philosopher Jean Hampton's retributive theory incorporates concerns for prevention); Markel, *Retributive Damages*, *supra* note 1, at 268 (arguing that under the theory of confrontational retributivism, the establishment of institutions furthering retributive justice will in practice facilitate the prevention of future wrongdoing).

²⁷⁶ Whether such an insurance market is likely to develop in practice or to become too concentrated are questions left for another day. Cf. Baer, *supra* note 273, at 1092-94 (addressing these questions).

²⁷⁷ This concern for prevention is an aspect of the *ex ante* function of retributivism, discussed in Part II of Markel, *Retributive Damages*, *supra* note 1, and this rationale applies, quite naturally, to permitting insurance for compensatory, aggravated, and deterrence damages purposes too.

ance is made more expensive and less available in the future; for commercial entities there is a loss of good will; and for any individuals involved, there are negative social consequences.²⁷⁸

In some ways, the question of insurance is also tied to what the state can reasonably ask of a defendant found liable for retributive damages. An analogy might be useful in explaining the limits of the argument that the signal for retributive damages is blunted through insurance. Imagine that we fined an individual offender \$10,000 for her criminal misconduct. There is no robust insurance market for criminal fines,²⁷⁹ probably because many fines are predicated on non-insurable risks. Notwithstanding the lack of insurance markets for such fines, we do not ask the offender to pay the fine using only her earned income. If the money is derived from a lucky investment, a racetrack gamble that paid off, or her husband's inheritance, the state would still take the money; the same is true if the money came from a loan from a neighbor or the bank. So why should we care so much when part of the money in a retributive damages scheme comes from insurance companies, given that the purchase of such insurance is also consistent with a signal of being considerate of others by reducing the risk of harm to them? In the end, it does not make sense to intrude too much into how the retributive damages are paid.

To be sure, some of the preceding remarks justify the practice of insurance for retributive damages based on reckless conduct in consequentialist terms. But I have also tried to defend the use of insurance as a vehicle by which the class of future defendants can show that they are considerate of those around them and respect people's rights to well-being and physical integrity. This alternative justification helps explain

²⁷⁸ Baker, *supra* note 252, at 112. Baker wrote this paragraph under the inspiration of Galanter and Luban's account of punitive damages as victim vindication. *See id.* at 110 ("As Galanter and Luban explain, the retributive purpose of punitive damages 'is to reassert the truth about the relative value of wrongdoer and victim by inflicting a publicly visible defeat on the wrongdoer.'" (quoting Galanter & Luban, *supra* note 3, at 1432)). But with the right adjustment—replacing "the value of the victim" with the "value of the social interest in equal liberty under law"—I think the rest of the paragraph is exactly correct in reflecting the public values of the retributive account in this project.

²⁷⁹ *But cf.* *Nw. Nat'l Cas. Co. v. McNulty*, 307 F.2d 432, 440 (5th Cir. 1962) (Wisdom, J.) ("It is not disputed that insurance against criminal fines or penalties would be void as violative of public policy. The same public policy should invalidate any contract of insurance against the civil punishment that punitive damages represent."), *superseded by statute*, VA. CODE ANN. § 38.2-227 (1986), *as recognized in* *United Servs. Auto. Ass'n v. Webb*, 369 S.E.2d 196, 197 (Va. 1988). Judge Wisdom, however, goes too far, for reasons that I hope my account explains.

the attractiveness of the policy in nonconsequentialist terms. Thus, one of this account's innovations is that, contrary to the reasoning of courts or commentators that have found otherwise,²⁸⁰ there is now reason to think that insurance for some conduct culminating in punitive damages is consistent with the retributive goals of punitive damages.

Of course, as I stated at the beginning, if there is evidence that the insurance company's agents were actually emboldening the defendant's reckless misconduct or giving advice on how to engage in the conduct with greater likelihood of concealment, then that would be a basis for thinking that the insurance company is somehow complicit in the wrongdoing.²⁸¹ But if the insurance company is simply helping to reduce the one-time sting of retributive damages and is otherwise incentivized and committed to reducing the incidence of the defendant's misconduct, then there is no good retributive reason to oppose insurance for retributive damages for liability based on insurable risks associated with reckless misconduct.

The logic of the preceding argument similarly applies to insurers' desire to cover insurable risks that lead to deterrence and aggravated damages. Since both aggravated and deterrence damages are designed to ensure appropriate cost internalization, there is good reason for the defendant to be able to acquire insurance for those purposes too, at least to the extent that insurers will continue to exclude intentional acts and cover only insurable risks.²⁸²

B. Settlement, Transparency, and Accountability

1. The Danger of Sweetheart Deals to Retributive Damages

Note that there is very little trouble with the conventional arrangement under which plaintiffs and defendants can settle their dis-

²⁸⁰ See, e.g., *id.* at 440 ("Where a person is able to insure himself against punishment he gains a freedom of misconduct inconsistent with the establishment of sanctions against such misconduct.").

²⁸¹ To facilitate the likelihood of greater monitoring and reduce the possibility of this kind of collusion, it might make sense for a jurisdiction to prefer insisting on an "occurrence" policy over a "claims made" policy. Under the former, the insurer at the time of the occurrence is required to pay, while under the "claims made" policy, the insurer who pays is the insurer at the time the claim is made. See Baer, *supra* note 273, at 1087 & n.255 (citing James D. Cox, *Private Litigation and the Deterrence of Corporate Misconduct*, LAW & CONTEMP. PROBS., Aug. 1997, at 1, 33).

²⁸² See Polinsky & Shavell, *supra* note 5, at 932 (expressing general support for insurance for the purpose of cost internalization damages); *id.* at 932 n.194 (citing others who share the authors' view).

putes over compensatory, aggravated, and deterrence damages. If a plaintiff is willing to gain in certainty through settlement what she might otherwise win at trial, then that should remain privately ordered without much more intrusion than what would normally occur under the current system of tort settlements.²⁸³

But with retributive damages paid largely to the state, the difficulties with settlement are much greater because of the possibility that sweetheart deals between defendants and plaintiffs would obstruct the goals of retributive damages. To use an example, a sweetheart deal in this context could occur when a plaintiff and defendant jointly expect a damages award to be worth ten dollars at the end of a trial—five in compensatory damages and five in punitive damages—and the state has a split-recovery scheme that would leave the plaintiff with seven dollars total, thus giving three dollars to the state. In this situation, a defendant might try to settle with the plaintiff for something greater than seven but less than ten dollars to deprive the state of the retributive damages.

Whether the sweetheart deal happens is probably a function of the legal shadow under which the parties will bargain.²⁸⁴ Unregulated sweetheart deals in cases involving misconduct warranting retributive damages trigger two negative consequences.²⁸⁵ First, settlements are often sealed and thus the defendant's behavior may not be adequately brought to light, preventing both legal condemnation and the sharing of relevant information with prospective plaintiffs who might have been similarly harmed.²⁸⁶ In other words, secret settlements may lead

²⁸³ But see *supra* note 91.

²⁸⁴ See generally Tom Baker, *Transforming Punishment into Compensation: In the Shadow of Punitive Damages*, 1998 WIS. L. REV. 211, 234-45 (listing factors that may alter incentives in allocating damages); Thomas Koenig, *The Shadow Effect of Punitive Damages on Settlements*, 1998 WIS. L. REV. 169, 172 (“[H]yperbole and simple confusion may shape settlements in a more powerful way than empirical truths.”).

²⁸⁵ But see Polinsky & Che, *supra* note 91, at 568 (noting that from an optimal deterrence perspective, settlements are good because they reduce litigation costs).

²⁸⁶ The clergy sex-abuse scandal manifested these problems. See TIMOTHY D. LYTON, *HOLDING BISHOPS ACCOUNTABLE: HOW LAWSUITS HELPED THE CATHOLIC CHURCH CONFRONT CLERGY SEXUAL ABUSE* pt. 2 (2008) (describing the use of sealed settlements); see also Scott Moss, *Illuminating Secrecy: A New Economic Analysis of Confidential Settlements*, 105 MICH. L. REV. 867, 870 (2007) (“With so many lawsuits beginning with allegations of grievous social harm but ending with the legal equivalent of ‘never mind,’ confidential settlements have drawn increasingly fierce criticism recently, attacked as ways defendants conceal serious misdeeds such as dissemination of hazardous products, discrimination, pollution, or sexual abuse.” (footnotes omitted)); Adam Liptak, *Judges Seek to Ban Secret Settlements in South Carolina*, N.Y. TIMES, Sept. 2, 2002, at A1 (“South Carolina’s 10 active federal trial judges have unanimously voted to ban se-

to nonpunishment in situations where punishment—through retributive damages—*should* exist.

Second, plaintiffs rarely insist that a defendant's settlement include money that would be listed on the punitive damages line. The money is treated simply as a transfer payment, with no particular description, or merely as a payment of compensatory damages. In the absence of any checks and balances, the government misses, among other things, two potential sources of revenue: first, the damages paid by a defendant in a settlement are treated simply as a cost of doing business and therefore are tax deductible; second, settlements in cases involving retributive damages might deprive the state of the retributive damages it otherwise would receive. Instead, the settlement creates a windfall in the plaintiff's wallet.²⁸⁷ In short, settlement offers a legalized form of deception that subverts some of the public goods that retributive damages are designed to achieve.

Here is a partial solution. First, in order for retributive damages to be available, the plaintiff must plead facts in the initial complaint that would, if true, evidence malice or recklessness. Permission to amend the complaint to allege such facts later should only be granted under compelling circumstances because otherwise, plaintiffs will omit such facts initially, threaten defendants during negotiations to amend the complaint with such facts afterward, and then settle for the sweetheart deal before the court or state representative realizes that important underlying facts were strategically omitted. If retributive damages are sought in a complaint initially, that could alert the judge

cret legal settlements, saying such agreements have made the courts complicit in hiding the truth about hazardous products, inept doctors and sexually abusive priests.”).

²⁸⁷ The compensatory part of personal-injury awards is not taxable income for the victim, but most federal courts have said that punitive damages awards are taxable income. See I.R.C. § 104(a) (2006) (providing that compensation received, whether by suit or agreement, on account of personal injuries or physical sickness is not included in gross income, with the exception of punitive damages). All the money in a settlement, however, would likely be described as “compensatory,” which creates a reason for the plaintiff to try to collude with the defendant against the state (or the insurance company). Cf. Baker, *supra* note 284, at 227-28 (“[B]oth plaintiffs’ and defense lawyers would prefer to see those aggravated damages ‘in the guise of compensatory damages’ rather than ‘in the guise of punitive damages.’ . . . For plaintiffs, the state share, the possibility of remittitur or reversal on appeal, and tax law are all important factors” (footnotes omitted)).

(and the state) to the greater role needed to supervise the settlement process.²⁸⁸

Second, following the filing of such a complaint, the court and a representative from the state attorney general's office would review and authorize settlements. If the state objected to the settlement because it seemed that the parties were engaging in a sweetheart deal and the defendant was trying to buy its way out of disclosing its misconduct, the state could exercise the option of buying the claim from the plaintiff (by paying the plaintiff a hypothetical "finder's fee" of \$10,000) and prosecuting the claim further than the plaintiff would have.²⁸⁹ This rule might especially make sense when the plaintiff simply wants compensatory damages and minimal sources of irritation.²⁹⁰

Because the state would have the choice of whether to seek retributive damages, defendants would not be able to induce a plaintiff to back down once the complaint is filed. But under the structure that I have discussed, even when parties negotiate payouts prior to the filing of the complaint, defendants will not agree to participate in a quiet shakedown by the plaintiff (i.e., settling a case involving reckless or malicious misconduct without paying retributive damages to the state as well).²⁹¹ Defendants will not agree to participate because there is always the chance that a private attorney general (PAG) will discover or already know about the defendant's misconduct and will share that information with the government. The government may then bring

²⁸⁸ Iowa has a split-recovery scheme in which the state plays an active role, for example, in monitoring litigation. For a description of this scheme, see Sharkey, *supra* note 5, at 435.

²⁸⁹ Perhaps the state should be able to sell the action to a third party too. See generally Michael Abramowicz, *On the Alienability of Legal Claims*, 114 YALE L.J. 697, 699 (2004) ("Courts increasingly have tolerated claim sales and have begun to view restraints on alienation skeptically."); cf. David Rosenberg, *Deregulating Insurance Subrogation: Towards an Ex Ante Market in Tort Claims* 307 (Harvard Law Sch. Pub. Law Research Paper No. 043, 2002), available at <http://ssrn.com/abstract=350940> ("Competitive ex ante claims markets (including secondary markets for claim re-sale and aggregation) provide the most efficient and effective means of achieving the deterrence ends of tort liability and generating the highest clearance price for sellers of potential tort claims.").

²⁹⁰ There is a risk that the government will not obtain all the information that it needs to make a good decision, but the government has various civil and criminal strategies to ensure that there is a full reporting by the lawyers involved.

²⁹¹ But cf. Sharkey, *supra* note 5, at 445 (noting that the risk of sham litigation in the context of the author's proposal for compensatory societal damages still "looms large"). On settlement pressure generally, see Charles Silver, *"We're Scared to Death": Class Certification and Blackmail*, 78 N.Y.U. L. REV. 1357 (2003).

suit for retributive damages, or, if it declines, the PAG may bring suit herself.²⁹²

Thus, if a plaintiff decided to allege retributive damages in the initial complaint, she would not be prohibited from subsequently settling. But this scenario would require plaintiffs to secure governmental approval to settle, and it would force defendants either to admit responsibility and pay some amount of retributive damages to the state or to deny responsibility.²⁹³ If the defendant denied responsibility, she would have to convince the state's representative that this particular claim was not worth pursuing because it lacked merit. Otherwise, the state—or, if the state declined, conceivably another PAG—could decide to sue the defendant. Clearly, the dynamics of settlement would change because defendants would have little incentive to settle without admitting liability. With these diminished incentives, plaintiffs would be unlikely to bring suits merely for the purposes of harassment. Under this structure, settlements for suits involving reckless or malicious misconduct could continue apace, but courts and the government would have to supervise the distribution of rewards and fees to plaintiffs and their lawyers to ensure a reasonably fair allocation consistent with what would happen were a trial to occur.²⁹⁴

Transparency with respect to settlement agreements would also be required in light of the public retributivist values articulated in Part II. A sealed settlement in a case where retributive damages were sought in the complaint might deprive the public of valuable information regarding the danger of the defendant's activity and its potential harm to others. Some jurisdictions are considering a prohibition on sealed settlements generally, with an eye toward assuring the public that the courts will not be facilitating patterns of complicit concealment.²⁹⁵

²⁹² See Markel, *Retributive Damages*, *supra* note 1, at 279-86.

²⁹³ Once the defendant admitted to wrongdoing and settled with the state, such information could be stored in a registry of the sort described by Jim Gash, *Solving the Multiple Punishments Problem: A Call for a National Punitive Damages Registry*, 99 NW. U. L. REV. 1613, 1617 (2005), though some modifications would need to be made to be compatible with the structure proposed here.

²⁹⁴ One might think that this supervisory role would necessitate large-scale hiring efforts on the part of government bureaucracies, but such fears seem overstated in light of the relative infrequency with which punitive damages are awarded in most jurisdictions. See Sebok, *supra* note 5 (surveying empirical literature on the infrequency of punitive damages). That said, the structure proposed here would change the litigation game substantially, so it is inappropriate to dismiss these fears altogether.

²⁹⁵ For discussions of the problems associated with concealing settlements, see generally Owen M. Fiss, *Against Settlement*, 93 YALE L.J. 1073 (1984); Susan P.

Whether such a general prohibition is a good idea is a subject for debate.²⁹⁶ However, it seems that prohibiting sealed settlements in cases where retributive damages are alleged in the complaint would be a good starting place, at least where the defendant's misconduct might have affected other citizens or might endanger them in the future.²⁹⁷

2. Settlement, Insurance, and the Public Interest

The introduction of liability insurance into any tort claim involving punitive damages further complicates settlement. Plaintiffs' lawyers are most focused on targeting a deep pocket so that they can collect damages. They also know that insurance policies exclude intentional acts. Hence, if there is a way to manipulate their client's claim so that the defendant is merely reckless, or even better, negligent, then the plaintiff's lawyer will seek to characterize the harm as recklessly or negligently caused. This strategic decision ensures that there is a deep pocket to cover the injury for settlements. If the defendant has insurance, she too is tempted to collude with the plaintiff to change the theory of the case so that an intentional tort is characterized instead as a negligence action. That collusion will ensure that the defendant does not have to pay the damages out of her own pocket.²⁹⁸ Indeed, this set of joint interests in transforming punishment into compensation may also explain various trial results.

There is no question that, to the extent that it would occur under the proposal here, this manipulation is subversive of retributive values because it undermines the public's interest in reducing Type II (false-negative) errors where the defendant escapes any appropriate censure for her misconduct. Additionally, claim manipulation obviously works against the interest of the insurance company, which might have had

Koniak & George M. Cohen, *Under Cloak of Settlement*, 82 VA. L. REV. 1051 (1996); Liptak, *supra* note 286.

²⁹⁶ See, e.g., Moss, *supra* note 286 (arguing that the economic perspective on sealed settlements is more ambiguous than previously conceived).

²⁹⁷ Judges already have a heightened obligation to supervise settlement classes, but they should not shirk their responsibility to consider nonclass settlements that may have profound effects on similarly situated litigants in the future.

²⁹⁸ See Adam Liptak, *Pain-and-Suffering Awards Let Juries Avoid New Limits*, N.Y. TIMES, Oct. 28, 2002, at A14 (analyzing ways in which lawyers and juries circumvent state limits on punitive damages, and noting a particular case where a jury awarded a woman who endured sexual harassment at work \$20 million in compensatory damages for pain and suffering).

a contractual exclusion for the kind of conduct now being settled and described in less noxious terms.

Such collusion and manipulation are constrained in several ways. Generally, the insurance company may do its own due diligence to see if there is any reason to suspect that the claim ought to be excluded under the contractual provisions. Moreover, the defendant's premium for insurance will increase as a result of past negligence as well as past recklessness or malice. That will reduce the defendant's ability to plausibly move all "punishment" into the insurance company's responsibility for "compensation." Additionally, the strategic behavior may be chilled both by threats of prosecution for insurance fraud, if discovered, and by codes of professional ethics that govern the conduct of lawyers.

Moreover, the retributive damages structure that I have proposed remedies the claim-manipulation problem in two respects. First, consider the insurance context. If the state denies *any* insurance for punitive damages, the pattern of "transforming punishment into compensation" continues largely unabated. But if the state permits insurance for retributive damages, as I have suggested that it should for certain risks, what happens? The parties lose some of their incentive to mischaracterize the claim because there will be insurance coverage for the retributive damages claim (minus the amount of coinsurance).

Second, and far more importantly, defendants will be less likely to collude with plaintiffs because without paying the public penalty associated with that conduct, the defendant will have little to no repose—depending on statutes of limitations—against any of the private attorneys general who might know of or discover the defendant's misconduct and subsequently bring suit for retributive damages only, even after the victim has been paid compensatory damages. The threat of the PAG, in other words, reduces the likelihood of settlements that obscure the misconduct from public assessment.

Finally, even if the retributive damages structure were subverted by the intersection of insurance and settlement incentives, plaintiffs would still be able to seek retributive damages in those cases where the defendant has the wherewithal to self-insure the retributive damages. While that regrettably reduces the available scope of retributive damages defendants, such a result is by no means a tragic one in light of the pronounced difficulty of punishing misconduct by the wealthy

and powerful.²⁹⁹ Moreover, if after trying this regime out we still became worried about how abusive or pervasive the manipulation of claims was, then we could experiment with increased transparency and judicial supervision of settlements even in claims where no retributive damages were alleged.

CONCLUSION

Despite the Supreme Court's increased tinkering with the constitutional boundaries of punitive damages over the last fifteen years, it has nonetheless left open to states a range of options regarding their structure and purpose(s). This project has sought to give states a blueprint for a pluralistic and attractive punitive damages regime that could serve separate purposes without those purposes crossing wires with each other or with the federal constitutional framework. While also examining the functions of cost internalization and victim vindication, I have paid particular attention to developing a strategy to advance the public's interest in retributive justice without upsetting the constitutional framework established by the Supreme Court. I hope that the project makes sense thus far with respect to procedural safeguards, settlements, and insurance. Nonetheless, there are still important questions to resolve pertaining to the appropriate treatment of punitive damages in complex matters regarding multiple injuries and lawsuits against entities. These questions are some of the issues that I address in the next article of this series. In the meantime, the Appendix summarizes some of the basic conclusions in the form of instructions designed to handle a pluralistic extracompensatory damages regime such as the one described in this Article.

²⁹⁹ See Markel, *Retributive Damages*, *supra* note 1, pt. IV (discussing the rationale of retributive damages as a tool to counteract misconduct by wealthy and powerful persons and entities).

APPENDIX: INSTRUCTIONS FOR ASSESSING
EXTRACOMPENSATORY DAMAGES

What follows is a distillation of the principal conclusions of this punitive damages project. This summary is meant as a rough draft for potential jury instructions within a given jurisdiction that has adopted this pluralistic scheme. These instructions are also designed to take into account the holding of the Supreme Court's recent decision in *Philip Morris*.³⁰⁰ The use of square brackets below is meant to indicate places where the polity has some flexibility with respect to the relevant instruction.

* * *

In considering the amount of extracompensatory damages on the defendant, you should determine whether three separate dollar amounts are necessary: (1) an amount to accomplish retributive justice against the defendant; (2) an amount to accomplish deterrence; and (3) an amount to vindicate the injury to the victim's personal dignity.

A. *Retributive Damages*

Retributive damages fulfill the punishment objective of extracompensatory damages. These instructions apply only to defendants who have committed misconduct that you have found to be malicious or reckless in nature. If you do not think, based on clear and convincing evidence, that the conduct in question was malicious or reckless in nature, do *not* award retributive damages.

Malicious conduct is conduct that was done with a purpose or knowledge of causing harm and for which no other legally recognized excuse or justification is available as a defense. A defendant acts recklessly when she consciously disregards a substantial and unjustifiable risk that harm will result from her conduct. The risk must be of such a nature and degree that, considering the nature and purpose of the actor's conduct and the circumstances known to the actor, her disregard involves a gross deviation from the standard of conduct that a

³⁰⁰ These instructions are a substantially modified version of the kind found in Polinsky & Shavell, *supra* note 5, app. In some places, having mostly to do with cost internalization, I expressly borrow the language from their proposed jury instructions.

law-abiding person would observe in the actor's situation. If there are multiple defendants, you must undertake this analysis separately for each of the defendants based on each defendant's misconduct. [A defendant corporation will not be held legally responsible for all of the misconduct of each of its employees.] You must ask whether each defendant's action was malicious or reckless.

If, and only if, you have determined that a particular defendant's misconduct was undertaken with malice or recklessness, the next step requires consulting the chart prepared by the state legislature. This chart should help you determine where on a scale of one to twenty—with twenty being the most reprehensible and one being the least—the defendant's misconduct lies. The chart tells you whether to add points to the scale based on various factors and whether to subtract points based on other factors. Your job is to assess the wrongfulness of the defendant's misconduct based on the reprehensibility chart. It is not your job to assess how much harm the defendant's misconduct has caused to society or other nonparties to this litigation. This finding of reprehensibility should also be accompanied by an explanation of what facts you considered relevant to your determination. Once you have determined the level of reprehensibility, the court will use a different chart to determine the amount of retributive damages that the defendant will pay based on your assessment of reprehensibility.

In determining the reprehensibility of the defendant's misconduct, you may, but are not required to, consider "evidence of actual harm to nonparties" because that can help show "that the conduct that harmed the plaintiff also posed a substantial risk of harm to the general public, and so was particularly reprehensible."³⁰¹ Similarly, you may also consider the harm or potential harm that the defendant's conduct caused to others in determining whether the defendant's misconduct was accidental or deliberate or part of a policy or pattern and practice. However, it is important that you not consider the mere fact that others were harmed as a basis for assessing a higher reprehensibility score. Those others who are not plaintiffs in this case can bring their own suits for compensatory and other damages.

Two facts are relevant to your task, although they should not inform your actual assessment of the reprehensibility of the defendant's misconduct. First, the plaintiff will personally receive no more than [\$10,000] of the retributive damages award. The balance will go to the state [to advance law enforcement objectives, including but not

³⁰¹ *Philip Morris USA v. Williams*, 549 U.S. 346, 355 (2007).

limited to providing services necessary for victims and offender reentry into society].

Second, the purpose of retributive damages is to make the defendant worse off than she would have been had she not undertaken her malicious or reckless misconduct. Thus, when determining the level of reprehensibility, do not consider the amount of other damages (whether compensatory, “aggravated,” or “deterrence,” described below). [If the defendant has made such payments or has been otherwise punished through the criminal justice system of this jurisdiction, then you ought to forego making any reprehensibility assessment.] [Note to judges: civil penalties already assessed against the defendant for this misconduct against this plaintiff should be credited against retributive damages. No retributive damages are available if the government has already criminally punished the defendant for the wrong to the particular plaintiff in this case.]

After you make your assessment of reprehensibility, the court [or you, the jury] will determine whether any other gains or profits by the defendant need to be forfeited in addition to the reprehensibility-based retributive damages award. The court may also make subsequent determinations regarding reasonable attorneys’ fees and costs to be determined in light of the risk, time, expense, and expertise related to this litigation. [It may also be your job to determine the financial condition of the defendant, or its net value if the defendant is an entity.]

B. Aggravated Damages for Repairing Personal Dignity Harms

In deciding the remedy for personal dignity harms, please first ensure that you have not already figured this amount into your assessment of compensatory damages, perhaps based on what you attributed under pain and suffering or emotional distress, or loss of enjoyment of life, or other noneconomic damages awarded to the plaintiff. Once you are certain that the amount of compensatory damages has not included an amount for insult to the plaintiff’s dignity, consider what action or amount of money is appropriate to vindicate the insult or injury to the plaintiff’s personal dignity. Injuries to personal dignity, as understood here, are those injuries in which the defendant specifically targeted her misconduct toward this plaintiff with an aim of diminishing the plaintiff’s dignity. If the defendant is a corporation, consider whether the injury to the plaintiff was part

of a larger course of conduct or whether it was specifically aimed at denigrating the dignity of this particular plaintiff.

[To facilitate review of your verdict and ensure consistency across similar cases, you are required to explain the basis for your reasoning in a few sentences or more.] The remedy you choose here may be an amount of money that you determine is appropriate to alleviate this particular injury to personal dignity. Bear in mind that the plaintiff (and, depending on the circumstances, her counsel) will receive the entirety of the amount that you award under this heading.

Additionally, or alternatively, you may require the defendant to apologize to the plaintiff for the injury to the plaintiff's dignity in person or via written communication. You may also suggest other possible actions that might repair the injury to the plaintiff's dignity as supplements or substitutes.

C. *Deterrence Damages*

In some cases, extracompensatory damages are desirable to ensure that defendants do not impose costs on others that the defendants should properly bear. This is called "cost internalization." In making your assessment for promoting cost internalization, bear in mind that you are not able to extract money from the defendant for harms that happened to persons or entities who are not parties to this litigation. You may only consider whether there is a likelihood that the defendant would have escaped having to pay this plaintiff for the harm caused to the plaintiff. Other possible victims of the defendant's misconduct may bring their own suits.³⁰²

Thus, ask yourself whether the defendant might have escaped having to pay for the harm for which she should be responsible to this plaintiff. For example, if the harm was substantial, noticeable, and likely to lead to a lawsuit, your estimate of the likelihood of escaping liability would be relatively low, perhaps even zero. But if the harm might not have been attributed to the defendant, or if the defendant tried to conceal her harmful conduct, your estimate of the likelihood of escaping liability would be relatively high. You should use the table below to determine the deterrence damages multiplier that corresponds to your estimated probability of escaping liability to this par-

³⁰² Note that these instructions accord with what I view to be the correct reading of *Philip Morris*, not what I think would logically be entailed by an unbridled prioritization of cost internalization.

ticular plaintiff. Then multiply the compensatory damages amount [plus an amount, if any, for compensating personal dignity harms] by your deterrence damages multiplier. The resulting number is the base amount for deterrence damages.

The base deterrence damages amount should *not* be adjusted upward or downward because of any of the following considerations:

- (a) reprehensibility of the defendant's conduct;
- (b) net worth or income of the defendant or net profits;
- (c) gain or profit that the defendant might have obtained from his or her harmful conduct;
- (d) litigation costs borne by the plaintiff; and
- (e) whether the harm included physical injury.

**Table 1: Probability of Escaping Liability:
Deterrence Damages Multiplier**

Probability of Escaping Liability	Deterrence Damages Multiplier
0%	0
10%	0.11
20%	0.25
30%	0.43
40%	0.67
50%	1.00
60%	1.50
70%	2.33
80%	4.00
90%	9.00

* * *

In sum, if you find that the conduct at issue was undertaken with malice or recklessness, you should make a finding of reprehensibility (using the chart and its commentary as well as guidelines provided by the state) based on a scale of one to twenty. Second, you should determine an amount of aggravated damages necessary, if any, to compensate the plaintiff for personal dignity harms that were not already covered by the compensatory damages or that would not be remedied by other measures, such as an apology, that you and the court deem

appropriate. Finally, if necessary, you should recommend the amount of deterrence damages needed to pursue cost internalization of the harm and potential harm caused by this defendant to this plaintiff. Recall that other victims of the defendant's conduct might bring their own suits and that you do not need to punish the defendant or extract compensation from the defendant based on harms inflicted upon these nonparties.³⁰³ Your selection of which damages are necessary, and in what amounts, should be accompanied by an explanation of what facts you considered relevant to your determination.

³⁰³ These amounts should ultimately be adjusted to reflect appropriate tax-sensitivity judgments, as developed in a subsequent article on the taxation of punitive damages that I am writing with Gregg Polsky. Markel & Polsky, *supra* note 9.