William & Mary Law Review

Volume 59 (2017-2018)
Issue 6

Article 3

5-15-2018

Accidents and Aggregates

Lee Anne Fennell

ACCIDENTS AND AGGREGATES

LEE ANNE FENNELL*

ABSTRACT

Tort law responds to discrete, harmful events—"accidents"—by converting unruly facts into a binary on/off judgment about liability. This operation, characteristic of much of law, resembles the "thresholding" process used to convert grayscale images to black and white. It embeds decisions about how to isolate and evaluate the sample of risk-related behavior connected to the accident. This Article focuses on the implicit but powerful role that aggregation—of behavior, precautions, and events—plays in the determination of liability. These aggregative choices determine how large a slice of an injurer's conduct tort law will capture within its viewfinder, and how tight the causal connection must be between the shortfalls observed there and the accident at hand. The analysis here also sheds light on questions of legal thresholding that emerge in other doctrinal areas.

* Max Pam Professor of Law and Ronald H. Coase Research Scholar, University of Chicago Law School. For helpful comments and conversations, I thank Kenneth Abraham, Matthew Adler, Kimberly Ferzan, Nuno Garoupa, Michael Gilbert, Mark Grady, Daniel Hemel, Bert Huang, Gregory Keating, Saul Levmore, Michael Livermore, Gary Lucas, Milan Markovic, Richard McAdams, Ariel Porat, Mildred Robinson, Daria Roithmayr, David Rosenberg, Steven Shavell, Thomas Ulen, and participants in workshops at Texas A&M School of Law, the University of Chicago Law School, the University of Virginia School of Law, and the 2017 American Law and Economics Association annual meeting. Research support from the Harold J. Green Faculty Fund and the SNR Denton Fund is also gratefully acknowledged. Early drafts of this Article were circulated under the title *Lumps and Lapses in Tort Law*.

TABLE OF CONTENTS

INTRODUCTION

Tort law deals in lumps. It responds not to the innumerable fine-grained acts of risk creation that each of us performs every day, but rather to large, discrete, harmful events—"accidents."[1] And it responds to those events in a binary way, converting unruly facts into an on/off judgment about liability.[2] The fact that tort law operates at the accident level rather than at the level of risk creation presents some complications, at least if we understand tort liability as significantly directed at providing appropriate incentives for action.[3] The accident, on this view, serves as a window into risk-creating (and risk-abating) behavior, and liability represents a rough-and-ready way of addressing that behavior—a kind of accounting shortcut that focuses on a realization event. How then should law isolate and evaluate the sample of risk-related behavior connected to the accident?

This Article provides fresh traction on this foundational question by examining the underappreciated role of *evaluative aggregation* in the liability determination. I focus on three aggregative choices: (1) how much behavior to compile for purposes of assessing due care;[4] (2) how to stack together units of precaution in examining a defendant's shortfall;[5] and (3) how to factor the actual or imagined

1. *See, e.g.*, ARIEL PORAT & ALEX STEIN, TORT LIABILITY UNDER UNCERTAINTY 103-10 (2001) (contrasting liability based on risk with liability based on damage, and associating the latter with prevailing tort law). I thank Gregory Keating for comments on the lumpiness inherent in tort law's harm-based approach to liability.

2. In this respect, tort law is not unique—law often has an all-or-nothing quality. *See, e.g.*, LEO KATZ, WHY THE LAW IS SO PERVERSE 139-56 (2011) (examining the either/or nature of law); Adam J. Kolber, *Smooth and Bumpy Laws*, 102 CALIF. L. REV. 655 (2014) (analyzing the law's choice between continuous and discontinuous outcomes).

3. Tort law also serves other functions, such as compensation and loss-spreading, and can be analyzed through corrective justice and civil recourse frames that make the accident dominant for independent reasons. The deterrence function of tort law will be my focus in this Article, although I recognize that it may exist in some tension with other approaches and goals.

4. *See infra* Part I.

5. *See infra* Part II.

repetition of an interaction into liability judgments.[6] These aggregative choices carry decisive weight by determining how large a slice of an injurer's conduct tort law will capture within its viewfinder, and how tight the causal connection must be between the shortfalls observed there and the accident at hand.

Consider, for example, a point emphasized in work on "lapses"[7] or "compliance errors"[8]: It is impossible for human beings to be perfectly consistent in taking precautions that must be repeated over and over in real time, such as alertly scanning the road while driving.[9] A single moment of inattention that produces an accident might be part of a larger pattern that represents as much care as any person could reasonably be expected to exercise.[10] Alternatively, the accident-causing shortfall might be a representative draw from an urn of chronically unreasonable conduct. Tort law does not distinguish between these cases because it takes as the relevant unit of analysis the single accident-causing moment, not a larger behavioral sample that might provide corroborating or mitigating evidence about the actor's overall level of care.[11] This does not necessarily mean that tort law *should* look at broader behavioral patterns—indeed, there are reasons to question that prescription—but it does mean that the

---

6. *See infra* Part III. Scholars have previously considered a variety of other aggregation puzzles in tort and other law, including how to aggregate different elements of a cause of action. *See, e.g.*, Saul Levmore, *Conjunction and Aggregation*, 99 MICH. L. REV. 723 (2001); Ariel Porat & Eric A. Posner, *Aggregation and Law*, 122 YALE L.J. 2 (2012); *see also* Porat & Posner, *supra*, at 9 n.8 (collecting sources touching on aggregation issues). Other works emphasize how law isolates the factors that will be treated as legally relevant, rather than as background conditions. *See* H.L.A. HART & TONY HONORÉ, CAUSATION IN THE LAW 33-35 (2d ed. 1985); James M. Anderson, *The Missing Theory of Variable Selection in the Economic Analysis of Tort Law*, 2007 UTAH L. REV. 255; Lee Anne Fennell, *Property and Half-Torts*, 116 YALE L.J. 1400 (2007).

7. *See generally* Robert Cooter & Ariel Porat, *Lapses of Attention in Medical Malpractice and Road Accidents*, 15 THEORETICAL INQUIRIES LAW 329 (2014).

8. *See, e.g.*, Mark F. Grady, *Res Ipsa Loquitur and Compliance Error*, 142 U. PA. L. REV. 887 (1994).

9. *See infra* notes 42-43 and accompanying text.

10. For this reason, some of what the law considers negligence may not really constitute negligence at all—at least not in the sense of representing an inefficient failure of due care. *See generally* Mark F. Grady, *Efficient Negligence*, 87 GEO. L.J. 397 (1998).

11. *See infra* Part I.A. Of course, juries may try to distinguish the cases through back-channel ways based on inferences that they draw about the defendant's character. *See infra* note 69 and accompanying text.

choice to focus on a single moment in assessing negligence carries consequences.

The consequences mount when we consider the possibility that certain durable technologies or mechanical processes might either substitute for or complement fast-eroding human precautions like paying attention while driving.[12] The emergence of autonomous vehicles represents an especially salient current example. A machine with a known error rate will seem to naturally invite an aggregate analysis, with attention focusing on whether the overall *pattern* of outcomes could have been cost-effectively improved. If not, then using the machine may appear nonnegligent, even though the machine might hiccup now and then and cause an accident, just as a generally cautious human might lapse. If the two cases are treated differently, certain technologies might be overused or underused depending on the degree to which their deployment demands sustained human attention.[13] It is possible to address this problem without widening tort law's behavioral viewfinder, but doing so requires recognizing the unsung role of aggregation choices in shaping liability, and hence incentives.[14]

Another margin for evaluative aggregation in tort law involves stacking together or breaking apart a set of precautionary steps that the actor in question chose to forgo. Suppose, for example, a cricket ball sails out of an unfenced cricket field at an altitude of seven feet ten inches and brains a pedestrian.[15] At trial, the pedestrian's estate shows that putting up an eight-foot fence—one high enough to have prevented this accident—would have cost only $1000 but would

12. *See infra* Part I.B. Cooter and Porat focus on the possibility that liability for lapses could lead to substitution of approaches and technologies that are less safe overall. *See* Cooter & Porat, *supra* note 7, at 350-55. Mark Grady instead emphasizes complementarity between durable and nondurable precautionary technologies, with advances in the former generating increased potential for compliance errors relating to the latter. *See* Grady, *supra* note 8, at 908-09, 933-35; Mark F. Grady, *Why Are People Negligent? Technology, Nondurable Precautions, and the Medical Malpractice Explosion*, 82 NW. U. L. REV. 293, 330-31 (1988); *see also* Cooter & Porat, *supra* note 7, at 353 n.54 (noting this divergence between their work and Grady's).

13. *See infra* Part I.B.

14. *See infra* Part I.C.

15. This is a standard example, prompted by the facts of *Bolton v. Stone* [1951] AC (HL) [850] (Eng.). *See, e.g.*, WARD FARNSWORTH & MARK F. GRADY, TORTS: CASES AND QUESTIONS 146-47 (2d ed. 2009); Marcel Kahan, *Causation and Incentives to Take Care Under the Negligence Rule*, 18 J. LEGAL STUD. 427, 428-29 (1989).

have saved an expected $1200 in accident costs over the useful life of the fence. This might look like an open-and-shut win for the plaintiff, at least if the jurisdiction follows the cost-benefit approach to negligence captured in the Hand formula.[16]

But not so fast. There is nothing inevitable about treating the full eight feet of fencing as an indivisible unit when analyzing the cricket club's behavior. Once we disaggregate that single, lumpy, all-or-nothing choice into incremental choices about fence heights, the simple case starts to look less airtight. Suppose the first six feet of fence height are really worthwhile, delivering $1100 in accident savings while costing only $800 in lumber and labor, but the last two feet require an extra $200 in construction costs only to save a marginal $100 in accident costs.[17] Even though it was negligent for the cricket club not to build a six-foot fence, it would not have been negligent for the club to stop at six feet—and a six-foot fence would not have stopped the fateful ball. On this account, the club's negligence (its failure to build the optimal six-foot fence) did not cause the accident. The correct doctrinal result would seem to be no liability.[18]

Yet again, not so fast. We must still consider how these aggregation decisions interact with an important architectural feature of tort law—the fact that liability falls to zero at the point of due care under a negligence standard.[19] One implication of that architecture has been well recognized: the possibility that a defendant would face a behavior-distorting cliff of liability if held to account not only for the harm that occurs *because* he is negligent but also for all of "the harm that occurs *when* he is negligent."[20] But there is another

16. The Hand formula was famously articulated by Judge Learned Hand in *United States v. Carroll Towing Co.*, 159 F.2d 169 (2d Cir. 1947). The formula calls for comparing the burden (*B*) of an untaken precaution (such as the unbuilt fence) against the expected probability of injury (*P*) multiplied by the expected magnitude of loss (*L*). *See id.* at 173. If *B* is less than *P* times *L*, it is negligent not to undertake the precaution. *See id.*

17. *See* FARNSWORTH & GRADY, *supra* note 15, at 147-48 (providing similar variations on the fence problem); *see also* RICHARD A. EPSTEIN & CATHERINE M. SHARKEY, CASES AND MATERIALS ON TORTS 175 (11th ed. 2016) (presenting similar hypotheticals that demonstrate the problem of "choosing the correct interval for assessing defendant's conduct").

18. *See* Kahan, *supra* note 15, at 429 ("[A]s a matter of common law, an injurer is only liable for accidents caused by his negligence. Therefore, the owner would *not* be liable for injuries from balls flying over the fence at heights exceeding [the efficient height].").

19. *See infra* Part II.B.2.

20. ROBERT D. COOTER & ARIEL PORAT, GETTING INCENTIVES RIGHT: IMPROVING TORTS,

implication of the negligence regime that has been widely ignored, though it also bears on whether the cricket club should be let off the hook. The law plunks a flat ledge of zero liability across the entire range of conduct falling beyond due care—a range of conduct that, in fact, generates different real-world accident rates.[21] This ledge also has distortive effects, and these distortions will ultimately require us to reverse our earlier intuition about whether to analytically disaggregate the unbuilt cricket fence.[22] The precaution-aggregation choice turns out to be a crucial lever for optimally adjusting the required causal relationship between negligence and harm.[23]

Finally, tort law contains some important puzzles that only become visible when a particular interaction is repeated many times (whether as a matter of fact or as a conceptual exercise). Deciding whether and how to "scale up" the liability analysis constitutes another domain for implicit aggregation choices. Repetition can reveal distortions and injustices that are muted at the individual-accident level—or it can do the opposite, washing out apparent anomalies.[24] Because statistical risks and expected payoffs become more meaningful and tractable under large-number conditions, results that appear intolerable at close range—a large chunk of liability for a relatively trivial act of negligence, say—may look more acceptable once we zoom out to capture a larger set of similar

---

CONTRACTS, AND RESTITUTION 26 (2014); *see also id.* at 17-31; Mark F. Grady, *A New Positive Economic Theory of Negligence*, 92 YALE L.J. 799, 804 (1983) (referring to these two possibilities as "the P*-cutoff rule" and "the full liability rule," respectively); Richard W. Wright, *The New Old Efficiency Theories of Causation and Liability*, 7 J. TORT L. 65, 84 (2014) (referring to the harming-because-negligent rule as "the Optimal Care rule"). Put in Robert Cooter's terms, charging a defendant for all harm that occurs when she is negligent would impose a discontinuous "sanction" for failing to exercise due care, not merely a "price" that taxes negligent conduct at the margin. *See generally* Robert Cooter, *Prices and Sanctions*, 84 COLUM. L. REV. 1523 (1984).

21. This might seem irrelevant, given that precaution costs exceed savings in accident costs across the entire range. But it turns out to matter under real-world conditions of uncertainty, for reasons Mark Grady previously articulated, *see generally* Grady, *supra* note 20, and that I revisit in depth below, *see infra* Parts II.A-B.

22. This conclusion follows from Grady, *supra* note 20, although the distortive "ledge effect"(as I call it) does not seem to have been widely appreciated. Here, I reframe the analysis to focus on the aggregation decision and show how it operates when decisions are chunky rather than continuous. *See infra* Part II.

23. *See infra* Part II.B.

24. *See infra* Part III.A.

interactions.[25] Here, it becomes important to consider whether insurance or specialized doctrines can synthetically replicate large-number conditions for individuals.[26]

Repetition may instead compound rather than counterbalance systemic shortfalls in liability patterns. Consider the negligence of doctors in subspecialties where patients routinely face high background risks of death. A doctor who negligently treats a population of patients who are overwhelmingly likely to die in any event will always be let off the hook by a more-likely-than-not standard, even if her negligence caused, say, twenty out of one hundred observed deaths. Although scholars have proposed various approaches to this well-recognized problem,[27] one underappreciated alternative deserves attention: instead of asking whether a given harm was more likely *than not* caused by the doctor's negligence, ask whether the harm was more likely to have been caused by the doctor's negligence *than the other harms* in the conceptual set generated by repetition of the interaction.

Here, we can take a page from "thresholding" in image manipulation.[28] When a grayscale image is converted into black and white, a continuous variable (shading) must be translated, pixel by pixel, into binary results.[29] A globally applied threshold will produce

---

25. This point tracks one that is made in the economic literature—that problems of lumpiness or indivisibility become less troublesome in high-volume contexts. *See, e.g.*, KENNETH J. ARROW & F.H. HAHN, GENERAL COMPETITIVE ANALYSIS 62 (1971) (noting that the "economic significance" of indivisibilities "is relatively less when the number of units is large" and observing that "the difference between one stamping mill and none is important, but if the relevant choice is between 100 and 101 shovels, the assumption of divisibility is unlikely to be seriously misleading"); HAGEN BOBZIN, INDIVISIBILITIES: MICROECONOMIC THEORY WITH RESPECT TO INDIVISIBLE GOODS AND FACTORS 2 (1998) ("The difference between the production of 100 000 or 100 001 cars is of little significance for an automobile company, whereas a household faces considerable consequences depending on whether it has got a car or not.").

26. *See infra* Part I.B.

27. *See infra* Part III.C.

28. *See, e.g.*, Alan C. Bovik, *Basic Binary Image Processing*, *in* HANDBOOK OF IMAGE AND VIDEO PROCESSING 39, 39-43 (Al Bovik ed., 2d ed. 2005).

29. *See id.* To the extent that law similarly renders continuous variables into binaries, it operates as a thresholding enterprise. *See, e.g.*, KATZ, *supra* note 2, at 157-81 (using the example of the legal treatment of death to illustrate the law's reliance on binary categories); Adam J. Kolber, *Smoothing Vague Laws*, *in* VAGUENESS AND LAW: PHILOSOPHICAL AND LEGAL PERSPECTIVES 275, 281 (Geert Keil & Ralf Poscher eds., 2016) ("[W]henever we use a bumpy law to govern a smooth phenomenon, we are rounding a continuous result to some nearby discrete option.").

unacceptable results where, for example, light or shadow falls across a portion of the image, making all pixels in a given region darker or lighter than the threshold that works best elsewhere in the image.[30] The key to successfully picking out foreground from background is to see which pixels are *local* standouts—hence, thresholding methods examine shapes, clusters, and pixel neighborhoods in order to determine the appropriate local threshold.[31] Translated into the tort realm, this would suggest lowering the liability threshold in contexts with high background risk to pick out those instances that were most likely to have been caused by the doctor's negligence.[32]

All these aggregation issues arise because tort law reaches behavior through the narrow and hazy window of observation afforded by the individual accident. If all instances of risk creation and mitigation could be perfectly tracked and priced, nothing special would transpire when one of those acts manifested in an accident, and there would be no need to worry about how to isolate the behavior relevant to the accident or assess any causal connections.[33] Likewise, the lumpiness of accidents would matter little if actors (and their insurers) could respond to *expected* costs that exhibited a predictable sensitivity to changes in behavioral inputs.[34] But this is exactly the problem: accidents demand binary liability responses that turn on just how the accident-relevant behavior is isolated, sliced, and analyzed—in other words, on issues of evaluative aggregation.

---

30. *See* Robert Fisher et al., *Adaptive Thresholding*, HYPERMEDIA IMAGE PROCESSING REFERENCE (2003), http://homepages.inf.ed.ac.uk/rbf/HIPR2/adpthrsh.htm [https://perma.cc/VU24-9LAE] (describing and depicting the use of "[l]ocal adaptive thresholding" to address situations in which "a strong illumination gradient" makes global thresholding perform poorly).

31. *See, e.g.*, Bovik, *supra* note 28, at 43-55 (discussing use of "*region correction* algorithms" and related approaches to address shortcomings of thresholding); Fisher et al., *supra* note 30 (describing techniques of "adaptive thresholding" that consider the local pixel neighborhood in setting thresholds and allow the threshold to vary over the image accordingly).

32. *See infra* Part III.C.

33. Any real-world effort at pricing risk creation would inevitably rely on imperfect proxies, however, introducing other aggregation issues, such as those associated with regulatory line-drawing. I thank Kimberly Ferzan for this point.

34. *See* Mark F. Grady, *Discontinuities and Information Burdens*, 56 GEO. WASH. L. REV. 658, 659 (1988) (reviewing WILLIAM M. LANDES & RICHARD A. POSNER, THE ECONOMIC STRUCTURE OF TORT LAW (1987)) (observing that the "abrupt" relationship between lapses and liability "does not necessarily create a discontinuity in *expected* liability").

Although this Article focuses on the way that aggregation issues play out in the torts field, related issues of aggregation run through all of law.[35] Property scholars wrestle with the "denominator problem" in regulatory takings doctrine, which likewise requires determining the proper unit of analysis against which to assess the impact of a particular governmental action.[36] Conceptually similar problems crop up in a variety of other fields, from constitutional law[37] to copyright[38] to criminal law.[39] Likewise, the broad run of legislative line-drawing exercises implicitly depend on—and can strategically manipulate—aggregation decisions.[40] Thus, the analysis here carries implications for other issues involving evaluative aggregation.[41]

The Article proceeds in four Parts. Part I focuses on the slice of behavior used to assess negligence liability. Part II considers how the aggregation of (untaken) precautionary steps interacts with tort law's liability structure. Part III examines the significance of event repetition where uncertainty exists about causation. Part IV turns to connections between these tort law aggregation problems and related problems that arise in other doctrinal areas.

---

35. I am exploring a number of these other contexts, and the connections between them, in a book currently in progress. LEE ANNE FENNELL, SLICES AND LUMPS: DIVISION AND AGGREGATION IN LAW AND LIFE (forthcoming 2019).

36. *See, e.g.*, Lucas v. S.C. Coastal Council, 505 U.S. 1003, 1016 n.7 (1992) ("Regrettably, the rhetorical force of our 'deprivation of all economically feasible use' rule is greater than its precision, since the rule does not make clear the 'property interest' against which the loss of value is to be measured."); *see also infra* Part IV.C.

37. *See* Daryl J. Levinson, *Framing Transactions in Constitutional Law*, 111 YALE L.J. 1311 (2002).

38. *See* Margot E. Kaminski & Guy A. Rub, *Copyright's Framing Problem*, 64 UCLA L. REV. 1102 (2017).

39. *See* Alon Harel & Ariel Porat, *Aggregating Probabilities Across Cases: Criminal Responsibility for Unspecified Offenses*, 94 MINN. L. REV. 261 (2009).

40. Line-drawing sometimes occurs quite literally, as through zoning, annexation, districting, and choosing areas for condemnation. Manipulating literal or figurative boundaries to make things come out right on net offers strategic possibilities because it allows bundling in negative value increments along with positive value ones, where the surplus associated with the former is sufficient to absorb the deficits associated with the latter. *See* Levinson, *supra* note 37, at 1326-32; *see also infra* Part IV.B.

41. *See infra* Part IV.

I. LUMPING LAPSES

Human beings cannot avoid occasionally falling short in their efforts to take due care. As Tony Honoré put it, "[I]n no activity or walk of life can people consistently maintain the high standard of skill and care required by law without variation."[42] Driving offers the most familiar example. Anyone who drives very often will experience the occasional lapse of attention.[43] While usually harmless and quickly forgotten, these lapses sometimes produce an alarming near miss, and in a tragic few cases, a terrible accident. How should these shortfalls be treated? Answering this inquiry requires making aggregation choices, at least implicitly.

A. Lapses and Accidents

Two dueling concerns surround the treatment of momentary lapses: the possibility that the observed lapse is an outlier, and the possibility that the observed lapse is representative. The first possibility makes liability for lapses unpalatable, while the second possibility makes exempting them problematic.[44] Both concerns are a product of tort law's focus on accidents rather than on risk creation as such. Because most lapses do not eventuate in accidents, the fact that a lapse happens to be causally connected to an accident already makes it an outlier among lapses. That fact alone cannot be a reason to exempt the actor from liability, at least if we wish to retain the accident-based structure of tort liability.[45] What we might

42. Tony Honoré, *Responsibility and Luck*, 104 LAW Q. REV. 530, 549-50 (1988); *see also* Jeffrey J. Rachlinski, *Misunderstanding Ability, Misallocating Responsibility*, 68 BROOK. L. REV. 1055, 1062 (2003) ("[P]eople's attention often lapses in the face of monotonous, albeit dangerous tasks. Despite tort law's requirement of reasonableness, it is difficult to maintain focus on a repetitive task.").

43. *See, e.g.*, Grady, *supra* note 8, at 900 ("It is impossible to drive a car for any period of time without missing a required precaution."). As Grady explains, the problem can be characterized in terms of the prohibitively high cost of achieving consistency in compliance. *See id.* at 899 ("[P]eople face a cost of consistent performance that is greater than the sum of the cost of all individual trials.").

44. *See, e.g.*, Cooter & Porat, *supra* note 7, at 331 & n.2; *see also* Grady, *supra* note 10, at 401-02.

45. There are alternatives, such as the New Zealand system for funding recoveries by accident victims. *See generally* Peter H. Schuck, *Tort Reform, Kiwi-Style*, 27 YALE L. & POL'Y

want to know, however, is whether the lapse is *also* a behavioral outlier for this actor.

Table 1 illustrates the possibilities by dividing actors into four categories based on their lapse frequency and accident outcomes.

Table 1. Lapse Rates and Outcomes

	Commonly Lapses	**Rarely Lapses**
Causes an Accident	I	II
Proceeds Uneventfully	III	IV

Tort law breaks the *rows* of Table 1 apart, but collapses the *columns*. The requirement that risk eventuate in harm means that only the lapsers occupying Cells I and II face potential liability. And the law's focus on the accident-causing moment rather than on a larger sample of the defendant's behavior means that Cells I and II are treated identically.[46] If actors were charged for their lapses rather than for the accidents that they caused, Table 1's two rows would collapse into one, but the columns would be broken apart for different treatment: all lapses would generate a payment obligation, but those who rarely lapse (II and IV) would pay less often than those who commonly lapse (I and III).

Examining a larger sample of behavior when making the negligence determination (assuming that this could be done) would also break apart the two columns by relieving Cell II actors of liability altogether. This would move tort law a half measure toward a risk-based model. Yet as long as tort law's primary distinction between the rows persists, liability still fails to track risk creation, and luck continues to govern individual outcomes.

Freeing Cell II lapsers from liability without any compensating adjustments to the other cells also means that injurers in the aggregate will face lower expected accident costs, and victims in the

---

REV. 187 (2008). But for purposes of the discussion here I assume that this element of the tort framework is fixed.

46. This, at least, is the doctrinal rule. Juries may in fact "forgive" lapses that they believe to be rare. *See, e.g.*, Mark F. Grady, Breach of Duty 15-20 (2017) (unpublished manuscript) (on file with author); *see also infra* note 69 and accompanying text.

aggregate must bear more of their own losses.[47] This may seem unproblematic if we suppose that the exempted accident-causers in Cell II should never have been considered negligent in the first place, so that the current legal approach is overcharging injurers as a whole by sweeping in some nonnegligent parties along with the negligent ones. But if so, why stop there? Surely some of the lapses that are committed by the common lapsers in Cell I could have been committed by even the most careful person—just not all of them. Yet we have no way of telling an "excess" lapse from one that comes from a human being's unavoidable allotment, when assessing the cause of an accident.[48] If we tolerate liability for unavoidable lapses when they are mixed in with avoidable ones committed by the same person, why not otherwise?

Moreover, might not charging *all* actors for these unavoidable lapses make up for other shortfalls of an accident-based negligence regime? A well-known concern about the negligence standard is that injurers will engage in risky activities to an excessive extent because they will never bear liability as long as they are sufficiently careful—even though their elevated activity level raises accident costs for victims.[49] This activity-level problem assumes that injurers can always comply with the negligence standard. But if lapses occur in proportion to activity levels, then it is impossible for an injurer to increase her activity level without also increasing her chance of being held liable for a lapse.[50] Thus, liability for lapses could help to check the tendency of injurers to overengage in risky activities by effectively taxing (in expected-value terms) the activity in proportion to its volume.[51]

---

47. Throughout the Article, I assume that victims and injurers are disjoint sets and focus only on the precautions available to injurers. This unilateral precaution assumption, although plainly unrealistic in many settings, simplifies the exposition to focus attention on the puzzles of aggregation explored here.

48. Grady alludes to this problem when he observes that "[m]any slips are like so many peas in a pod; the efficient ones look the same as the uneconomic ones." Grady, *supra* note 8, at 905-06.

49. *See* STEVEN SHAVELL, ECONOMIC ANALYSIS OF ACCIDENT LAW 21-26 (1987). The converse concern attaches to a strict liability regime. Here, *victims* are likely to engage in excessive levels of activity, even though this raises costs for injurers. *See id.* at 26-32.

50. *See* Grady, *supra* note 12, at 309 (observing that liability for compliance errors can check activity levels).

51. *See id.* ("A rule that forgave reasonable memory lapses would be much less effective than the actual rule in controlling activity levels."); *see also* Mark F. Grady, Justice Luck in

With these points in mind, we can turn to the question of whether the individual accident-causing moment is the right or wrong unit of analysis upon which to base liability.[52] Answering this question requires considering both whether this is a suitable or unsuitable way of linking liability to human behavior as an absolute matter, and whether it introduces a distortive inconsistency in the relative treatment of different categories of actors or actions. I will start with the latter question, which has formed the basis for critiques of the law's treatment of lapses,[53] before circling back to the former.[54] Although I conclude that inconsistent aggregation choices can indeed produce distortions, there is more than one way to resolve differences in relative treatment—and there is a reasonable case for doing so in a way that retains the law's current treatment of lapses.

B. Distorted Decision-Making

Robert Cooter and Ariel Porat have flagged the potential for the legal treatment of human lapses to produce behavioral distortions, if the law treats the shortcomings of other kinds of precautionary technologies more leniently.[55] This divergent treatment boils down

---

Negligence Law 20 (Feb. 12, 2018) (unpublished manuscript), https://papers.ssrn.com/sol3/papers.cfm?abstract_ id=3124996 [https://perma.cc/33UQ-ZTGG] (observing that current law "basically imposes a 'stochastic tax' on" lapses that could not be cost-effectively avoided).

52. Grady characterizes the problem of efficient lapses as one of an insufficient period of observation but sees practical difficulties with expanding the time scope. *See, e.g.*, Grady, *supra* note 10, at 400-02. Some of the ways that Cooter and Porat would implement a lapse defense would effectively expand the observation period, either synthetically through probabilistic reasoning or through examining repeated lapses within a particular time frame. *See* COOTER & PORAT, *supra* note 20, at 70-72.

53. *See infra* Part I.B.

54. *See infra* Part I.C.

55. Cooter & Porat, *supra* note 7, at 350-55. They also observe that activity levels will be suppressed for lapse-prone activities that generate positive externalities. *See id.* at 348-50; *see also* Keith N. Hylton, *Duty in Tort Law: An Economic Approach*, 75 FORDHAM L. REV. 1501, 1505-10 (2006) (suggesting that liability may be calibrated in ways that account for externalized benefits as well as externalized costs of behavior). This point is answered well by an argument Steven Shavell has recently made in urging an expansion of the domain of strict liability: if activities produce positive externalities, then actors should be subsidized in accordance with those positive externalities, not relieved of liability for some set of accidents that they produce. *See* Steven Shavell, Why Strict Liability Should Apply to All Dangerous Activities—Both Common and Uncommon 37-39 (Jan. 13, 2017) (unpublished manuscript) (on file with author). It would be happenstance if relief from liability happened to match up to positive externalities, and separating out these elements allows for better incentives to take

to differences in the aggregation of behavioral samples for evaluation. Suppose, to use an example from Ward Farnsworth and Mark Grady, that a surgeon will accidentally leave a sponge inside a patient in one out of every one million surgeries.[56] When that one-in-a-million case occurs, the doctor will appear negligent for failing to undertake a simple check that would have avoided a significant risk of loss.[57] Now suppose a machine can be purchased that keeps track of sponges used in surgery and counts them as they are removed. If the machine malfunctions one time per million surgeries, should its lapse be treated the same as or differently from the lapse of the surgeon?[58]

Cooter and Porat suggest that machine errors are likely to be treated differently from human errors.[59] Whereas the surgeon's mistake will be evaluated in isolation by asking whether it would have been cost effective for her to spend an extra moment ensuring that she had all the sponges out before closing up the patient (a question that will always be answered in the affirmative), the machine's error will be examined in terms of whether it would be cost effective to design a machine that had a lower error rate (to which the answer may well be negative). Although they do not put it in quite these terms, Cooter and Porat suggest that actors may choose mechanized precautions that will be assessed in the aggregate over individualized human actions that will be assessed one at a time, even if the cost-benefit ratio is better for the human precautions than for the mechanized precautions.[60]

The activity-level issue raised above only heightens the problem that Cooter and Porat identify. A party who cannot avoid using a technology—human judgment—that will generate actionable lapses as she increases the volume of a given activity may indeed curtail

---

care along dimensions that are not well policed by the negligence system. *See id.* at 38. One of those dimensions is activity levels, which liability for lapses might help control. *See supra* notes 49-51 and accompanying text.

56. FARNSWORTH & GRADY, *supra* note 15, at 158.

57. *Id.* ("[I]f a surgeon mistakenly leaves a sponge inside a patient, there is no room for him to argue that in fact he is a very careful person and that this was a once-in-a-lifetime slipup.").

58. *See id.* (asking but not answering this question).

59. Cooter & Porat, *supra* note 7, at 352-53.

60. *See id.* at 350-55.

her activity level.[61] But a party who can substitute a mechanized precautionary technology for a judgment-intensive approach will have no such incentive.[62] Rather, she can be certain that no matter how high her activity level, the technology will always protect her from liability if its overall low failure rate is deemed to meet the due-care standard.[63] Thus a driver who relies on cruise control (or, soon, automated cars) may be able to increase the number of miles driven without fear of expanded liability, while a driver whose own decisions about speed or other factors will be evaluated one by one could not do so.[64]

The distortion that concerns Cooter and Porat, then, is foundationally a mismatch between the degree of aggregation that is employed in evaluating due care in human actions versus that which is employed in evaluating due care in automated processes.[65] They provide examples of technologies—cruise control, traffic lights, and so on—that generate patterns of results over time.[66] Because the full pattern can be observed, errors that would look like lapses in the individual human actor case are not picked out and evaluated in isolation but are instead aggregated with the far more prevalent

61. *See supra* notes 49-51 and accompanying text.

62. A wrinkle explored at length by Grady is the potential complementarity between technological developments and "nondurable" precautions like checking dials and instruments—the omission of which amounts to compliance errors. *See generally* Grady, *supra* note 12. Grady hypothesizes that these complementarities are ubiquitous and typically cause technological developments to increase rather than decrease compliance errors and associated findings of negligence. *See id.* at 311-14. If all technologies involved such complementarities and heightened risks of lapses, then the concern about substitution raised by Cooter and Porat would not occur—instead, substitution away from new technologies would be expected because these technologies would introduce more opportunities for human beings to fall short. However, Grady also discusses resort to what he calls "risk dumping" technologies, which seek to overcome or compensate for the human shortcomings that generate compliance errors. *See id.* at 297-98, 309-13, 334.

63. *See infra* note 96 and accompanying text.

64. *See, e.g.*, COOTER & PORAT, *supra* note 20, at 61-62.

65. Cooter and Porat's concern about substitution toward machines and away from humans is only one of the distortions that they discuss. They are also concerned with substitution away from lapse-prone procedures toward procedures that do not invite lapses but that are riskier overall. *See id.* at 69-70 (giving examples of alternative ways of delivering babies and treating cancer). These cases can also be viewed as problems of aggregation: the choice *between* procedures is a bundled one, but the choices made *within* a given procedure are considered à la carte. *See infra* note 96.

66. *See* COOTER & PORAT, *supra* note 20, at 61-65.

instances of correct operation for purposes of evaluating whether the technology meets the due-care standard.

Importantly, such technologies embed a verification function: they ensure that we see the full universe of acts, and that we can observe the true ratio of good outcomes to bad outcomes. The hiccup of a machine with a known error rate differs from an observed human failing in that it is possible to immediately see—and verify—the error's position within a larger pattern of outcomes.[67] In the case of a human lapse, by contrast, fact-finders are observing a single draw from an opaque urn of behavior, a single frame from an ongoing movie that is otherwise hidden from view. Jurors will be uncertain whether the behavior they are seeing is indeed a rare lapse, analogous to the error of a generally sound machine, or whether they are instead glimpsing one instance of an overall pattern of unreasonably risky behavior.[68] For this reason, it is perhaps unsurprising that jurors would latch onto even legally irrelevant cues that might help them form a judgment of the defendant's character.[69]

To remove the distortion that presently exists between different modes of precaution, Cooter and Porat recommend a kind of "lapse defense" that would effectively permit defendants to rely on aggregated information to demonstrate that they were suffering from an atypical lapse rather than a typical shortfall.[70] Taken to its logical

---

67. On this information problem, see, for example, Grady, *supra* note 12, at 306-07. Another difference is that machines, unlike people, lack awareness of their own track record and cannot strategically decide that they can "afford" a slip. I thank Michael Gilbert for this point.

68. This is why, in Grady's view, allowing a lapse defense would amount to "taking character evidence." Grady, *supra* note 10, at 402. Another concern is whether we can get good insight into lapse rates in a manner that does not permit gaming. *See id.* at 401-02.

69. *See* Janice Nadler, *Blaming as a Social Process: The Influence of Character and Moral Emotion on Blame*, 75 LAW & CONTEMP. PROBS., no. 2, 2012, at 1, 15-31 (presenting and discussing experimental results suggesting that attributions of responsibility for harm are sensitive to extraneous information bearing on an actor's perceived moral character); Grady, *supra* note 51, at 6-12, 19 (observing that courts allow juries to forgive lapsing defendants, and that "the jury will look such a defendant in the eye" to decide whether forgiveness is appropriate).

70. *See* COOTER & PORAT, *supra* note 20, at 60-73. Grady previously considered such an approach, but appears to have largely rejected it as impractical. *See* Grady, *supra* note 10, at 401-02 (considering the possibility that "the court would expand the relevant period of time during which the actor's behavior is analyzed" to account for inevitable (and thus efficient) lapses, but suggesting this would be impractical because it "essentially involves taking character evidence, which could be extremely self-serving to the defendant"); *see also* Grady,

conclusion, this would amount to a change in the unit of analysis for assessing whether due care has been taken. It would no longer be possible to answer that question based on observation of a one-off event, but would instead require looking at a larger sample of behavior. But how would this approach be implemented?

One possibility would be to focus on some facet of the defendant's behavior that operates at a higher level of aggregation than the individual lapsing moment—what Cooter and Porat term "second-order precautions."[71] Examples might include routines, plans, or checklists that the defendant follows as standard operating procedure, even if there was a shortfall on a particular occasion. Grady has suggested that courts already tend to treat "precaution plans" differently from individual precautions and will, for example, view more leniently a surgeon who has adopted a plan of counting sponges but lapses on a single occasion than a surgeon who lapses in the absence of such a plan.[72] An explicit lapse defense could build on this approach, although it might induce wasteful expenditures on formalizing plans that are not consistently followed or that would perhaps even be inefficient to follow.

Another avenue for obtaining a larger sample of injurer behavior is suggested by the immense amounts of data that smartphones can collect (and indeed already collect). For example, some insurance companies offer discounts to drivers who submit to smartphone-enabled monitoring of their driving habits.[73] The data collected on risky behaviors like hard acceleration and braking, speeding, and erratic movements currently allow premiums to better reflect risk

---

supra note 8, at 905 (observing that a system in which "an erring driver or surgeon could have friends and colleagues testify that the lapse in question was reasonable given the defendant's normally careful habits" would carry a large administrative cost, although one that might perhaps eventually become worth bearing); Grady, *supra* note 12, at 333 ("A more modest reform would be a rule that allowed a doctor to prove that, notwithstanding his inadvertent negligence on the occasion in question, he was maintaining a reasonable rate of compliance.").

71. Cooter & Porat, *supra* note 7.

72. *See* Grady, *supra* note 46, at 38-39.

73. *See, e.g.*, *Drivewise from Allstate*, ALLSTATE INS. COMPANY, https://www.allstate.com/drive-wise.aspx [https://perma.cc/CK6W-QWDX]; *Drive Safe & Save Mobile*, STATE FARM, https://www.statefarm.com/insurance/auto/discounts/drive-safe-save/mobile-app [https://perma.cc/L4NQ-W9CK].

profiles.[74] But these data could easily be repurposed in service of a lapse defense.[75]

An interesting wrinkle surrounding these new monitoring systems is their capacity to track miles driven.[76] This is exactly the sort of activity-level information that the negligence standard currently ignores. Having it more granularly priced into insurance premiums could improve incentives for drivers. But whether it will indeed be properly priced into liability premiums would seem to be endogenous to the legal treatment of lapses. If lapses give rise to liability, as they do presently, then the pricing of liability premiums should reflect the fact that any driver who drives a lot will unavoidably lapse from time to time and may cause an accident. If a lapse defense became available, however, then it would no longer be the case that driving a great deal would translate into greater expected liability due to unavoidable lapses.[77] On the contrary, it might seem that driving a large number of miles could be *helpful* to the lapsing driver's case because it would increase the base against which a singular lapse would be assessed.

Another possibility for amassing a larger behavioral sample might be a variation on the information-escrowing approach that Ian Ayres and Cait Unkovic have proposed for sexual harassment and other types of wrongdoing,[78] or the "How's My Driving" approach that Lior Strahilevitz has recommended for driving and

74. *See* ALLSTATE INS. COMPANY, *supra* note 73; STATE FARM, *supra* note 73.

75. Such an approach could work a hardship on those who do not own a smartphone or are unwilling to have their every movement tracked. Submitting information about past behavior would presumably be voluntary, but a fact-finder could draw a negative inference if it were not supplied.

76. *See, e.g.*, STATE FARM, *supra* note 73 ("The fewer miles you drive and the safer you drive, the more you could save on auto insurance.").

77. The discussion in the text refers just to *liability* premiums (third-party insurance). We would expect the mileage information to be priced into *collision* premiums (for damage to the insured's own car) regardless of the legal regime because this exposure exists independent of liability judgments. Indeed, if a lapse defense were widely available, we would expect collision premiums to be even more sensitive to miles driven because there would be fewer instances in which another driver would be liable for the damage to the insured's vehicle, and more time on the road would mean more exposure to other drivers' lapses.

78. Ian Ayres & Cait Unkovic, *Information Escrows*, 111 MICH. L. REV. 145 (2012). For an earlier discussion of a similar approach inspired by the "recorder" system used in competitive contract bridge, see Frederick Schauer & Richard Zeckhauser, *On the Degree of Confidence for Adverse Decisions*, 25 J. LEGAL STUD. 27, 48-52 (1996).

other behaviors.[79] A record that is clean of complaints over an extended period in which reporting is widely available could be used to show that a given lapse was an aberration. Again, however, we might worry if extensive engagement in an activity like driving were considered an exonerating factor, given its contribution to risk creation. Of course, miles driven could also be the basis of a Pigouvian tax or similar measure designed to overcome such distortions—though this moves us a step away from an accident-centric model.

More broadly, making current liability turn on past risk creation would attenuate the causal relationship upon which tort law is premised; liability would depend on risk creation over time, and not just on how risk was created in the incident before the court. This is not necessarily a bad thing, but it does represent a shift from a model that relies on pairing up injurers and victims on an accident-by-accident basis. Due care would no longer be evaluated solely based on a defendant's behavior toward a specific injured counterparty, but would instead be evaluated across time based on her interactions with innumerable potential counterparties. Contrary to the standard assumptions of tort law, proof of negligence *would* be found "in the air"[80]—or at least in the modus operandi of the defendant.

C. A Disaggregating Alternative

There is another way to alleviate the imbalance between the law's treatment of human lapses and other kinds of failure rates: by imposing strict liability for accidents caused when mechanized or routinized systems are used to control particular operations.[81] This

79. Lior Jacob Strahilevitz, *"How's My Driving?" for Everyone (and Everything?)*, 81 N.Y.U. L. REV. 1699 (2006).

80. *See* Palsgraf v. Long Island R.R. Co., 162 N.E. 99, 99 (N.Y. 1928) (Cardozo, C.J.) ("Proof of negligence in the air, so to speak, will not do." (quoting FREDERICK POLLOCK, THE LAW OF TORTS 455 (11th ed. 1920))).

81. This is how products liability already works in the context of manufacturing defects. The individual defective product forms the unit of analysis for establishing liability, and it is no defense that due care was taken in mass-production techniques. *See* John C.P. Goldberg & Benjamin C. Zipursky, *The Strict Liability in Fault and the Fault in Strict Liability*, 85 FORDHAM L. REV. 743, 773-74 (2016) (analogizing the inevitable errors of mass production to the inevitable errors of a driver or surgeon over a long enough span of time).

is the inverse of Cooter and Porat's proposal. Instead of compiling a larger sample of human behavior, the "behavior" of the machine or procedure would be effectively disaggregated by treating each failure as a basis for liability. This would remove the distortion as effectively as would the creation of a lapse defense, because there would no longer be an incentive to substitute an automated or routine process solely for the purpose of taking advantage of evaluative aggregation.[82] If the mechanical procedure were really superior to a particular set of human decisions, however, it would still be employed. Negligence would remain the standard for human-mediated actions, but human lapses would still give rise to liability in the manner that currently occurs, with the attendant (salutary) pressure on activity levels.[83]

This approach raises two questions. First, can the law justify holding people responsible for lapses that no one could possibly avoid? Second, can the application of a nominally divergent standard for human and automated acts be justified? The key to the first question is to see a negligence standard as a safe harbor from liability that the law extends to actors under certain instrumentally defined circumstances,[84] not an inalienable human right to inflict all the harm on others that due care would not prevent (or even all the harm that one cannot personally help causing). The answer to the second question turns on the implicit role of aggregation and disaggregation in evaluating care, which translates nominally divergent standards into substantively equivalent ones.

To start, consider liability for human lapses, which is currently the law,[85] and which follows from the practice of evaluating each lapse in isolation. This does mean holding actors responsible for harm that they could not help causing—conditional on having

82. *See supra* text accompanying notes 59-64.

83. *See* Grady, *supra* note 12, at 309.

84. For one discussion of the implications of safe harbors in tort law, see Jason Scott Johnston, *Uncertainty, Chaos, and the Torts Process: An Economic Analysis of Legal Form*, 76 CORNELL L. REV. 341, 355-56, 364 (1991). For a recent illuminating discussion of safe harbors in general (and their converse, areas of per se liability), see generally Susan C. Morse, *Safe Harbors, Sure Shipwrecks*, 49 U.C. DAVIS L. REV. 1385 (2016).

85. There are arguably some limited exceptions in the law, such as for emergencies or certain momentary distractions. *See* COOTER & PORAT, *supra* note 20, at 63-66; Grady, *supra* note 8, at 901-02. But these do not serve to generally excuse the shortfalls in attention or forgotten precautions that plague many repetitive tasks.

chosen to engage in the activity in question at a given level. But holding people responsible for shortcomings they could not prevent is already an entrenched and pervasive feature of our negligence regime, which does not tailor liability to match the particular abilities of different actors.[86] Consider the inherently clumsy person who cannot meet the ordinary standard of reasonableness, but who is nonetheless held liable if his clumsiness causes harm.[87] As a number of scholars have noted, the law's failure to tailor the due-care standard introduces a "pocket of strict liability" into the negligence system.[88] Where an actor cannot reach the safe harbor of due care, she is required to cover the (unavoidable, for her) costs of her actions.

An actor who cannot meet the standard of due care for a given activity may therefore find herself forced, as a liability-limiting measure, to reduce her participation in the activity or drop out of it

---

86. *See, e.g.*, Omri Ben-Shahar & Ariel Porat, *Personalizing Negligence Law*, 91 N.Y.U. L. REV. 627, 636 (2016) ("Current law does not personalize standards of care. It adheres, instead, to a regime of uniform, nonpersonalized standards."). Some limited tailoring does occur, as for children engaged in children's activities, but there is no general scaling of the due-care standard to match the particular abilities of each actor. *See, e.g.*, *id.* at 637-44 (discussing instances in which the law "permits some partition of the reference group against which an actor's behavior is judged"). The law's approach to this question has been the subject of some analysis and critique. *See generally id.* (arguing that negligence law should be "personalized" to match individual differences in skill and riskiness); Warren F. Schwartz, *Objective and Subjective Standards of Negligence: Defining the Reasonable Person to Induce Optimal Care and Optimal Populations of Injurers and Victims*, 78 GEO. L.J. 241 (1989) (analyzing standards of care in negligence and concluding that, despite some areas of concern, the approaches taken by courts tend to roughly reflect trade-offs in the costs of misincentives and of obtaining information about optimal care levels).

87. As Oliver Wendell Holmes, Jr., famously put it,

> If, for instance, a man is born hasty and awkward, is always having accidents and hurting himself or his neighbors, no doubt his congenital defects will be allowed for in the courts of Heaven, but his slips are no less troublesome to his neighbors than if they sprang from guilty neglect. His neighbors accordingly require him, at his proper peril, to come up to their standard, and the courts which they establish decline to take his personal equation into account.

O.W. HOLMES, JR., THE COMMON LAW 108 (1881).

88. *See, e.g.*, Kenneth S. Abraham, *Strict Liability in Negligence*, 61 DEPAUL L. REV. 271, 272 & n.6 (2012) (collecting sources on this point); Grady, *supra* note 8, at 896-98, 897 n.24, 898 ("In the 20 years that have elapsed since [Guido] Calabresi published *The Costs of Accidents*, his theory about the pocket of strict liability has guided practically all thinking on the subject."); *see also* Goldberg & Zipursky, *supra* note 81, at 746-48, 754-57 (discussing this view and characterizing such unavoidable conduct as "strict liability wrongs").

altogether.[89] Yet unless there is some reason that people should be encouraged to engage in certain activities at particular levels, the fact that people will respond to a regime that holds them liable for harms they cannot help causing by doing less of an activity or dropping out of the activity altogether seems like a point in favor of the approach.[90] To be sure, there are instances in which the desire for broader participation in certain activities essential to human flourishing, or the need for children to go through a learning phase, or the high social value associated with certain forms of inclusiveness, will militate in favor of tailoring. But the fact that we already have a system that holds many actors responsible for shortfalls they cannot help is often regarded as a functional aspect of tort law rather than a cause for concern.[91]

Lapses, too, represent a way in which strains of strict liability work their way into a negligence system.[92] As already suggested, the resulting metering effect may be desirable since the negligence system on its own fails to effectively control activity levels.[93] What seems most problematic, I posit, is not the fact of expected liability for lapses, but rather the high variance in individual results associated with those lapses.[94] Where insurance markets are available

89. *See, e.g.*, Schwartz, *supra* note 86, at 242 ("When we use an objective standard to measure the adequacy of the care taken while engaging in the activity, people who are poorly equipped to take care exclude themselves from the activity because they view the applicable standard of care as excessively onerous.").

90. As Schwartz notes, it is also possible to use a subjective standard of care to screen out those who should not participate in the activity by making the question of engagement in the activity an explicit part of the negligence inquiry. *See id.* While this approach may be sensible under some conditions, it often confronts serious information problems, including an inability to know the value that individual actors glean from engaging in an activity, or even what their actual activity levels were. *See* SHAVELL, *supra* note 49, at 25-26; Schwartz, *supra* note 86, at 279. Some of these information problems may be abating, however. *See generally* Ben-Shahar & Porat, *supra* note 86 (discussing the role of "Big Data" in advancing a personalization agenda).

91. The degree to which unattainable standards of care will induce overinvestment in precautions by potential injurers depends on the extent to which the law imposes a "cliff effect" at the point of due care by holding negligent actors liable for accidents that their negligence did not cause. *See* Ben-Shahar & Porat, *supra* note 86, at 652-53. This issue is taken up in Part II.

92. *See, e.g.*, Grady, *supra* note 12, at 299-300, 307.

93. *See supra* notes 49-51 and accompanying text.

94. Grady notes a number of other concerns, including making defendants into insurers of plaintiffs or forcing plaintiffs to (effectively) buy excessive amounts of insurance against injury. *See* Grady, *supra* note 12, at 308-09, 333.

to translate the high-variance individual results into more predictable expected value "taxes" on activity levels, the results may seem less troublesome. To be sure, we still must worry about the distortive effects that Cooter and Porat identify if different preventative technologies give rise to different levels of underlying liability, since the mere translation of realized results into expected values does not address that disparity.[95] This brings us to the second half of the proposed approach, which would remove this distortion by applying strict liability to automated processes that produce errors at predictable rates.

Here, we can start by asking why an ordinary negligence standard does not already reach errors made by machines.[96] The reason, presumably, is that the machine's errors are evaluated in the aggregate (as an "error rate") and the process is deemed to be non-negligent so long as there is no cost-effective precaution that could have reduced the overall loss rate. A plaintiff might attempt to argue that the specific failure in question could have been cost-effectively avoided by a trivial intervention had a person been present to catch the failure. Yet a person's presence cannot be instantly conjured up only at the moments of failure; catching the machine's error would require continuous human presence throughout the machine's operation. If it would not have been cost effective to hire a person for the entire running of the machine, the error could not have been avoided in a cost-effective manner. And it seems unduly

95. *See supra* Part I.B.

96. A separate inquiry is why the law would not treat *the decision to use* the automated process as itself negligent if it were in fact true that overall expected accident costs would be lower with the human-mediated process (notwithstanding the lapses). Doing so would make the machine-user liable for all harms caused by that choice. The (unsatisfying) doctrinal answer is that the decision to use a particular technique or procedure is unlikely to be second-guessed (especially in medical malpractice) as long as each procedure is well-accepted in the field and meets customary standards of care. *See* Cooter & Porat, *supra* note 7, at 335-36, 353-55; Grady, *supra* note 12, at 317. This means that the fixed or inherent risks associated with a given procedure do not typically give rise to a finding of negligence, even if another procedure would have fewer of these fixed or inherent risks, and more "variable" or lapse-based risks. Leveling the playing field in this context would require applying strict liability to accidents that occur under the inherently riskier procedure (or, put differently, treating it as negligent simply to have chosen the riskier procedure, given the existence of another procedure that entails less overall risk). Viewed in this light, we can understand the "automated processes" situation as a special case of a more general problem: that the bundled choice between procedures is typically given less scrutiny than errors that occur within procedures.

anthropomorphic to suggest that the machine *itself* behaved negligently in those moments of failure.

In the case of human lapses, the same reasoning might be attempted. For example, a driver might argue that she would effectively have to hire a "second self" (an assistant) to watch over her shoulder to prevent the inevitable occasional lapse, and that this would not be cost effective. But because one's own attention to the task at hand is viewed as a resource that can be expanded and contracted at will, it is assumed that one could always have paid just a bit more attention at the time of failure—whether or not this is realistic. Thus, in the machine case, two different lumpy precautions become conceptually severable: one for a machine with a certain error rate, and the second for a full-time human overseer. Having invested in the first, investing in the second is not cost justified.[97] In the human case, the operator with a certain error rate and her on-demand double-checker are believed to be embodied in the same individual. The two lumpy precautions in the machine case are replaced with a continuous spectrum of caretaking that is thought to be open to the human individual as she goes about a judgment-intensive task like driving.

Now that we see the nature of the problem, the rationale becomes clear for moving to strict liability in the machine case while keeping the negligence standard in place in the human-error case. The assertion that doing so will chill innovation and thwart safety-enhancing moves to automation[98] ignores the implicit role of evaluative aggregation in determining liability outcomes. It assumes that the human-to-machine comparison is apples to apples, when it is really apples to orchards.[99] Given the tendency toward evaluative

97. This might not always be the case; there might be some machines that perform so poorly that adding an attendant would be cost justified. *See infra* Part II (examining issues surrounding additive and substitute precautions).

98. *See, e.g.*, Ryan Abbott, *The Reasonable Computer: Disrupting the Paradigm of Tort Liability*, 86 GEO. WASH. L. REV. (forthcoming 2018) (manuscript at 4-5), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=%202877380 [https://perma.cc/X45S-4LZW] (contending that a negligence standard for human-caused injuries coupled with a strict liability standard for computer-caused injuries "discourages automation").

99. Cooter and Porat are, in effect, urging an orchards-to-orchards comparison—one that would either aggregate behavior into larger patterns or look at the use of durable "second-order precautions" designed to produce particular patterns over time. *See* COOTER & PORAT, *supra* note 20, at 70-72; Cooter & Porat, *supra* note 7, at 339-48 (discussing how "second-order

disaggregation in the human case and the tendency toward evaluative aggregation in the mechanized case, a (nominal) negligence standard in the former paired with strict liability in the latter merely levels the playing field. And it does so in a way that conserves on information costs because it eliminates the need to investigate the human's overall "error rate."

This approach dovetails with other policy considerations that support strict liability for automated processes.[100] The use of a routinized or mechanized process suggests significant repetition of actions and the capacity to determine in advance overall levels of exposure. Repeat play and risk-bearing capacity are often invoked to support strict liability, framed as enterprise liability.[101] And, in fact, the law tends to treat repeat-play enterprises carried out on a broad scale differently from one-off activities carried out by individuals. For example, modern products liability law has evolved a strict liability standard for manufacturing defects. Even though a manufacturer may choose a process that has a very low error rate, and no further cost-justified precautions are available, liability still follows in cases of failure.[102] The very predictability associated with the low failure rate enables effective cost reduction and appropriate choices about activity levels. The broader point is that the lumpiness in outcomes associated with accidents is mediated by repetition, transforming what might in the individual case be an outlier observation into part of a larger pattern.

---

precautions" can reduce the probability of a lapse and urging a "second-order reasonableness defense" to lapses).

100. *See, e.g.*, Omri Ben-Shahar, *Should Carmakers Be Liable When a Self-Driving Car Crashes?*, FORBES (Sept. 22, 2016, 11:36 AM), http://www.forbes.com/sites/omribenshahar/2016/09/22/should-carmakers-be-liable-when-a-self-driving-car-crashes/ [https://perma.cc/N7L2-9QCU] (offering an insurance rationale for automaker liability and suggesting that it would not produce distortive effects).

101. *See, e.g.*, Gregory C. Keating, *Products Liability as Enterprise Liability*, 10 J. TORT L. 41, 74 (2017) ("When activities are actuarilly large, the accidents that they engender will likewise be predictable and regular, and the costs of those accidents can be factored into the costs of conducting the enterprise."). For historical background and a critique of enterprise liability, see generally George L. Priest, *The Invention of Enterprise Liability: A Critical History of the Intellectual Foundations of Modern Tort Law*, 14 J. LEGAL STUD. 461 (1985).

102. *See, e.g.*, COOTER & PORAT, *supra* note 20, at 66; Keating, *supra* note 101, at 79-80.

II. STACKING CARE

Aggregation also enters into judgments about untaken precautions. Recall the example of the unfenced cricket field and the brained pedestrian.[103] An eight-foot fence—one high enough to stop the ball—appeared cost justified when considered as a unit. But the last two feet of the fence were not cost justified at the margin, and the optimal six-foot fence would not have prevented the accident. Should the defendant be able to avoid liability by demonstrating this fact, or should the plaintiff be free to choose how to aggregate together or break apart the precautionary steps that the defendant failed to take?

This example presents the precaution-aggregation puzzle with unusual clarity. The neatly stacked fence heights with known costs and projectile-stopping properties enable us to partition the defendant's untaken precaution—the failure to build any fence at all—into two parts: a negligent omission (failing to construct the first six feet), and an efficient omission (failing to build beyond six feet). Moreover, because the hypothetical states the height of the exiting ball with certainty, it is readily apparent whether the defendant's negligent omission or merely her efficient omission was responsible for the accident. This atypically unambiguous setup makes the liability question appear deceptively simple.

A. Marginal Analysis and Causation

One of the most basic lessons of economics is that analysis must be conducted "at the margin."[104] Whether you are deciding how tall

---

103. *See supra* text accompanying notes 15-18.

104. This principle, familiar from economics, turns up in some decided tort cases. *See* WILLIAM M. LANDES & RICHARD A. POSNER, THE ECONOMIC STRUCTURE OF TORT LAW 99-100 (1987) (observing that the court's analysis in *Blyth v. Birmingham Waterworks Co.* (1856) 156 Eng. Rep. 1047, 1049; 11 Exch. 781, 784-85, evidenced "an implicit concern with marginal or incremental rather than total costs of care"). As Landes and Posner explain,

> The [*Blyth*] court was not interested in whether the total costs of burying the main to a depth at which it would not have burst even in the unusually severe frost of 1855 were less than the expected accident costs of the pipes' bursting. It was interested in whether, given that the mains had been buried to a depth that would prevent their freezing in any ordinary frosts, the incremental expense of

to build a fence, how many widgets to make, or how many cookies to eat, a surefire way to get the wrong answer is to continue building, manufacturing, or eating until your *total* costs begin to outweigh your *total* benefits. That stopping point will be too late in any situation where marginal costs are rising, marginal benefits are falling, or both. Under these conditions, which are highly typical, the first units in the sequence will have a more favorable ratio of benefits to costs than the later units in the sequence. In making decisions about each subsequent unit, what matters is whether the benefits *of that unit* outweigh the costs *of that unit*.[105] As soon as the answer stops being yes, it is time to immediately stop adding units. Continuing beyond that point will mean adding units that are more costly than beneficial, even though it may take some time for the deficits that they introduce to eat away the surplus of benefits over costs associated with the earlier units and bring the total cost-benefit ratio into the red.

Applying marginal analysis to our cricket case identifies the point at which an actor who was taking due care would have ceased adding height to the fence.[106] In our example, that stopping point would be six feet. Thus, it might seem like a red herring to assert that an eight-foot fence, considered as a unit, would produce total benefits that exceed total costs. But the fact that six feet rather than eight feet is the correct stopping point for an actor who is exercising due care does not, on its own, tell us what the legal response should be to a defendant who never started building a fence at all.

It might seem axiomatic that in a negligence regime, a defendant cannot be held liable unless her *negligence* caused an accident.[107] That premise suggests the following two-step operation: (1) identify the optimal fence height using marginal analysis; and (2) see if the defendant's failure to build a fence of that optimal height caused the

---

> protecting against an unusually severe frost would be justified by the incremental reduction in accident costs resulting from such an expense.

Id. at 100.

105. For a discussion of marginal analysis as it applies in legal contexts, see WARD FARNSWORTH, THE LEGAL ANALYST: A TOOLKIT FOR THINKING ABOUT THE LAW 33-35 (2007).

106. *See id.* (discussing the application of marginal analysis to a cricket fence example inspired by *Bolton v. Stone*).

107. *See, e.g.*, COOTER & PORAT, *supra* note 20, at 23 ("Under prevailing negligence law and causation principles, liability should be imposed only for harms *caused* by the injurer's *negligence*."); Kahan, *supra* note 15, at 428.

accident. If not, then there should be no liability. But this analysis embeds a questionable assumption about the respite from liability that a negligence regime provides. On an efficiency analysis, the shape of that respite should depend on the incentives that it produces for actors, which depend in turn on how different doctrinal formulations interact with tort law's specific architecture.

If we view a negligence regime not as setting the outer boundaries for appropriately imposed liability, but rather as offering a safe harbor[108] when a defendant has (actually) done all that might efficiently be done to prevent an accident, the problem looks different. In other words, it is not self-evident that a negligent defendant is entitled to have her omissions partitioned into negligent and efficient increments when she has failed to undertake the partitioning work herself by exercising due care. This reframing allows us to turn to the question of how best to minimize the costs of accidents, including the costs of preventing accidents and of administering the system.[109] The next Section takes up that inquiry. Drawing on Mark Grady's analysis,[110] I show how the aggregation choices used in assessing precautions represent a powerful but underappreciated way to fine-tune tort law incentives—with surprising results for our cricket example.

B. Unifying Untaken Precautions

In many areas of law, a defendant becomes liable only after her behavior aggregates to a certain point or crosses a certain line—yet once she does cross the line, her behavior is treated (without comment) as an indivisible unit. Next to nothing separates the driver who is just over the line that demarcates driving while intoxicated and the driver who is just below that line. Still, a driver who is just over the line will suffer consequences that are harsher than could be explained by comparing her behavior's expected social costs with those of her trivially less intoxicated counterpart. The line, we understand, has to be drawn somewhere—but this does not mean

108. *See supra* note 84 and accompanying text.

109. *See* GUIDO CALABRESI, THE COSTS OF ACCIDENTS: A LEGAL AND ECONOMIC ANALYSIS 26-31 (1970).

110. *See generally* Grady, *supra* note 20.

the line-crosser is entitled to have everything below the line cleared from her account, or to partition her intoxication into the segments falling above and below the legal limit.[111] Thus, she cannot limit her exposure by demonstrating that she would have driven almost as badly had she not taken the last sip that put her over the legal limit. Having actually crossed the line, she forfeits the protective benefits the line otherwise would have afforded to her below-threshold conduct.[112]

In the drunk driving context, as in other criminal contexts, we may tolerate or even welcome a cliff effect in order to get meaningful deterrence. To use Robert Cooter's distinction, "sanctions" seem more appropriate than mere "prices."[113] Tort law is different in that there are significant social costs to overshooting as well as undershooting due care. But even when a price (and not a sanction) is what we seek, that price still must be set appropriately. For reasons to be explained below, permitting defendants to partition an untaken precaution identified by plaintiffs into negligent and nonnegligent omissions would apply a distortive discount to the tort system's pricing mechanism. To see why, we must first step back to examine how a negligence regime operates.

1. Negligent Defendants and Efficient Accidents

Unpacking the cricket fence problem first requires pinpointing the essential difference between negligence and strict liability. The two approaches differ in their treatment of accidents that happen even though everyone is exercising due care. To prevent all accidents would require a level of precaution that would cost society far more than it would be worth.[114] This is why we do not set all speed

111. Indeed, if asked to articulate what the drunk driver did wrong, we would probably say something like, "downing three beers in quick succession and then getting behind the wheel," not "driving after taking the tiny sip that put her blood alcohol level over the legal limit."

112. This result is not inevitable. The law could be made more continuous in its effects by punishing people only for the marginal impacts of their over-the-limit conduct. *See* Kolber, *supra* note 29, at 293 ("If one is especially concerned about the horizontal inequity of giving the slightly less culpable no punishment at all, we could eliminate the inequity simply by punishing all offenders only to the extent that their conduct exceeds the threshold."). But the fact that this is not how law usually operates remains suggestive. *See also infra* Part IV.C.

113. *See generally* Cooter, *supra* note 20.

114. *See generally* CALABRESI, *supra* note 109.

limits at ten miles per hour, require cars to be built like tanks, or mandate the wearing of body armor at all times.[115] The costs of "efficient accidents"[116]—those that occur even when everyone is being as careful as it makes sense to be—must fall on someone. Negligence law leaves these costs to fall on the victims, while strict liability assigns the losses from these efficient accidents to the injurer.[117] When an injurer in a negligence regime falls short of due care, however, she exposes herself to liability. But how much liability?

To return to our example above, should the cricket club that has negligently failed to build any fence at all be liable for the harm caused by all balls that leave the field (regardless of altitude), only those errant balls that would have been stopped by an optimal six-foot fence, or some group of escaping orbs in between these two alternatives? The academic literature on this topic notes correctly that the first alternative, which we might call the "harming while negligent" (HWN) approach, creates a cliff of liability[118]—fall one iota short of due care and all bets are off! Under this approach, the actor would face liability for the efficient accidents that otherwise would have been charged to the victim, as well as the ones that the actor's negligence itself actually caused. Such a discontinuity puts a great deal of pressure on actors to make sure they meet the due-care

115. *See* GUIDO CALABRESI, IDEALS, BELIEFS, ATTITUDES, AND THE LAW: PRIVATE LAW PERSPECTIVES ON A PUBLIC LAW PROBLEM 9 (1985) (using the example of cars "built like tanks" that "cannot go faster than ten miles per hour" to emphasize that the costs of safety can sometimes be unacceptably large).

116. *See, e.g.*, Isaac Ehrlich & Richard A. Posner, *An Economic Analysis of Legal Rulemaking*, 3 J. LEGAL STUD. 257, 269 (1974) (referring to "nonnegligent" accidents as "efficient" ones). Accidents that could not have been prevented through the exercise of due care are often referred to in the legal literature as "unavoidable," but this is a misnomer. *See* Grady, *supra* note 8, at 910 n.73 (describing this usage and observing that "most unavoidable accidents can indeed be avoided, but at excessive cost"). "Efficient accidents," while closer to the mark, is potentially misleading as well. Some of the means through which these accidents might be reduced, such as curtailing activity levels, do not get picked up in standard due-care analyses but might be efficient for actors to undertake; if so, the accidents thus avoided would not have been "efficient" ones, even though "due care" (as standardly understood) could not have prevented them.

117. *See, e.g.*, Richard A. Posner, *Strict Liability: A Comment*, 2 J. LEGAL STUD. 205, 208-09 (1973).

118. *See, e.g.*, COOTER & PORAT, *supra* note 20, at 21; *see also supra* note 20 and accompanying text.

standard; it imposes a sanction, and not a mere price, on falling short.[119]

Yet it seems doctrinally wrong. Surely the cricket club should not be liable in a negligence regime for a freak cricket ball that flies out of the field at an altitude of sixty feet since no reasonable fence could have stopped it from happening. And, in fact, it appears that courts do not impose liability on negligent actors when it is clear that behaving with due care would not have kept the accident from happening.[120] This does not mean defendants always benefit from a surgical separation of negligent omissions and efficient omissions. On the contrary, two kinds of uncertainty allow those categories to be routinely blurred together. One is uncertainty about which increment of the defendant's omission caused the harm. Often, the negligent portion of the defendant's conduct caused the harm with some positive probability but not with certainty, as will be discussed below.[121] For now, it is sufficient to observe that this form of factual uncertainty likely causes negligent defendants to be liable for more harm than would be statistically associated with their negligent omissions alone.[122]

A second source of uncertainty concerns the position of the due-care standard, which marks the dividing line between a defendant's negligent omissions and her efficient omissions. In many cases, it will be evident that a defendant's conduct fell below any plausible standard of due care, yet remain unclear exactly how much more was necessary to achieve due care. Here, we might expect a systematic skew in the direction of liability, especially if due care is evaluated in a manner endogenous to the choices the defendant

119. *See* Cooter, *supra* note 20, at 1538-39.

120. *See* Wright, *supra* note 20, at 84 ("[T]he decisions of the courts ... have consistently limited negligence liability to injuries that were caused by the negligent aspect of the injurer's activity."); *see also* COOTER & PORAT, *supra* note 20, at 23-25 (distinguishing cases of perfect and imperfect causal attribution).

121. *See infra* Part III.

122. It is possible that factual uncertainty about causation is resolved symmetrically. *See* Kahan, *supra* note 15, at 442 (considering the possibility that finders of fact might "semi-automatically" find causation once negligence was established, but finding it "also plausible to assume that some fact finders resolve these difficult questions in the favor of injurers"). The question is an empirical one, but it seems likely that a fact-finder will err on the side of liability when faced with a plainly harmed plaintiff, a plainly negligent defendant, and factual uncertainty running only to *which facet* of the defendant's conduct produced the harm. *See infra* Part III.

actually made about precautions. In the absence of a clear external standard for due care (such as a speed limit), attention tends to focus on "the untaken precaution"—the content of which depends on what the defendant actually did.[123] The cricket hypothetical enables us to observe this potential for endogeneity at close range and evaluate whether it represents a feature or a bug.

The balance of this Section considers how to approach the liability of negligent defendants where marginal analysis establishes that taking due care would not have prevented the harm in question. The HWN standard discussed above will be compared with a "harming because negligent" (HBN) approach, as well as with a hybrid alternative falling between these extremes—what Mark Grady calls a "cost-benefit rule,"[124] and which I will refer to here as the "unified untaken precaution" (UUP) rule.

2. Cliffs and Ledges

Consider a rule that would limit the defendant's liability to those accidents that were actually *caused* by her negligence—in other words, an HBN regime. Such an approach eliminates the cliff of liability that can confront defendants who fall just a little short of due care under an HWN approach. To be sure, the cliff provides a powerful incentive to reach the requisite level of due care, and perfectly informed and rational actors would simply meet that standard without fail. But actors who are uncertain about the location of the due-care standard (or who fear mistakes by fact-finders) may take excessive care in an effort to avoid disproportionately harsh results stemming from relatively minor shortfalls (or perceived shortfalls).[125]

Yet the HBN alternative leaves in place a *ledge* that also proves problematic: liability sits at zero for all harms that due care would

123. *See* Mark F. Grady, *Untaken Precautions*, 18 J. LEGAL STUD. 139, 143 (1989) ("[Courts] take the plaintiff's allegations of the untaken precautions of the defendant and ask, in light of the precautions that had been taken, whether some particular precaution promised benefits (in accident reduction) greater than its associated costs."). A failure of due care may also be inferred from the fact of the accident in some cases, following the doctrine of res ipsa loquitur. *See id.* at 141 & n.7 (collecting literature).

124. *See* Grady, *supra* note 20, at 814-21 (describing and depicting this approach).

125. *See, e.g.*, *id.* at 806-12.

not prevent. Not only is an actor who failed to exercise due care excused from the costs associated with the accidents that due care would not prevent, an actor who goes beyond due care gets no further credit for reducing these accident costs.[126] There is, at the point of due care, a sharp drop-off (to zero) in the private returns that the actor internalizes from taking more increments of care, even though the societal benefits associated with those added increments of care may remain quite significant.[127] *That* discontinuity—what we might call a ledge effect—matters too.[128] Uncertainties that present difficulties for the HWN regime reemerge here to generate drawbacks for the HBN approach, as Grady has carefully demonstrated.[129]

Consider an actor like the cricket club that must decide whether to err on the side of going above or falling below due care, where there is uncertainty about exactly how a court will evaluate its conduct.[130] The club's owner has to shell out money for each

126. The fact that due care has been reached does not mean that no further reductions in accident costs are possible; it simply means that those reductions cost more than they are worth. Although the lack of net social value associated with these further accident reductions means that they should not be undertaken in a world of perfect information, they nonetheless matter when an actor must decide whether to err on the side of more or fewer precautions. *See id.*

127. This can be readily demonstrated by considering a case where risks and precautions are perfectly continuous. As the actor approaches the level of due care, she spends $1.00 to gain $1.01 in accident reduction benefits and should stop when the marginal costs equal the marginal benefits (spending $1.00 to save $1.00). If she goes $1.00 further, accident costs fall by $0.99, but she internalizes none of those social gains. Her liability is zero whether she spends the extra dollar or not. What is from society's standpoint an overexpenditure that costs just one penny (net) is from the actor's perspective an error that costs a full dollar because she gets nothing in return.

128. Although Grady's analysis does not put matters in quite these terms, his graphical representation of the operation of his P*-cutoff rule illustrates this discontinuity. *See* Grady, *supra* note 20, at 812-13, 813 fig.3 (showing that the costs internalized to the actor rise sharply as she moves past the point of due care, P*, creating an asymmetry with her cost profile before reaching due care).

129. *See id.* at 801-06, 802 fig.1, 812-13, 813 fig.3.

130. *See id.* at 806-21 (analyzing and depicting the choice between underprecaution and overprecaution, given uncertainty, under different legal rules). Significantly, this uncertainty is the primary reason why the HBN versus HWN choice matters. In the absence of such uncertainty, the actor would simply comply with due care, and that would be that. *See id.* at 806-09. Of course, all actors experience some inevitable lapses, and some actors cannot meet the due-care standard easily or at all. For these reasons, a negligence standard already inserts an element of strict liability—one that the HWN regime would exacerbate. *See supra* notes 88-91 and accompanying text.

increment of fencing height. If she builds too high, she will be wasting her money; her liability is already at zero once due care is achieved, and she receives no credit for going further. If she builds too low, however, she will save on fencing costs and will expose herself only to the incremental accident costs associated with the amount by which she fell short. As a result, mistakes in the direction of too much care will cost her more than mistakes in the direction of too little care.[131] Thus, actors in an HBN regime may be inclined to err on the side of stopping short of due care.

A numeric example will clarify. Suppose the schedule of marginal costs and benefits for increasing fence heights between five and eight feet is as shown in Table 2.

Table 2. Fencing Costs and Benefits

	Marginal Cost of Construction	**Marginal Benefit to Actor (HBN Regime)**	**Marginal Benefit to Society**
Increase from 5 to 6 feet	100	175	175
Increase from 6 to 7 feet	100	0	75
Increase from 7 to 8 feet	100	0	25

Here, due care under the Hand formula[132] requires building to, but not above, six feet (assuming that only integer fences are possible). Beyond that point, construction costs exceed savings in accident costs.

Now consider the problem from the viewpoint of a cricket club in an HBN regime that is uncertain ex ante what fence height will be judged optimal. An upward deviation from the optimal six-foot fence

131. For a graphical demonstration of this point, see Grady, *supra* note 20, at 812-13, 813 fig.3; *see also id.* at 801-06, 802 fig.1.

132. *See supra* note 16.

to a seven-foot fence will cost the cricket club an extra $100 in construction costs but will not save it anything in accident costs since liability is already sitting at zero. So it loses $100 on net by guessing too high about the standard of care. A downward deviation from the optimal six-foot fence to a five-foot fence will mean bearing an extra $175 in expected accident costs, but it will also mean saving $100 in construction costs. The net cost of guessing too low is just $75. The actor will, therefore, err on the side of too little precaution.

From society's standpoint, the calculation looks different. The cost of the upward mistake is just $25. Even though an extra $100 is spent on construction, there is a $75 savings in accident costs to help offset it. The downward mistake, on the other hand, costs society the full $75 that it costs the actor. On these numbers, society would prefer that the actor make the upward mistake rather than the downward mistake.[133] Of course, a downward mistake will not invariably be more costly than an upward mistake—that will depend on the specific profile of costs and benefits associated with the available changes in fence heights.[134] The important point is that the ledge effect baked into the negligence standard will cause injurers' errors to skew low under an HBN rule, rather than center on the optimal point.[135]

Compare next the calculation the club owner would make under an HWN standard that makes a negligent actor liable for all injuring balls, regardless of altitude. Now, she will choose to overshoot rather than undershoot due care.[136] If she builds to only five feet, she will be liable not just for the $175 in expected costs for the

133. Of course, society would prefer that the actor stop at exactly the level of due care, since going further costs more in precaution than it saves in accident costs. But if a mistake is to be made, it would prefer the less costly of the two.

134. The two *types* of errors (undershooting and overshooting) are equally socially costly, but the *magnitudes* of these mistakes may differ in a given setting, owing to factors like lumpiness in precautionary technologies. Here, the fact that fences come only in integer heights—one cannot add just a few inches—makes fencing precautions lumpy. In some cases, lumpiness in precautionary alternatives will actually help to *remove* distortions in decision-making by producing greater convergence between the privately optimal strategy and the socially optimal strategy, while in other cases it does the opposite. *See* Lee Anne Fennell, *Slicing Spontaneity*, 100 IOWA L. REV. 2365, 2377-82 (2015). None of this changes the fact that an HBN regime has a tendency to make precautions skew low, but lumpiness in precautionary steps can ameliorate or exacerbate the effects of that tendency in certain cases.

135. *See* Grady, *supra* note 20, at 812-13, 813 fig.3.

136. *See id.* at 809-12, 810 fig.2.

increment in between five feet and six feet, but also the accidents that occur when balls exit the field at seven, eight, and sixty feet (and all other heights). This means she will also be on the hook for the additional $75 in expected costs that would have been saved by the six-foot to seven-foot increment, the additional $25 in expected costs associated with balls exiting in the seven-foot to eight-foot range, and all the costs for balls higher than that. To save $100 in construction costs, she exposes herself to at least $275 in liability, or a net cost of $175 or more. Making a mistake in the other direction and building to seven feet is less costly in this regime. True, she spends $100 more on construction and does not get any benefit from society's $75 in cost savings. But a $100 mistake is cheaper than a $175 mistake.

Grady's analysis makes the case for an intermediate alternative, one that he suggests hews most closely to what courts actually do.[137] This approach would impose liability on a negligent defendant only when the plaintiff can identify an untaken precaution that satisfies two criteria: (1) it would have prevented the accident; and (2) taken as a whole, the precaution's benefits in terms of accident reduction exceed its costs.[138] In short, the precaution (considered as a unit) must be both causally effective and cost effective. Under such a regime, liability would follow in the hypothetical where the cricket club constructed no fence at all, the ball sailed out at just under eight feet, and the eight-foot fence, taken as a whole, would have been cost justified. Because this approach relies on artificially "unifying" the untaken precaution for purposes of evaluating negligence liability, I call it the "unified untaken precaution" (UUP) rule.

The workings of this rule can be explained intuitively. To be incentivized to dead-center her behavior on due care, an actor must bear equivalent costs for overshooting due care and undershooting

137. *Id.* at 814-29 (describing his "cost-benefit" approach).

138. *See id.* at 815 (explaining that under his approach, courts would "find an injurer negligent whenever the costs of at least one specific untaken precaution are less than the reduction in expected harm that would have resulted from that precaution" and "would then impose liability if the specific act of negligence used to prove the breach of duty was the cause in fact of the harm the victims suffered"); Grady, *supra* note 34, at 673 (describing the cost-benefit analysis and the causal inquiry as "two different rings of fire" that a proposed untaken precaution must clear).

due care. But in an HBN regime, as we have seen, an actor bears *more* costs than society does when she overshoots due care, but the *same* costs as society does when she undershoots due care. The actor's liability is zero across the whole range of conduct above the due-care line, so every penny she spends on excess precaution buys her no benefits at all, even though it does still buy society additional benefits (albeit ones that are not worth their full cost). This divergence between the cost of the overshooting mistake for society and its cost for the actor herself will create a skew in the direction of insufficient care under an HBN regime.

To correct that skew requires increasing the costs that attach to undershooting due care *just enough* that the actor is left indifferent between overshooting and undershooting.[139] The HWN approach cannot accomplish this sensitive task. Instead, it bluntly overcorrects with a cliff of unlimited liability that creates a skew in the opposite direction, toward too much care. Under an HWN rule, anyone who falls even a little short of due care will be liable for *any* accidents she causes. The UUP approach, by contrast, makes the degree of extra exposure added to undershooting errors dependent on the degree to which the actor undershot due care.[140] This is accomplished by allowing a victim to specify—and, crucially, treat as an indivisible unit—any untaken precaution that meets the criteria of cost effectiveness and causal effectiveness.[141]

When an injurer has been negligent, such a qualifying untaken precaution can extend beyond the point of due care.[142] By treating the proffered untaken precaution as a single unit, a court is effectively comparing the precaution's total costs with its total benefits.[143] As we saw above, total benefits can continue to exceed total costs well beyond the optimal stopping point on a marginal analysis,

---

139. Alternatively, were it possible, society could rebate to the actor the accident cost savings associated with overshooting due care.

140. *See* Grady, *supra* note 20, at 815-16.

141. *See id.* at 815-17.

142. *See* Grady, *supra* note 34, at 661 ("Untaken precautions beyond the efficient set appear cost-beneficial only when the injurer has used less precaution than due care.").

143. To be clear, the court is examining the marginal cost *of the precaution as a whole* with the marginal benefit *of the precaution as a whole*. But the marginal steps contained within the envelope of "the precaution" are ignored, as are any marginal differences between a lesser precaution and the proffered one. *See id.* at 670-71 (distinguishing "marginal analysis" from a cost-benefit assessment of "discrete untaken precautions").

because the surpluses from the earlier (efficient) units subsidize the deficits in the later (inefficient) units.[144] Still, the requirement of cost effectiveness provides an important constraint: the actor can be held liable for untaken precautions that exceed due care, but only to the extent of the surplus of benefits over costs associated with the untaken increments that actually fall short of due care.[145] The larger the defendant's error, the more socially costly her shortfall, the greater the potential for liability for untaken precautions extending beyond due care.[146]

Under certain assumptions, including perfect linearity, the UUP approach adds consequences to the insufficient care side of the line that exactly make up for the gap between social costs and private costs on the excessive care side of the line.[147] But even under more realistic conditions where precautions are lumpy and their effects nonlinear, the UUP approach is likely to outperform the other alternatives. Unlike in a pure HBN regime, a negligent defendant is not allowed to partition her omissions into negligent and nonnegligent components when a qualifying untaken precaution can be identified that spans the two categories. But, unlike HWN, the UUP approach does not generate an unlimited cliff of liability. The pedestrian struck by a ball exiting the field at a height of sixty feet has no claim because a sixty-foot fence would not be worth its cost, even when considered as a unit and compared with nothing at all.

Instead, as Grady's elegant graphical representations of this point demonstrate, a plaintiff's ability to identify an untaken precaution that exhibits both cost effectiveness and causal effectiveness depends crucially on what the defendant has done and failed to do.[148]

---

144. *See supra* Part II.A; *see also* Grady, *supra* note 34, at 671-72 ("When courts conduct cost-benefit analysis of alternative untaken precautions, the surplus on the movement up to the efficient level allows the precaution to extend into a zone where the last part of it is producing a deficit.").

145. *See* Grady, *supra* note 34, at 671-72.

146. *See id.* at 671 ("The lower the defendant's actual level of precaution, the more social surplus that exists from moving up to the efficient level.").

147. The result is a symmetrical penalty on errors on either side of due care. *See* Grady, *supra* note 20, at 817-21, 819 fig.4, 820 fig.5. The counterbalancing may not be perfect under real-world conditions where nonlinearities and lumpiness exist.

148. *See id.* at 815-16 ("[A]s the injurer takes less precaution, he creates more opportunities for the victim to show a breach of duty."); *id.* at 818-21, 819 fig.4, 820 fig.5; *see also id.* at 801-06, 802 fig.1.

Suppose that the cricket club had constructed the optimal six-foot fence. In that case, there would have been no way that a plaintiff could identify an untaken precaution that would have been cost justified and also would have stopped the ball in question. If incremental increases beyond six feet are not cost justified, then it is immaterial that they would have stopped this accident. But when the defendant fails to meet the standard of due care, the increment by which she falls short effectively grants the plaintiff some added running room to identify a precaution that meets the doctrinal criteria. So by failing to construct any fence at all, the cricket club (properly, in Grady's view) leaves itself open to the lumpy precaution that fits into the gap created by its own shortfall.[149]

As a final demonstration of how this approach works, consider an actor who constructs a five-foot fence rather than the optimal six-foot fence. Suppose two pedestrians are injured by cricket balls in the same week; one of these balls exited the field at a height of just under eight feet, as in the original example, and the second one exited the field at a height of just over six feet. Under Grady's approach, the second plaintiff, but not the first plaintiff, could make out a successful case. The first plaintiff would be unable to show an untaken precaution that would have been cost justified and also would have prevented this accident. Only an eight-foot fence would have stopped this accident, but increasing the fence height from five feet to eight feet is not cost justified on the figures given in Table 2. It would cost $300 in construction costs but would save society only $275 in accident costs.[150] By building to five feet, the actor has shielded herself from liability she would have faced had she built no fence at all.

The second plaintiff, struck by a ball just over six feet in altitude, would have more luck. Here, the cost of building the extra two feet ($200) compares favorably with the accident cost savings associated with those extra two feet ($250);[151] considered as a unit, the extra

149. *See id.* at 815-16; Grady, *supra* note 34, at 671 ("Whenever a plaintiff is allowed to add a lump of untaken precaution to the defendant's suboptimal care level, he can show that marginal benefits exceed costs even when the lump extends somewhat beyond due care.").

150. This figure is obtained by adding the marginal benefit for each of the one-foot increases ($175 + $75 + $25 = $275).

151. This figure is obtained by adding the marginal benefit for the two one-foot increases ($175 + $75 = $250).

two feet of fencing is cost justified. It may seem harsh to hit the defendant with liability for an accident that the optimal fence would not have stopped, but the defendant opened herself up to this result by failing to build the optimal fence.

This approach is attractive, but it presents some conceptual and doctrinal difficulties. As the examples above show, liability might result, variously, from failing to build a six-foot, seven-foot, or even eight-foot fence, depending on what kind of fence, if any, the defendant actually built. Does this pattern mean that the due-care standard floats about depending on what the defendant has actually done or left undone? This is indeed what Grady proposes.[152] His framing of his cost-benefit test hews to the HBN philosophy but allows what counts as "negligence" to become more or less demanding depending on the defendant's actual inputs.[153] Disconcertingly, however, this requires us to accept an understanding of due care that appears to eschew marginal analysis[154] and demand (at times) unjustified incremental investments in safety.

There is an alternative way to characterize this approach. We could see it as leaving due care in a fixed position (here, the optimal six-foot fence) but adopting a looser and more instrumental understanding of the relationship between failure to take due care and the imposition of liability. When the UUP approach makes the owner of an unfenced cricket field liable for the ball that an eight-foot fence would have stopped, this is not due to any belief that due care requires an eight-foot fence, or that fences are only available in eight-foot increments. Instead, unifying the precaution and comparing it to what the defendant actually did is merely a neat trick designed to make up for other incentive misalignments produced by the negligence regime.[155] On this reading, the UUP is a sort of "harming while negligent plus" (HWNP) rule, where the "plus" factor that produces liability is the plaintiff's ability to identify an

152. *See* Grady, *supra* note 20, at 814 ("This [cost-benefit approach] retains the concept of legal causation of the P*-cutoff rule, but defines breach of duty in a manner altogether different from the P*-comparison approach.").

153. *See id.* at 816 ("[T]he cost-benefit approach permits alternative proofs of breach of duty, and the levels of untaken precaution available for the victim's proof increase as the injurer's actual precaution decreases.").

154. *See supra* note 143 and accompanying text.

155. *See generally* Grady, *supra* note 20.

untaken precaution that is both cost effective and causally effective.[156] As we will see, a conceptually similar HWNP approach can be pursued in contexts in which the failure to exercise due care produces factual uncertainty about whether due care would have prevented the harm in question.[157]

The UUP standard not only corrects for the distortive ledge effect built into the negligence standard, but also economizes on information. To determine liability, it is not necessary (nor especially useful, nor generally even possible) to calculate the exact location of the due-care line in the abstract.[158] Instead, a court can simply look at whether an untaken precaution exists that would have stopped this accident and that would have also been cost effective on the whole.[159] Consistent with this point, we might understand due care as a step good that, like a bridge, generates certain societal benefits when (and only when) it is provided by defendants in full.[160] These

156. The notion of "causally effective" that I have in mind here would include not just "but for" causation but also standard notions of proximate cause that require that the accident stem from a type of risk that is increased by the activity in question. The examples in the text easily satisfy this standard, so it is not at issue. However, causation would fail on these independent grounds if, for example, a cost-effective reduction in speed would have prevented the accident not because of its effects on accident risks, but simply because it would have caused the vehicle to be somewhere else when an exogenous risk (such as a tree randomly falling) produced harm by coincidence. *See* Berry v. Sugar Notch Borough, 43 A. 240, 240 (Pa. 1899); *see also* Ariel Porat, *Offsetting Risks*, 106 MICH. L. REV. 243, 246-50 (2007); *cf.* Keating, *supra* note 101, at 56 (explaining that enterprise liability is limited to "the 'characteristic risks' of activities").

157. *See infra* Part III.

158. *See* Grady, *supra* note 34, at 660 (discussing the "excessive information" that courts would require to implement an approach based on optimal precaution levels). A related problem is that there may exist no precaution that perfectly aligns with the optimal level of precautionary expenditures, given that precautions are often chunky in nature and must be supplied in particular quantities, if at all. Calculating the point where smooth cost-benefit lines would hypothetically cross provides a technical answer that may have no real-world counterpart—or one that might be too costly to uncover. *See* Gregory C. Keating, *Reasonableness and Rationality in Negligence Theory*, 48 STAN. L. REV. 311, 328-32 (1996) (rejecting a "razor's edge" interpretation of the Hand formula that would determine negligence based on a penny's difference in marginal costs and benefits).

159. *See* Grady, *supra* note 34, at 661 ("This same untaken-precaution approach also reduces courts' need for technical information because they no longer have to identify the precautions that produce the global minimum of social cost; they need only examine the costs and benefits of the precautions that the plaintiff has actually alleged that the defendant failed to take.").

160. *See also infra* Part III.B.

benefits, which include reduced information costs,[161] entitle non-negligent defendants to a safe harbor from liability for efficient accidents.[162] Because defendants who fall short need not be extended this same immunity, tort law is free to shape their liability in the way that will best achieve its objectives. The UUP approach appears to fit the bill, exposing negligent defendants to a limited form of strict liability[163] in order to align their incentives and reduce the costs of adjudication.

This approach might seem vulnerable to another criticism, however: that it invites strategic behavior by actors when precautions come in lumpy increments. The next Section takes up this point, which is not unique to the UUP proposal but rather emerges whenever actor-selected preventative measures enter into the liability assessment.

---

161. Attention to information costs can also help address another question that the UUP approach raises: What untaken precautionary steps can be unified together as a single "precaution" for purposes of the UUP analysis? As the fence example shows, the UUP binds together two untaken increments: a cost-justified increment that is not causally connected to the accident (first six feet) and a marginally unjustified increment that is causally connected to the accident (last two feet). But surely this does not mean that a plaintiff can meld an unrelated but serious shortfall of the defendant cricket club (such as its failure to maintain the brakes on the club's car) with some highly inefficient precaution that would have stopped the plaintiff's accident (such as a thirty-foot-high fence) under the UUP approach. Instead, the UUP approach calls for a logical unity that makes it appropriate to think of the two pieces as a single precaution. This will be the case when the precautionary steps are so closely related that it would typically be difficult or impossible to disentangle which subset was responsible for the accident. Treating these entwined untaken steps as a unified precaution economizes on information costs and underscores the information function that is performed when an actor takes due care. Once the efficient steps have actually been taken, it is no longer difficult to tell whether they were sufficient to prevent the accident. I thank Michael Livermore and Mildred Robinson for pressing me on this point.

162. A safe harbor tends to produce convergence at the point of safety—no one gains from going beyond what is necessary, but no one wants to be left outside and exposed to liability, either. *See* Morse, *supra* note 84, at 1389-90 (noting this "two-way convergence"). That result depends, however, on perfect information about the location of one's conduct relative to the standard, or, if that information is lacking, symmetrically arrayed consequences for overshooting and undershooting.

163. The UUP approach imposes strict liability to the extent that the defendant may be held liable for some accidents that due care would not, in fact, have prevented. *See supra* notes 142-46 and accompanying text. It is, however, a limited form of strict liability that reaches only those accidents that qualifying untaken precautions would have prevented.

C. Strategic Slicing

Suppose the precautions available to an actor in a given context are not continuous in nature like fence heights, but instead involve binary or lumpy choices about things like buying a new piece of safety equipment or adding another employee.[164] A focus on untaken precautions measured relative to a defendant's actual actions might seem to allow defendants to strategically select precautions that fall short of due care, yet not *enough* short of due care to leave room for a cost-effective and causally effective untaken precaution, given the chunks in which precautions must be taken.[165] There are ways to address this concern, albeit imperfectly.

At the outset, it is worth emphasizing that this issue arises only in cases lacking a clear external standard that defines due care, such as a statutorily required piece of safety equipment. If a life preserver is required on board, the fact that the defendant did something less, such as having a rope on board, will not shield her from liability if the absence of the preserver caused the harm, regardless of how the marginal analysis plays out between rope and preserver or between preserver and doing nothing. The trickier cases are ones in which the location of the due-care standard must be determined based on a Hand formula analysis. Here, a comparison of different possible precautions may be likely to take center stage.

Imagine that a defendant faces a situation similar to that presented in the famous *Carroll Towing* case and must decide whether to hire a bargee who could intervene in the event a boat breaks loose from its moorings.[166] Suppose further that there are three possible

164. Even the fence height examples above assumed a certain degree of lumpiness, insofar as noninteger fences were ruled out. *See supra* Part II.B.2. But it is easy to imagine far less divisible precautions.

165. To keep the analysis as simple as possible for purposes of isolating the implications of lumpiness, the discussion here focuses on scenarios in which precautions are open to only one party, the potential injurer. Numerous other strategic possibilities arise from the interactions between precautionary choices made by different parties. *See, e.g.*, Dhammika Dharmapala & Sandra A. Hoffmann, *Bilateral Accidents with Intrinsically Interdependent Costs of Precaution*, 34 J. LEGAL STUD. 239 (2005); Ehud Guttel, *The (Hidden) Risk of Opportunistic Precautions*, 93 VA. L. REV. 1389 (2007); Alan J. Meese, *The Externality of Victim Care*, 68 U. CHI. L. REV. 1201 (2001); Steven Shavell, *Torts in Which Victim and Injurer Act Sequentially*, 26 J.L. & ECON. 589 (1983). The lumpiness and durability of precautions can play important roles in these contexts as well.

166. *See* United States v. Carroll Towing Co., 159 F.2d 169 (2d Cir. 1947). The example in

precautions: hiring a daytime bargee (Precaution A); hiring a night-shift bargee (Precaution B); and hiring a full-time bargee who would live on the boat and be on call around the clock (Precaution C). Imagine our defendant chooses Precaution A, the daytime-only bargee, and the accident occurs after hours, when the bargee is ashore. Precaution B, the night-shift bargee, or Precaution C, the full-time bargee, would have stopped the accident from happening. Suppose further that, starting from the baseline of a daytime-only bargee, the marginal improvement associated with adding a night-shift bargee would not justify the cost. Can a towing company that hires the daytime bargee immunize itself against the claim that a full-time bargee would have been cost justified, based on the demonstration that adding a night-shift bargee would not be cost justified?

Not necessarily. Although the towing company's precautionary choice might at first look precisely analogous to a cricket club's optimal fence, for which no further improvements would be efficient, there may be indivisibilities or economies of scale associated with paying for a full-time bargee who will live on the boat. Suppose the costs and benefits look like the ones in Table 3.

Table 3. Bargee Precautions

	Cost of Precaution	**Expected Benefit**
A: Daytime Only	100	120
B: Night-Shift Only	100	85
C: Full-Time	180	205
D: A + B	200	205

In this case, hiring one person to be on call around the clock is cheaper than hiring two shift workers to split the work. This makes Precaution C, the full-time bargee, the optimal precaution—even

---

the text is a stylization of the actual facts in *Carroll*. In the case itself, the towing company *had* hired a bargee, just not a very reliable one. *See id.* at 173-74 (noting that the bargee had been absent from the barge without excuse for about twenty-one hours); *see also* Grady, *supra* note 12, at 301 (noting that the slacking bargee is often viewed as the functional equivalent of "not hiring a bargee in the first place (a fictional durable precaution that helps us understand the case in terms of the theory)").

though it would be inefficient to add Precaution B to existing Precaution A.

This example shows why it would be unworkable to use a defendant's actual conduct as the sole baseline from which to assess whether there was an efficient untaken precaution that would have stopped the harm from occurring. In any case involving indivisibilities or economies of scale, the defendant could choose a partial measure like Precaution A and be shielded from liability by the unavailability or nonexistence of a Precaution B that, when added to Precaution A, would meet the criteria of being both causally effective and cost effective. But this result does not follow from the UUP approach described above.[167] The untaken precaution sufficient to establish the negligence liability of the defendant might be either a precaution that could be added to the defendant's actual efforts or a substitute precaution for the efforts that the defendant actually undertook—but only if the *upgrade* is a cost-effective one.

Going back to the bargee precaution numbers in Table 3, we can see that the defendant who has actually hired a daytime bargee will be able to rebut any claim that she should have *also* hired a night-shift bargee. Doing so would cost an extra $100 but save only $85 in accident costs. This is so even though the total benefits of the two shifts together (Precaution D) exceed their total costs. But the plaintiff can nonetheless show that the defendant should have substituted a better precaution (full-time bargee) for the precaution she took (daytime bargee).

Substituting a full-time bargee for a daytime-only bargee is a cost-justified move, and it would have stopped the accident in question. Making the switch requires expending an extra $80 in precaution costs but yields an extra $85 in expected benefits. Thus, a proper analysis will compare what the defendant could have done with what the defendant actually did to see if there is room for an improvement that would have mattered in the case. The defendant who actually hired a bargee for the day shift can indeed shield herself from claims that she should have also hired a night-shift worker, but not from claims that she should have undertaken a

167. *See supra* Part II.B.2.

different approach from the outset: hiring a full-time live-aboard worker.[168]

Lumpiness in precautions may make some kinds of strategizing easier, however. David Gilo and Ehud Guttel focus on the possibility of strategically low activity levels: if installing a particular kind of smokestack becomes cost justified only at a given production level, a factory might hold production below that threshold in order to avoid having to incur the cost of the smokestack.[169] Here, the defendant is effectively choosing one precaution—curtailing her activity level—in order to make another precaution prohibitively expensive relative to its benefits when added to that earlier precaution. But it is not as simple as it was in the bargee case to say that the defendant should have replaced the precaution she selected (suppressed activity) with a different one (smokestack) from the get-go. Activity levels are generally not second-guessed at all in a negligence regime, given the difficulty of a court determining whether a particular level of activity was or was not worthwhile for the defendant.[170] Saying that a defendant should have done a given

168. This conclusion is sensitive to the numbers used in the example. Suppose instead that a change in labor laws causes the cost of a full-time worker to rise to $200, the same cost as purchasing the two shifts in combination. Under the UUP approach, the barge owner who has actually retained a day-shift worker would be shielded from liability because there would not be any marginal improvement that could be made in that case. *See* The Kathryn B. Guinan, 176 F. 301, 301-02 (2d Cir. 1910) (upholding a judgment that a night watchman was not required in addition to the day-shift scow master, albeit on grounds of custom). On the other hand, the barge owner who has not retained any bargee would be held liable for nighttime as well as daytime losses because a full-time bargee (or the combination of day-shift and night-shift bargees) would, taken as a whole, be a cost-justified improvement over what she actually did, which was nothing. *See supra* Table 3.

169. David Gilo & Ehud Guttel, *Negligence and Insufficient Activity: The Missing Paradigm in Torts*, 108 MICH. L. REV. 277, 280-81 (2009) [hereinafter Gilo & Guttel, *The Missing Paradigm*]. Gilo and Guttel argue that their claims are not limited to lumpy precautions, but their theory seems to fit best with such precautions, which also form their primary examples. *Compare* Mark Grady, Response, *Another Theory of Insufficient Activity Levels*, 108 MICH. L. REV. FIRST IMPRESSIONS 30, 30-32 (2009), http://repository.law.umich.edu/cgi/viewcontent.cgi?article=1064&context=mlr_fi [https://perma.cc/S76Y-7YQQ] (contending that Gilo and Guttel's theory depends upon lumpiness), *with* David Gilo & Ehud Guttel, Response, *Insufficient Activity and Tort Liability: A Rejoinder*, 108 MICH. L. REV. FIRST IMPRESSIONS 64, 65 (2009), https://repository.law.umich.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1061&context=mlr_fi [https://perma.cc/J6DJ-DVTQ] ("[T]he cost of precaution need not be 'lumpy' for the insufficient activity result to materialize.").

170. *See, e.g.*, Gilo & Guttel, *The Missing Paradigm*, *supra* note 169, at 278-79. Usually, this incapacity to effectively police activity levels translates into the concern that defendants will select excessive activity levels. For example, a driver might make an unnecessary trip to

activity *more* seems both hard to prove and at odds with usual notions of autonomy.[171]

Similar issues arise with decisions that are distributed over time. An actor may sink costs into precautions that are optimal under current conditions but that are destined to soon become suboptimal. Technological changes might make new precautions superior to ones that were optimal when previously adopted, yet the marginal improvement might not repay the retrofit.[172] Should the fact that the precaution was optimal when selected be an effective rebuttal if tacking on an upgrade later would not be cost justified? In these cases, the costs of strategic behavior and the moral hazard of ignoring future trends must be weighed against the advantages of encouraging people to undertake precautions optimal for the conditions they presently confront. That the law can always step in with regulatory solutions[173] mitigates these concerns to some degree

---

a store for a frivolous reason, when doing so carries trivial value for her compared to the risk that she generates for others even if she is careful. *See* Shavell, *supra* note 55, at 26. Because courts are ill-positioned to determine whether a given trip was valuable or valueless, *see id.* at 25, the driver gets the benefit of the negligence regime's safe harbor regardless. The fact that drivers may be unable to avoid occasional lapses from the standard of due care tends to buffer this effect, however, as does the widespread use of insurance that may be better able to meter and price activity level risks. *See supra* Part I.

171. *See* Kenneth S. Abraham, Response, *Insufficient Analysis of Insufficient Activity*, 108 MICH. L. REV. FIRST IMPRESSIONS 24, 26-28 (2009), http://repository.law.umich.edu/cgi/viewcontent.cgi?article=1065&context=mlr_fi [https://perma.cc/4AR3-BG3G]. To be sure, regulatory solutions can get at this issue from a different direction, arguably with just as much interference with autonomy, although framed less controversially. *See id.* at 26-27 (noting the importance of "optics"). For example, the government could simply mandate a smokestack for widget production operations, and those operators who could not cover its cost at their current production levels would be forced to either increase production or go out of business. *See* Gilo & Guttel, *The Missing Paradigm*, *supra* note 169, at 311.

172. For example, suppose a defendant used a certain type of glass that was state of the art at the time of its installation, T1. Later, at T2, an improved type of safety glass is invented that offers a better overall ratio of benefits to costs. Even though choosing the original type of glass would amount to negligence if undertaken at T2, it is possible that the cost of upgrading from the current type of glass would not be worth the marginal improvement that would be derived from doing so. *Cf.* Trimarco v. Klein, 436 N.E.2d 502, 506 (N.Y. 1982) ("[I]t was ... for the jury to decide whether, at the point in time when the accident occurred, the modest cost and ready availability of safety glass and the dynamics of the growing custom to use it for shower enclosures had transformed what once may have been considered a reasonably safe part of the apartment into one which, in the light of later developments, no longer could be so regarded.").

173. *See* Giuseppe Dari-Mattiacci & Luigi Alberto Franzoni, *Innovative Negligence Rules*, 16 AM. L. & ECON. REV. 333, 338 (2014) (discussing the possibility that old technologies might

by giving actors an incentive to guess correctly about unfolding conditions.

Allowing defendants to capitalize on marginal analysis to dodge liability when, and only when, they have actually taken the intermediate precautionary step forecloses other sorts of opportunism. Consider the much-studied case of *Haft v. Lone Palm Hotel*.[174] Under applicable law, the motel was required to provide either a lifeguard at its pool or a "statutory substitute": a sign warning that no lifeguard was on duty.[175] Plaintiff's decedents were a father and his young son who went into the pool without knowing how to swim, and drowned.[176] There was no lifeguard present and no sign posted about the absence of a lifeguard.[177] The defendants prevailed at trial, but the California Supreme Court reversed and remanded for a new trial, holding that once the defendant's negligent noncompliance was established, the burden shifted to the defendant to show that the lack of a *lifeguard* was not a proximate cause of the deaths.[178]

Although academic treatments disagree about which of the alternative precautions (if either) constituted optimal care,[179] one way of understanding the case is illustrated by the numbers in Table 4.

---

be prohibited by law, making their use negligence per se).

174. 478 P.2d 465 (Cal. 1970).

175. *See id.* at 470-72, 470 n.8.

176. *See id.* at 466-67.

177. *Id.* at 467-68.

178. *See id.* at 470.

179. *Compare* Grady, *supra* note 20, at 822 (stating that the sign has "at least a plausible claim" to be the optimal care level, especially given the cost of retaining a lifeguard during the off-season, when the accident occurred), *with* Wright, *supra* note 20, at 88 (suggesting that the lifeguard was viewed as the optimal precaution, with the statute allowing the sign, if posted, to support an assumption of risk defense). Levmore's "recurring misses" analysis treats the failure to post the sign as the relevant failure. *See* Saul Levmore, *Probabilistic Recoveries, Restitution, and Recurring Wrongs*, 19 J. LEGAL STUD. 691, 705-10 (1990).

Table 4. Swimming Pool Precautions

	Cost of Precaution	**Expected Benefit**
Sign Only	100	120
Lifeguard Only	500	510

Assume that both the sign and the lifeguard are lumpy on-off choices, and that it is only sensible to do one or the other, not both.[180] Even though both precautions "pay for themselves" in that total benefits exceed total costs, the sign is the optimal precaution on the numbers given. This can be demonstrated by considering an "upgrade" from the sign to the lifeguard. It would cost an extra $400 but would produce only an extra $390 in benefits. Thus, a defendant who had posted the sign would have no difficulty showing that it was not negligent to stop there—even if there were no statute on the books to back up her choice. But a defendant who had not taken either of the precautions would have no such defense. The plaintiff could show that the lifeguard would have been a cost-justified precaution compared to what the defendant actually did since an "upgrade" from doing nothing entails benefits of $510 and costs of only $500.

The rationale for this approach, and the strategic possibilities that it avoids, become clearer if we imagine a variation on the *Haft* facts. Suppose there were two *equally cost-effective* ways of meeting the due-care standard[181]: providing a life buoy at poolside, and making life jackets available for checkout at the motel desk. With respect to any particular drowning, it might be clear that one precaution would have been causally effective and the other would not have been. For example, the person who drowned might have been an unaccompanied nonswimmer who would have almost certainly worn an available life jacket (based on well-documented past behavior) but who did not have anyone on hand to throw him a life buoy.[182] If the motel takes neither precaution and a motel

180. Although it would be possible to hire a part-time lifeguard and post the sign for the balance of the time, the fixed costs of hiring the lifeguard might make a full-time position the only viable alternative.

181. I am indebted to Ariel Porat for raising this point.

182. Or, conversely, a victim with a history of eschewing life jackets might have benefited from a life buoy due to the presence of a nonswimming companion.

guest drowns, the motel should not be able to strategically specify which of the two precautions it "would have" taken had it not been negligent, so as to evade a showing of causation. The potential for underdeterrence is clear if this were permitted, even though the motel can meet the due-care standard by actually taking either of the precautions.

The *Haft* court's reasoning, and its ultimate bottom line, afford more than one possible interpretation. It is almost (but not entirely) certain that a posted sign would not have prevented the drownings.[183] It is also very likely that a lifeguard on the scene could have saved the pair. If an HWN rule applies,[184] the difference is academic. But if the rule is HBN,[185] one of two difficulties must be confronted. One must either find a way to get past the slim odds that the sign would have saved the decedents (fudging causation), or else find a way to impose liability for the failure to provide a lifeguard (fudging due care). Saul Levmore has focused on the former alternative under the head of "recurring misses" since it is, after all, conceivable that some rare individuals would be deterred by a sign from going into an unattended swimming pool.[186] But the second alternative fits neatly into Grady's analysis: liability attaches for falling far enough short of due care to afford space for a cost-effective and causally effective alternative—the lifeguard.[187]

Perhaps it is not really necessary to choose between these two alternatives in understanding *Haft*. One might instead return to the idea of a "harming while negligent plus" (HWNP) regime,[188] where the plus factor can be either the omission of a precaution that meets causation and efficiency criteria when compared with the plaintiff's conduct, or the omission of the optimal precaution under circumstances where it *might* have made a difference. The next Part delves further into this question of omitted precautions that might have, but did not necessarily, cause harm. Here, we see how another form of aggregation—actually or conceptually repeating the interaction—sheds light on the evaluation of risk.

183. *See* Levmore, *supra* note 179, at 705 (putting the odds at about 10 percent).
184. *See supra* Part II.B.
185. *See supra* Part II.B.
186. *See* Levmore, *supra* note 179, at 705-06.
187. *See* Grady, *supra* note 20, at 822-23 (discussing *Haft* in terms of his theory).
188. *See supra* notes 155-57 and accompanying text.

III. REPEATING RISKS

The cricket fence example was atypical in that it permitted certainty about whether a given accident would have been prevented by a fence of a given height. In many scenarios, by contrast, we know only that a given accident would have been prevented with some probability if the actor were not negligent. To use one of Cooter and Porat's examples, suppose that the negligent heating of a vat used to make hot chocolate causes a valve to crack and cause harm.[189] But there is only some positive probability that the crack and resulting harm were caused by the defendant's negligence; the valve would crack a certain number of times when heated nonnegligently.[190] What to do in that case? One possibility is to charge the accident to the defendant when the defendant was negligent if (and only if) it is more likely than not that the negligence caused the harm.[191]

This is an answer, but it may not be a satisfying one if the same accident-causing procedure will be negligently repeated over and over, with the negligence causing and not causing accidents in stable and predictable proportions.[192] Figuring out what to do about such probabilistic cases involves another form of evaluative aggregation, one that asks what is the right set of accidents—or potential accidents—to hold in mind in assessing liability.

A. Compounding and Offsetting

A standard move in the economic analysis of tort law is to create a kind of conceptual population explosion to assess how a particular approach will work when scaled up. Instead of looking just at the

189. *See* COOTER & PORAT, *supra* note 20, at 20.

190. *See id.*

191. *See id.* at 20-21.

192. *Cf.* David Rosenberg, *The Causal Connection in Mass Exposure Cases: A "Public Law" Vision of the Tort System*, 97 HARV. L. REV. 849, 858 (1984) ("[T]he [preponderance-of-the-evidence] rule is neither a rational nor a just means of resolving the systematic causal indeterminacy presented by mass exposure cases involving defendants whose tortious conduct has caused or will cause a statistically ascertainable increase in the incidence of a particular disease.").

case at hand, where a defendant's negligence might have caused the plaintiff's injury (but probably did not), one hauls in a huge set of replicating mirrors and imagines the same scenario playing out again and again and again, hundreds or thousands of times.[193] Mass-producing the interaction between plaintiff and defendant generates large-number trials that make probabilistic risk more tractable.[194] Doing so also helps to identify and highlight systemic effects that could not be observed in a one-off case.

Consider how this operation plays out in the vat example introduced above. If the probability that the defendant's negligence was to blame for a cracked vat is always, say, 40 percent, a more-likely-than-not standard will produce the phenomenon Levmore has termed "recurring misses."[195] Out of one hundred repetitions of the accident, forty will be caused by negligence, but none will be charged to the defendant.[196] This can produce an insufficient degree of deterrence.[197] The opposite problem, which we might call "relentless hits," occurs if the probability that negligence was to blame is always, say, 60 percent.[198] Here, the defendant is always charged with the accident even though her negligence was not to blame in forty of the one hundred cases. This produces a cliff effect similar to the one discussed above, which again puts increased pressure on reaching the due-care standard.[199]

The cracked vat example goes to uncertainty about the causal role of negligence, where it is clear that the defendant's acts (hot

193. In some cases, such as mass toxic exposure cases, this exercise of imagination is not necessary—the interaction has already been mass-produced in reality. *See generally id.*

194. It also subtly changes the framing of the risk being produced. Where an individual interaction might generate merely a risk of harm to a given individual, the replication of this interaction enough times produces the certainty of harm to a roughly determinate number of (unspecified) individuals. As Barbara Fried has recently argued, we cannot categorically equate such statistical risk creation with the creation of certain harm on a given occasion (for to do so would rule out everything, no matter how low the risk, if it were repeated enough times), nor can we simply ignore the risk in all cases, but rather must take into account the level of risk. Barbara H. Fried, *Facing up to Risk* 17-19 (Stanford Pub. Law, Working Paper No. 2850587, 2016), https://ssrn.com/abstract=2850587 [https://perma.cc/QQ8M-HE2N].

195. *See* Levmore, *supra* note 179, at 705-10.

196. *See id.* at 705-06.

197. *See id.* at 706.

198. *See* COOTER & PORAT, *supra* note 20, at 20-22; *see also* Ariel Porat, *Misalignments in Tort Law*, 121 YALE L.J. 82, 109-12 (2011).

199. *See* COOTER & PORAT, *supra* note 20, at 20-22; *see also supra* Part II.B.

chocolate making and associated vat-heating) caused the harm. Causal uncertainty can also run to whether a particular defendant's acts caused the plaintiff's harm *at all*.[200] Whether causal uncertainty relates to the causal effect of the defendant's negligence specifically or of the defendant's acts more generally, a crucial question is whether replicating interactions like the one in question will tend to produce compounding or offsetting errors. As a first cut, we must consider what is meant by "interactions like the one in question." Obviously, if the interaction before us involves a 40 percent probability of a causal connection, repeating that exact same scenario hundreds of times will compound the error and produce a 40 percent error rate if we stick with a preponderance-of-the-evidence standard. So we cannot quite mean that.

The question, rather, is whether the scenario before us, with its particular level of causal uncertainty, is representative of the kinds of cases that we would expect to see replicated over time. Is there something about the vat apparatus (such as a stable background failure rate when heated properly) that will *always* make it 40 percent likely that the negligently rapid heating is what caused the vat to crack?[201] Or, instead, should we expect that the kinds of interactions that defendants and plaintiffs like the ones before us will have over time will tend to occupy a broad spectrum of probabilities? In many settings, we might expect the errors introduced by a rote application of the preponderance-of-the-evidence standard to roughly offset, with the 40 percent cases that generate no liability counterbalanced by the 60 percent cases that generate liability. We may accept some level of error in individual cases but become alarmed when scaling up suggests a systematic pattern of either recurring misses or relentless hits.[202]

The significance of such systemic causal errors varies, depending on whether we are talking about attribution errors that run to negligence-causation or more broadly to act-causation. In the former case, relentless hits merely amount to a pocket of strict liability, as

200. *See infra* Part III.C.

201. Levmore makes this point about the facts in *Haft*. *See* Levmore, *supra* note 179, at 706 ("The background statistics on such matters as drowning and the efficacy of lifeguards and signs are sufficiently stable to ensure that the preponderance rule will systematically 'miss' ongoing instances of antisocial behavior that it should deter.").

202. *See, e.g.*, *id.* at 705-06.

will be discussed below, and recurring misses are relatively unproblematic to reach without introducing serious distortions. Systematic problems with act-causation, although occurring in relatively few domains,[203] present the more troubling specter of holding defendants liable when their acts were not causative at all or, alternatively, letting defendants systematically off the hook when they did cause, on average, a certain quantum of harm. Various alternatives have been tried or proposed to reach these issues, including prorating damages to account for uncertainty or shifting the burden of proof.[204] Unsurprisingly, none of these approaches is fully satisfactory. There is a fundamental tension between the scaling up that is necessary to clearly see (and potentially solve) these problems and the one-to-one pairing of wrongs and injuries that animates tort law.[205]

An especially compelling variation of this basic problem is illustrated by Jonathan Cohen's "paradox of the gatecrasher," in which 1000 people are seated in a rodeo arena but only 499 have paid the admission fee—and it is impossible to tell which ones.[206] Under a preponderance-of-the-evidence standard, each person at the rodeo (and hence *every* person at the rodeo) could be ejected or sued for the admission fee, since it is more likely than not in each case that the person is a gate-crasher rather than a paying customer.[207]

---

203. In some cases, the uncertainty is due to a high background risk given by nature, which may be either episodically high (persons overboard on rough seas) or persistently high (persons being treated for late-stage cancers). In other cases, it is clear that harm was caused by some defendant, but there is uncertainty about which one (cases of market share liability or alternative liability). *See infra* Part III.C.

204. *See, e.g.*, Summers v. Tice, 199 P.2d 1, 3-5 (Cal. 1948) (shifting the burden of proof in an alternative-liability scenario, where both defendants were negligent but only one could have caused the harm); Herskovits v. Grp. Health Coop. of Puget Sound, 664 P.2d 474, 486-87 (Wash. 1983) (Pearson, J., concurring) (detailing a damage-prorating approach based on reduced chance of survival, drawing on analysis in Joseph H. King, Jr., *Causation, Valuation, and Chance in Personal Injury Torts Involving Preexisting Conditions and Future Consequences*, 90 YALE L.J. 1353 (1981)). For a survey of possible approaches to recurring misses, see Levmore, *supra* note 179, at 706-10.

205. *See infra* Part III.C.

206. L. JONATHAN COHEN, THE PROBABLE AND THE PROVABLE 75 (1977). This problem has generated an extensive academic literature. *See, e.g.*, Levmore, *supra* note 179, at 694-95, 695 n.6 (discussing this scenario and citing prior literature on it).

207. *See* COHEN, *supra* note 206, at 75. To be sure, a no-liability result could be premised here on the promoter's contribution to the information shortfall—its failure to come up with an appropriate system of ticketing or stamping hands to distinguish paying guests from gate-

Here, multiplication of uncertainty generates a manifestly absurd level of injustice, yet it is no different in kind from that which can and indeed must occur in the individual close case.[208] The difference is that the act of multiplication enables us to see a larger slice of the world and observe its systematic skew.

That still does not quite explain why what is tolerable in the single case may be intolerable in the aggregate. An economic explanation would be that as long as cases generally wash out over time, incentives will be in at least rough alignment, and we can do no better than stick with the preponderance-of-the-evidence test.[209] On this account, close cases are tolerated on the assumption that there is no systematic skew overcharging or undercharging defendants. When the curtain is pulled back to show that such a skew exists (within some identifiable corner), it becomes plain that our hope for mistakes to wash each other out is unfounded.

Yet why stop with *this* rodeo when setting the bounds for acceptable error cancellation? Suppose there is another rodeo arena across town, run by the same inept promoters, that is simultaneously seating 1000 persons, 501 of whom bought tickets, and 499 of whom are gate-crashers. Here, recovery would be impossible against any of the rodeo-goers on a preponderance-of-the-evidence standard, even though nearly half are gate-crashers. The errors cancel if we consider the rodeos together, but we hardly feel better. The promoters recover the right amount, but not from the right people. Repeat this pair of rodeos hundreds of times and gate-crashers find themselves getting ejected regularly, but paying customers are ejected almost as often.

---

crashers. *See* Levmore, *supra* note 179, at 694-95. The analysis in the text assumes that no such precautions were available, which might require positing some additional facts such as excellent ticket counterfeiting by the gate-crashers or unexpectedly disappearing ink for handstamps.

208. *See* COHEN, *supra* note 206, at 75-76. A possible response might be that facts tend to cluster in a relatively lumpy fashion, making hairline close cases a rarity, with most fact patterns yielding answers far above or below the 50 percent mark. *Cf.* Bovik, *supra* note 28, at 41 fig.4 (providing image histograms used in thresholding to illustrate the difference between "[w]ell-separated modes" and "[p]oorly separated or indistinct modes" with respect to gray levels). However, the *tried* cases may tend to cluster closer to the 50 percent mark, regardless of the overall distribution, due to selection effects. *See generally* George L. Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 J. LEGAL STUD. 1 (1984).

209. *See, e.g.*, Levmore, *supra* note 179, at 693-98.

Perhaps what makes the aggregate case more viscerally problematic than the individual case is simply that aggregation changes the unit of analysis when we begin to examine mistakes. In an individual one-off, we have someone who is (in the fact-finder's eyes) partly at fault and partly not; a binary decision in either direction does not strike us as a severe miscarriage of justice (in a civil case, anyway) if for no other reason than we can do no better. Scaling up changes what was a probabilistic assessment about an individual to a social judgment that is 100 percent wrong as to an identifiable number of actual people (even if we cannot say which ones). With an individual in a close case, we doubt a mistake is being made at all; we think anyway it is less likely than not. With a crowd, we are certain that mistakes are being made, and being made at the whole-person level for a number of identifiable individuals.[210]

There is at least one setting, however, in which we can formulate a relatively nondistortive approach to causal uncertainty: the case in which act-causation is clear but the causal role of negligence is not. The next Section explains how understanding due care as a step good helps to resolve the difficulties surrounding negligence-causation.

B. The Bridge of Due Care

Consider again the case of the negligently heated hot chocolate and the cracked vat, where it is 40 percent likely that the negligence caused the crack and resulting damage. If we take a "harming because negligent" (HBN) approach and require preponderance of the evidence, the negligent chocolatiers escape liability. And if we repeat the experiment a hundred times, they would escape liability every time, while their negligence would have without doubt caused forty of the cases of harm. It seems as if we must choose between letting them get away with negligence repeatedly and adjusting the required causal relationship.

A "harming while negligent" (HWN) standard escapes these vexing difficulties because it loosens the fixation on whether negligence caused a particular harm and asks instead whether negligence was

210. *Cf.* Fried, *supra* note 194 (discussing how risk translates into statistical certainty of harm when scaled up).

present and the act caused the harm.[211] It is worth reconsidering the doctrinal and normative objections to this approach, to see whether they hold water. Here we might again consider viewing the negligence rule not as defining the outer limits of liability, but rather as setting out a safe harbor for which a defendant can qualify only if she fully meets the standard of due care.[212] As suggested above, this recasting would treat due care as a kind of lumpy or step good that must be supplied in full in order to deliver its liability-constraining benefits to the actor.

One rationale for treating due care as a step good is that it has certain public good characteristics that are realized only when it is supplied in full. When an actor satisfies the due-care standard, not only are the costs of accidents (including prevention costs) minimized, society is saved the trouble of having to figure out whether her shortfall (as opposed to her act in general) was causally responsible when harm results. In short, the actor is delivering an information-cost benefit to society by fully meeting the due-care standard.[213] The reward for providing this benefit is that she will be immune from liability even though her acts in the world will continue to generate accidents, albeit efficient ones. On this view, the law truncates an actor's liability once she meets the standard of due care not because of any normative principle that people are entitled to freely externalize all harms that due care would not prevent, but rather because she has qualified herself for special treatment by supplying the valuable lumpy good of due care.

There are at least two drawbacks to this approach. First, it may be difficult for actors to know when the due-care standard has been met, especially when it does not correspond to any obvious discontinuity in the world. Keeping a vat of chocolate below the boiling point is one thing, but making sure that heating occurs at a certain speed or with a certain frequency of monitoring is another.[214]

---

211. *See supra* Part II.B.1.

212. *See supra* Part II.B.2.

213. For a recent discussion of the role of information costs in tort law, see generally Jennifer Arlen, *Economics of Tort Law*, *in* 2 THE OXFORD HANDBOOK OF LAW AND ECONOMICS 41 (Francesco Parisi ed., 2017).

214. *See* COOTER & PORAT, *supra* note 20, at 17-20, 18 fig.1.1, 20 fig.1.2 (providing examples involving "natural continuity" (a vat leak) and "natural discontinuity" (a boil that spoils the batch)).

Second, and closely related, it is nearly impossible for actors to avoid falling below due care now and then, and such lapses can lead to large amounts of liability.[215] These problems exist under any negligence standard, but they are sharpened under an HWN standard because a moment's lapse or a trivial miscalculation can expose an otherwise careful defendant to unlimited liability for not just the harm that was caused by her shortfall, but also for all the harm she causes that could not have been prevented by due care.[216]

An intermediate position between HWN and HBN would assess liability only when there is at least an appreciable chance that the defendant's negligence caused the harm in question. If it is obvious that due care would have done nothing to stop the harm from occurring, then recovery would not be available. This fits well with the information-cost account above because fact-finding difficulties will tend to cluster around cases where the defendant's negligence could have plausibly caused or not caused the accident, not the outlier freak accidents that due care could not have done anything to prevent. Such an "appreciable chance" approach in cases of causal uncertainty could be coupled with the unified untaken precaution approach (UUP) above to short-circuit many factual difficulties and avoid the distortions that the "zero liability at due care" ledge would otherwise produce.[217] And it may be quite close to the approach courts already take.

Consider the case of *Maddocks v. Bennett*, a 1969 Alaska Supreme Court decision involving an alleged allergic reaction to a beauty salon's hair dye treatment.[218] The manufacturer's instructions for the dye product specified use of a patch test twenty-four hours in advance to screen for allergic reactions, but the patch test was omitted.[219] The plaintiff first began remarking symptoms of itching and fatigue about twenty-nine hours after the dye treatment (although she also reported having felt unwell or "funny" for much of that day).[220] The symptoms subsequently escalated: the plaintiff's eyes

215. *See supra* Part I.
216. *See supra* Part II.B.1.
217. *See supra* Part II.B.2.
218. 456 P.2d 453 (Alaska 1969).
219. *Id.* at 454.
220. *Id.* at 461.

swelled, her scalp itched and burned, and all of her hair fell out.[221] The defendant conceded that failure to perform the patch test twenty-four hours in advance was negligent, but disputed that a properly administered patch test would have revealed an allergic reaction in time to stop the planned dye treatment.[222]

The court held that the plaintiff bore the burden of "show[ing] the truth of the following statement: if the patch test had been given, results would have occurred within the 24 hour waiting period indicating an allergic reaction."[223] The court did not deem the delayed emergence of symptoms (twenty-nine hours after the treatment) dispositive since had a patch test been given, it would have been on a different and more sensitive area of skin, and there would have been active monitoring for—as opposed to passive noticing of—any allergic reactions.[224] The court concluded that "[b]ecause a reasonable person could conclude that more likely than not the test would have shown some indication of an allergic reaction, appellee did establish cause in fact sufficiently to take the case to the jury."[225] While this outcome nominally heeds both the allocation of burden of proof to the plaintiff and the use of the more-likely-than-not standard for cause in fact, sending the case to the jury is very likely (given hindsight bias, a clearly negligent defendant, and a clearly harmed plaintiff) to yield a recovery under circumstances where the patch test *could* have made a difference.

Contrast this case with *Peterson v. Nielsen*, a 1959 Utah Supreme Court decision involving a highway accident.[226] In that case, the defendant was clearly negligent for pulling out in front of the plaintiff as she traveled on an arterial highway, and the only question was whether the plaintiff's speed (which was somewhat in excess of the legal limit) was a contributing factor that would bar her recovery.[227] The court answered this question in the negative because it found that she could not have avoided the accident had

221. *See id.* at 454.
222. *See id.* at 459.
223. *Id.* at 460.
224. *See id.* at 460-61.
225. *Id.* at 461.
226. 343 P.2d 731 (Utah 1959); *see also* Wright, *supra* note 20, at 86-87 (discussing this case).
227. *Peterson*, 343 P.2d at 732-33.

she been traveling at the legal speed: "In the instant case, any reasonable analysis of the specific findings as to speeds and distances, and the facts necessarily incident to them, will show that the plaintiff could not have avoided the accident by exercising due care."[228] After carefully reviewing the evidence on skidmarks and distances, the court concluded that "she would have collided with him had she been traveling at the lawful speed, or even at a considerably lesser one."[229]

What this pair of cases suggests is this: where an actor was negligent and that negligence could have had no effect on the accident that occurred, there will be no liability; but when it is an open question whether the negligence could have mattered, that may be enough to get to the jury.[230] Ruling out instances in which an actor's negligence could have had no impact softens the cliff effect of a pure HWN standard, without creating the distortions and information-cost issues of a pure HBN standard.[231] Imposing liability when the actor was negligent, caused harm, *and the negligence might have mattered* thus constitutes another flavor of a "harming while negligent plus" (HWNP) approach.[232] Here, the plus factor is an epistemically difficult to reach but factually plausible

228. *Id.* at 733.

229. *Id.* at 734. Wright presents this case as a rebuttal of Grady's claim that courts will impose liability when a cost-effective and causally effective precaution has been identified, relative to what the actor actually did. *See* Wright, *supra* note 20, at 86-87. In fact, it does not appear that any such precaution was ever identified in the case. No slower speed was named that would have stopped the accident, so there was no analysis about whether dropping to that speed would have had benefits in excess of costs when considered as a unit from the baseline representing the driver's actual speed. Even if this had been shown, however, the existence of a statutory speed limit might well have afforded a safe harbor to the actor if it were clear that the harm would have occurred even at that speed. This example shows why it matters whether we understand the UUP approach to be actually shifting the due-care line (a move that becomes unavailable when an external standard for due care exists) or instead altering the causation requirement associated with a fixed due-care line (which could be doctrinally combined with a fixed due-care standard such as a speed limit). *See supra* text accompanying notes 152-56.

230. Wex Malone made a similar observation over sixty years ago in the context of omitted fire safety measures. Wex S. Malone, *Ruminations on Cause-in-Fact*, 9 STAN. L. REV. 60, 78 (1956) ("[W]hen the defendant is unable to make a positive showing that a compliance with the law would have been futile under the circumstances, and where the matter is fairly open to speculation, the issue of cause almost always reaches the jury.").

231. *See supra* Part II.B.

232. *See supra* text accompanying notes 156-57.

causal relationship between the negligent increment of the actor's conduct and the harm that resulted.[233]

A discontinuity in liability remains under this approach, insofar as a defendant can be liable for more than what her negligence actually (more likely than not) caused. But this discontinuity is automatically softened in another way: the closer one is to due care, the fewer one's expected accident costs will be because the frequency and severity of accidents will in fact be reduced as one approaches due care. Contrast this result with the cliff effect that the law might construct in a criminal or regulatory context, in which being only one hairsbreadth over a given legal limit or one hairsbreadth beneath the applicable standard could yield results that are just as severe as being quite far over the line or below the standard.[234] In tort, one is not held liable merely for generating undue risk; one must also *actually cause harm*.[235] Thus, one gains an expected benefit by coming nearer to due care than by remaining further from due care by virtue of the effects one produces in the world.

These gains are concretely experienced only by those engaged in a great deal of repeat play—or those who have access to insurance that creates similar results—because only through the law of large numbers will actual accident costs come to approximate expected accident costs. For most people, accidents are infrequent, lumpy events. The theoretical ability to reduce accident costs through close-to-optimal care will be of little comfort to someone who has the misfortune of having a minor miscalculation or momentary lapse generate catastrophic loss.[236] But it must again be emphasized that

---

233. Such information concerns appear to have driven the analysis in *Zuchowicz v. United States*, 140 F.3d 381 (2d Cir. 1998). In that case, a patient was negligently prescribed an overdose of a drug. *Id.* at 384. There was strong evidence that the drug caused the patient's injury, but it was much less clear that the drug *overdose* (that is, the defendant's negligence) was specifically responsible. *Id.* at 390-91. This epistemic gap, Judge Calabresi concluded, did not preclude a finding of liability:

> [W]hen a negative side effect is demonstrated to be the result of a drug, and the drug was wrongly prescribed in an unapproved and excessive dosage (*i.e.* a strong causal link has been shown), the plaintiff who is injured has generally shown enough to permit the finder of fact to conclude that the excessive dosage was a substantial factor in producing the harm.

Id. at 391.

234. *See supra* Part II.B; *infra* Part IV.C.

235. *See, e.g.*, Wright, *supra* note 20, at 67.

236. *See, e.g.*, Jeremy Waldron, *Moments of Carelessness and Massive Loss*, *in* PHILOSOPH-

this potential result is already a feature of the tort system as we know it and would exist even under an HBN regime with very demanding causation requirements.[237] Addressing that larger issue brings us full circle to questions of aggregation in evaluating conduct—the place we began in studying lapses.

C. Matching Up Accidents

I have deferred until now the most difficult type of causation problem: the one that arises when it is unclear whether the defendant's risk-generating acts caused the harm at all. This question of act-causation presents issues distinct from those raised above, where it was certain that the defendant's acts caused the harm and the only question was whether her negligence did so. Act-causation emerges as a problem when the background risk given by nature or produced by the acts of other parties make it difficult to attribute causal responsibility to the defendant.[238]

The conceptual exercise of repetition has interesting effects here. If the defendant was negligent, and we imagine her doing an act over and over again with harm (of a sort that her negligence would be expected to cause) repeatedly occurring, we might surmise that she caused some of that harm, even if we do not know which specific harms were her doing. Instead of asking whether she was more likely than not to have caused each of the harms (a question that will in some contexts give us a repeated negative answer), we should ask: Which of these harms are *most likely to be hers*? If we could see the full universe of possibly caused harms, and we also knew roughly the amount of harm the defendant was responsible for, we could look for features that would tip us off as to which ones were more likely *than the others* to have been the fault of this defendant.

ICAL FOUNDATIONS OF TORT LAW 387 (David G. Owen ed., 1995).

237. *See id.* at 398-401 (describing a scenario in which a defendant's negligence clearly caused the accident, but nonetheless generates an amount of liability that seems disproportionate to the severity of the lapse).

238. These problems tend to cluster around drownings, toxic exposures, and medical misadventures—settings where background risk is high, other potential blame-bearers are prevalent, or both. *See, e.g.*, Rosenberg, *supra* note 192, at 856-58 (noting these problems in mass toxic exposure cases).

We plainly have no such capacity, yet the ordinary preponderance-of-the-evidence standard can be understood in something like these terms. By partitioning cases based on whether the harm was more likely than not caused by the defendant (the usual focus of our attention), the standard also automatically identifies which defendant-harm pairings are more likely *than other possible defendant-harm pairings*. If we wish to maintain an accident-centered system, then we need a thresholding algorithm that will charge a defendant with the amount of harm that she (statistically speaking) caused. This happens automatically under a preponderance-of-the-evidence standard when the strength of the causal connections between harms and injurers are symmetrically arrayed around the centerpoint.

In settings where causation will always be less likely than not, yet the defendant is plainly responsible for some of the harm, a different approach is required. Proportional liability seems like a logical response.[239] But it fits uneasily within an accident-centric system.[240] What should we do, for example, when it is absolutely clear that a particular defendant could not have caused *this* plaintiff's harm? The relentless application of a proportional liability rule in this context means abandoning any semblance of a causation requirement, but relaxing it risks distortions in the overall system

239. Under this approach, damages are prorated to factor in causal uncertainty. For example, if a doctor's negligence was 20 percent likely to have caused the plaintiff's death, the plaintiff's estate would recover 20 percent of the amount of damages that would normally be available in such a case. If repeated over time, the doctor would ultimately be liable for amounts equaling 20 percent of the deaths, even though every plaintiff would recover in part and no plaintiff would recover in full. *See, e.g.*, Levmore, *supra* note 179, at 692, 697-98. A different way of characterizing proportional liability is to redefine the harm as a "loss of a chance" of survival—an approach that allows a court to retain a preponderance-of-the-evidence standard. *See, e.g.*, Matsuyama v. Birnbaum, 890 N.E.2d 819, 832 (Mass. 2008). But when the injury is defined in this way, conceptual consistency would seem to require opening the courthouse doors to those who had suffered no tangible ill effects at all, if a doctor's negligence nonetheless made their survival less likely—a move that seems foundationally at odds with an accident-based system. I thank Bert Huang for discussions on this point.

240. Proportional liability is just one possible node along a spectrum of aggregative approaches that would shift the focus from the individual accident to the systemic effects of risk creation. For arguments in favor of aggregative "public law" mass tort adjudication processes, see Rosenberg, *supra* note 192.

of deterrence.[241] An alternative would attempt to deliver incentives by finding the relatively best defendant-harm pairings.

Here, instead of tort law's more-likely-than-not inquiry, we would ask a slightly different question: Is the strength of the causal connection between the harm and the defendant in the case before the court stronger or weaker *than the average causal connection* between defendants engaged in this type of activity and harms of this type?[242] If desired, the verbal formulation could be adjusted to require a larger gap between the average causal connection and the liability-generating one. The point is a basic one: if we are trying to match up defendants who cause harm with the harm that they cause, it is the *relative* strength of the causal connection, and not its absolute strength, that should matter. This corresponds to the intuitive inquiry of whether there will be better opportunities than this one to hold the defendant to account for her risk-generating behavior.[243]

The answer to that question directs us to a type of bundling that is implicit in the repetition exercise. Many kinds of precautions are "durable" in that they cannot be readily tailored across situations and conditions.[244] For example, the preventative measure of putting a life buoy on a boat before it goes out to sea cannot be altered during the boat's journey depending on factors that might bear on its efficacy.[245] The buoy on board in calm, predator-free waters is bundled with the buoy on board in stormy, shark-infested waters.[246]

---

241. *See, e.g.*, Hymowitz v. Eli Lilly & Co., 539 N.E.2d 1069, 1076-78 (N.Y. 1989) (considering and rejecting exculpation in market-share liability litigation).

242. This inquiry closely resembles one method of "adaptive thresholding" in image manipulation. *See* Fisher et al., *supra* note 30 (illustrating how using "the *mean* of a 7×7 [pixel] neighborhood" as the local threshold generates much better results than applying a global threshold where there is "a strong illumination gradient" in the image).

243. The question of whether there is a better scenario or better plaintiff is a common one in tort (and other) law and can explain a variety of doctrines. *See, e.g.*, Levmore, *supra* note 179, at 705-10 (discussing the problem of "recurring misses," which assumes a dearth of opportunities to hold the defendant to account); Richard M. Re, *Relative Standing*, 102 GEO. L.J. 1191 (2014) (discussing the advantages of a "most interested plaintiff" standing rule over an "adequacy-based approach").

244. *See* Mark F. Grady, *Marginal Causation and Injurer Shirking*, 7 J. TORT L. 1, 16-20 (2014) (discussing the significance of durability and divisibility in precautions).

245. *See id.* at 16-17 (using an example involving a life buoy and heterogeneity in swimming ability to make this point).

246. Grady terms this a "victim-aggregating effect," but a precaution need not actually aggregate different victims as long as it aggregates a variety of possible accident scenarios

Failing to supply the buoy generates a real risk of liability because its absence in the calm, predator-free waters is very likely to make a difference, even if it is unlikely to matter in the state of the world where the swells are high and sharks circle. Here, the inability of the defendant to choose precautions separately in the two states of the world helps to preserve appropriate incentives.[247] Often the buoy would not have mattered, but we are able to easily pick out the times when it was most likely to have mattered by using a more-likely-than-not standard.

But suppose that a boat *only* plies stormy, sharky waters while perpetually carrying a crew of terrible swimmers.[248] It is possible that leaving off the life buoy in this case is not actually Hand formula negligent[249] at all, but let us suppose that the buoy is so inexpensive and occasionally so effective that it is negligent not to include it. Now we want to ask not whether it is more likely than not that the missing buoy caused an overboard sailor's death, but rather whether *this* overboard sailor's death was more likely (or, perhaps, *much* more likely) to have been caused by the missing life buoy than the *typical* such overboard sailor's death occurring under the conditions in which the boat regularly operates.

A more straightforward example appears in medical contexts that are characterized by a high background risk of death.[250] Here too, we want to identify the harms that were the most likely ones, among those occurring in this specialized high-risk setting, to have been caused by a doctor's error. That might be done by lowering the liability threshold to something close to the average causal connection

---

involving different amounts of background risk and hence different degrees of precaution efficacy. *See id.* at 26.

247. *See id.* at 16-20. The implicit bundling associated with durable precautions is not always beneficial, however. For example, it might dictate carrying a bulky piece of safety equipment that is not cost justified most of the time and only gets in the way of the crew, but is so useful on rare occasions that it is cost justified overall. In such a case, the equipment can be analogized to a cricket fence that is *only* available in a height that exceeds the optimum. Conversely, a given precaution might be extremely efficacious under certain rare circumstances, but because these circumstances are so rare, the precaution is *not* cost justified overall and will not be provided—even though it would be worthwhile to provide if it could be selectively conjured up just in the states of the world where it is useful.

248. *Cf. id.* at 19 (presenting a hypothetical in which a boat owner can strategically tailor the provision of a life preserver to the swimming abilities of the boat's crew).

249. *See supra* note 16.

250. *See, e.g.*, Herskovits v. Grp. Health Coop. of Puget Sound, 664 P.2d 474 (Wash. 1983).

between a doctor's negligence and negative medical outcomes. If a doctor's negligence would, if repeated, cause death in twenty cases out of one hundred, a court might ask whether, in the case before it, the causal connection between the defendant's negligence and the plaintiff's harm exceeds (or, perhaps, *significantly* or *greatly* exceeds) the baseline causal connection of 20 percent.

We might worry that making doctrinal adjustments like this one will deter people from entering fields that involve high background risk, such as medical specialties where survival chances are chronically low.[251] The concern is compounded if we also think that these same fields involve a high risk of lapses, relative to other fields. This is similar to the worry that people will underutilize valuable technologies if they generate many opportunities for human lapses, even if they increase overall safety or survival chances.[252] The answer to all these points is essentially the same: we may wish to subsidize people who enter certain fields or use certain technologies to account for the overall beneficial effects of these choices on social welfare.[253] What we want to leave in place, however, are incentives for people to take appropriate levels of care at the margin.

If we wish to deliver incentives through a tort system that makes binary judgments at the accident level (an open question), we will at times need to settle for a rough and functional understanding of causation. Just as the thresholds for converting grayscale pixels to black and white may need to be adjusted within portions of an image to compensate for the effects of light and shadow,[254] we may need to alter the relevant thresholds for assigning liability under certain background conditions. Scaling up the interactions can help us see how and why to make these adjustments.

IV. EXTENSIONS AND CONNECTIONS

The analysis above has demonstrated how aggregation choices, often made without reflection or comment, can decisively shape tort

251. *See* Cooter & Porat, *supra* note 7, at 348-50.

252. *See* Grady, *supra* note 12, at 297 (giving the example of a dialysis machine, which reduced the overall risk of death but also increased the opportunities for humans to make liability-generating compliance errors).

253. *See* Shavell, *supra* note 55, at 37-39.

254. *See* Fisher et al., *supra* note 30.

liability and incentives. Although the particulars vary, similar points can be made in other areas of law. In this last Part, I briefly consider connections between the aggregation puzzles examined here and some that arise in other doctrinal areas. These other legal contexts offer useful points of comparison and the potential for transferable lessons in both directions.

A. Setting the Viewfinder

The problems of evaluative aggregation examined here illustrate a deep and important question that runs through all of law: How wide or narrow should the evaluative viewfinder be? Sometimes widening the frame to encompass a larger slice of behavior allows evidence to pile up in support of a conclusion that could not be reached, or could not be reached to a sufficient level of confidence, based on a narrower window of observation. At other times, widening the frame enables other observations to dilute or offset the ones initially isolated in the viewfinder. Conversely, a constriction in the evaluative frame might remove either corroborating or mitigating evidence from view, or do some of each.

The lapse analysis thus shares common ground with questions that arise in other evaluative contexts: Is the observation before us typical or atypical, and could widening the frame provide a definitive answer? The fact that law often determines liability based on what amounts to a single draw from an otherwise opaque urn raises questions about the feasibility and legitimacy of basing decisions on more draws. A core normative question is whether it is appropriate in a given context to, effectively, "tak[e] character evidence" in this manner.[255]

Indeed, to even speak of behavior as a draw from an urn implies that human beings constitute the same "urn" over time, and that observations that form a pattern are not independent of each other.[256] As a legal proposition, this is controversial. To the extent that liability is meant to attach to a given act and not to one's overall urn-pattern, peering into the urn or plucking additional draws may

255. *See* Grady, *supra* note 10, at 402.

256. This issue relates to larger philosophical questions about the durability of personal identity. *See, e.g.*, DEREK PARFIT, REASONS AND PERSONS (1984).

appear illegitimate—both beside the (legal) point and potentially prejudicial.[257] Yet in many contexts, people try to do exactly that, and it is generally regarded as not only sensible but essentially required. For example, it might seem rash to fail to obtain references about a potential hire. Likewise, an employer who suspects an employee of dishonesty might watch that employee closely to see if corroborating or exonerating patterns appear.[258] How we handle lapses in tort law may bear on treatment of these other matters, and vice versa.

B. Bundling and Strategizing

In tort law, as we have seen, the capacity to treat untaken precautions as indivisible bundles may be functional at times. This upends our usual affinity for marginal analysis. By allowing the high social value of omitted *efficient* precautionary steps to absorb the social waste associated with omitted *inefficient* precautionary steps, the latter as well as the former can become a basis for liability when the two appear together in a single proffered untaken precaution.[259] This may seem unconventional, until we recognize that legislative line-drawing, logrolling, and dealmaking of all sorts very typically involve similar moves.

Elements that are inefficient or disadvantageous at the margin are frequently allowed to cannibalize some of the surplus from the elements that are worthwhile. The larger the excess of benefits over costs for the worthwhile elements, the larger the flaws of the unworthwhile elements may be without sinking the overall package. In making an up-down assessment about a policy or program, total costs and total benefits are typically compared—and because total benefits can exceed total costs well beyond the stopping point

257. For an interesting exploration of this issue in a criminal law context, see Sean P. Sullivan, *Probative Inference from Phenomenal Coincidence: Demystifying the Doctrine of Chances*, 14 LAW, PROBABILITY & RISK 27 (2015) (discussing *Rex v. Smith*, 84 L.J.K.B. 2153 (1915), an infamous case involving a defendant's alleged bathtub drowning of his wife, where evidence surfaced that the defendant's two other marriages had also ended with bathtub drownings).

258. *See* Schauer & Zeckhauser, *supra* note 78, at 28 (contrasting the "discrete event" approach of criminal law with employment decisions in which patterns of conduct are often deemed relevant).

259. *See supra* Part II.B.2.

indicated by marginal analysis, a plan that appears valuable overall may include elements that erode rather than augment its value.

Deference to legislative judgments suggests that courts will rarely question the precise position of a line or ask whether it could have been drawn in a different way to advance a greater amount of social value.[260] Is this desirable or undesirable? Recall that the rationale for unifying precautions in the tort context turned on the otherwise asymmetric treatment of errors falling above and below due care.[261] In the legislative or administrative realm, errors may involve doing too much or doing too little. If we think that inaction or insufficient action is systematically punished less severely than excessive action, then there may be an analogous reason to accept bundling that embeds some inefficient increments.[262] We have seen that the best defense against the proffer of a unified untaken precaution is to take due care in the first place, or as near to due care as one can get.[263] Similarly, moving legislation or regulation toward the optimal line offers protection against bundled deals that will trade on the surplus of an untaken reform.[264]

The potential for strategic behavior also exists, however. As new regulations move the baseline forward, this may effectively constrain the size of the next available change that would satisfy

260. *See, e.g.*, Village of Euclid v. Ambler Realty Co., 272 U.S. 365, 388-89 (1926) (explaining that laws are not invalid simply because their prohibitions "include individual cases that may turn out to be innocuous in themselves").

261. *See supra* Part II.B.

262. Of course, actions can move in a deregulatory direction as well as a regulatory one. *See, e.g.*, Jonathan S. Masur & Eric A. Posner, *Cost-Benefit Analysis and the Judicial Role*, 85 U. CHI. L. REV. (forthcoming 2018) (manuscript at 7), https://ssrn.com/abstract=2915063 [https://perma.cc/Y3HB-JRLM].

263. *See supra* Part II.B.

264. Getting as close as possible to optimality will help to immunize a legislative or regulatory agenda both against excessive future moves in the same direction and against repeals that move in the opposite direction. This is because there will be fewer moves in either direction that would qualify as beneficial overall (to the extent that is required under cost-benefit analysis or otherwise). Conversely, the further a legislature or agency overshoots optimality by tacking on worthless elements that erode value, the more opportunities it provides for a later repeal that removes not only the worthless elements but also some of the valuable ones as well. Here, the gains from taking away worthless elements produce surplus that can absorb some deficits produced by taking away valuable ones. Having courts directly impose marginal analysis on both new measures and repeals could police excessive moves in both directions, *see generally* Masur & Posner, *supra* note 262, although this would not address inaction.

a cost-benefit analysis—especially if the steps in which regulators must move are inherently chunky. Thus, incremental moves might in some cases stymie rather than catalyze further moves.[265] By manipulating the unit of analysis—unifying or subdividing moves—the space between optimality on a marginal analysis and acceptability on a cost-benefit analysis can be exploited or eliminated.

C. Consequential Line-Crossings

Aggregation choices can determine whether a particular actor has crossed a legal line. Such choices become more important when cliff-like consequences attach to line-crossing. Consider regulatory takings law, where the government's liability turns on whether it has, in the words of Justice Oliver Wendell Holmes, gone "too far."[266] This inquiry often depends, at least in part, on how much of the owner's property was taken.[267] How should a property owner's holdings be grouped together or broken apart in evaluating whether the government has overreached?[268] And what should the consequences be if the government has, indeed, gone too far?

The first question requires a multifactor analysis, according to the U.S. Supreme Court's recent ruling in *Murr v. Wisconsin*.[269] The doctrinal answer to the second question, however, seems relatively straightforward: once a taking is found, the government must pay for all that it has taken, not just the "too far" increment.[270] This

265. However, incremental changes may also alter political coalitions in ways that increase the likelihood of future changes. *See generally* Saul Levmore, *Interest Groups and the Problem with Incrementalism*, 158 U. PA. L. REV. 815 (2010). The capacity for early reforms to alter the cost-benefit balance of future changes may offer a counterweight to these effects.

266. *See* Pa. Coal Co. v. Mahon, 260 U.S. 393, 415 (1922).

267. Under *Lucas v. South Carolina Coastal Council*, a regulation that eliminates "all economically beneficial use" will be a taking that requires just compensation unless "the proscribed use interests were not part of [the owner's] title to begin with." 505 U.S. 1003, 1027 (1992). Under the multifactor *Penn Central* test that applies in most situations, the question of how much was taken comes up in assessing the economic impact on the owner and the degree of interference with the owner's "distinct investment-backed expectations." Penn Cent. Transp. Co. v. New York City, 438 U.S. 104, 124 (1978).

268. This issue is familiar to property scholars as a "denominator problem" or an issue of "conceptual severance." *See supra* note 36 and accompanying text; *infra* notes 276-77. For further discussion of aggregation and division in the context of regulatory takings, see Lee Anne Fennell, *Lumpy Property*, 160 U. PA. L. REV. 1955, 1972-75 (2012).

269. *See* 137 S. Ct. 1933, 1945-46 (2017).

270. *See* Suitum v. Tahoe Reg'l Planning Agency, 520 U.S. 725, 748 (1997) (Scalia, J.,

generates a formidable cliff effect.[271] The approach is analogous to an HWN rule in the tort realm, to the extent that the causal link between the "too farness" and the harm to the landowner will not be parsed or the effects partitioned.[272] One way to eliminate the cliff of liability would be to effectively "continuize" the government's payment requirement and impose liability for all diminutions of value.[273] Another tack would challenge the idea that the government is fully liable for the landowner's loss once its impositions cross the line and amount to a regulatory taking.[274]

If we accept a cliff effect in the takings context but resist it in the tort context, it is perhaps because we are more conscious of ledge effects in the former setting, and of their pernicious effects on incentives. In the regulatory takings arena, the government enjoys a

---

concurring in part and concurring in the judgment) ("Whereas once there *is* a taking, the Constitution requires just (*i.e.*, full) compensation, a regulatory taking generally does not *occur* so long as the land retains substantial (albeit not its full) value." (citations omitted)). Even if the government chooses to lift the regulation rather than pay for the taking, it is still liable for the time slice during which the regulation was in force. First English Evangelical Lutheran Church v. County of Los Angeles, 482 U.S. 304, 321 (1987).

271. *See Lucas*, 505 U.S. at 1019 n.8 (recognizing the potential for takings doctrine to generate an "all-or-nothing" result in which "the landowner with 95% loss will get nothing, while the landowner with total loss will recover in full"). The complexities of valuing property for purposes of compensation can, however, soften this effect by altering what it means to recover in full. *See* Christopher Serkin, *The Meaning of Value: Assessing Just Compensation for Regulatory Takings*, 99 NW. U. L. REV. 677, 692-94 (2005) (observing that preexisting restrictions on the property are taken into account in determining its fair market value, and that potential unenacted regulations may be considered as well).

272. *See supra* Part II.B.1. Causation can become relevant in some regulatory takings cases. *See, e.g.*, Lemmons v. United States, 496 F.2d 864, 875 (Ct. Cl. 1974) (holding that recovery was unavailable where a permissible permit revocation caused the plaintiff's loss, not the alleged taking of the plaintiff's leasehold).

273. This view is most prominently associated with the work of Richard Epstein. *See, e.g.*, RICHARD A. EPSTEIN, TAKINGS: PRIVATE PROPERTY AND THE POWER OF EMINENT DOMAIN (1985); *see also* KATZ, *supra* note 2, at 145-51 (discussing efforts to "continuize" different areas of law).

274. Some commentators have proposed valuation approaches that would effectively isolate the "too far" portion for compensation. *See* Douglas W. Kmiec, *Regulatory Takings: The Supreme Court Runs out of Gas in* San Diego, 57 IND. L.J. 45, 75 (1982) (advocating a "least economically viable use value standard" for calculating permanent damages because it "satisfies the constitutional minimum by supplying the value increment necessary to cure the regulatory taking"); *see also* John J. Costonis, *"Fair" Compensation and the Accommodation Power: Antidotes for the Taking Impasse in Land Use Controversies*, 75 COLUM. L. REV. 1021 (1975) (recommending a "Reasonable Beneficial Use standard" for measuring compensation under an "accommodation power" theory).

no-liability ledge right up until it has gone too far. If the government had to pay only for the too-far increment, the argument would run, it would err on the side of "too far" every time since it bears all the costs of going less far but internalizes no monetary benefit from its restraint (it pays zero, no matter how close to the "too far" line it may be). Of course, there are some questionable assumptions built into this line of reasoning, including issues of governmental responsiveness to payment obligations and political pressures.[275] But this core concern may nonetheless help explain the doctrine.

Alternatively, or in addition, perhaps there is a normative view that when the government goes too far, this constitutes a unified, indivisible act which cannot, or at least should not be, split into its component "not too far" and "too far" pieces for settling up. The unified untaken precaution approach embodies a similar judgment by treating a defendant's negligent omission and her nonnegligent omission as a unified liability-generating event, at least when the two share space within a cost-justified untaken precaution that would have prevented the accident. Considered side by side, the two contexts suggest the need for more explicit thinking about when certain legal or factual events will be treated as divisible or indivisible.

Regardless of its rationale, the cliff effect in regulatory takings analysis generates great pressure around the "too far" line. This makes matters of aggregation paramount. Thus, courts and commentators struggle with "the denominator problem"[276] and its discredited alter ego, "conceptual severance,"[277] in attempting to define

---

275. *See, e.g.*, Ronit Levine-Schnur & Gideon Parchomovsky, *Is the Government Fiscally Blind? An Empirical Examination of the Effect of the Compensation Requirement on Eminent-Domain Exercises*, 45 J. LEGAL STUD. 437 (2016); Daryl J. Levinson, *Making Government Pay: Markets, Politics, and the Allocation of Constitutional Costs*, 67 U. CHI. L. REV. 345, 414 (2000).

276. The denominator problem takes its name from the implicit fraction-construction exercise that regulatory takings doctrine invites when it asks how severely the owner's property was burdened or whether "all economically beneficial use" was taken. *See* Lucas v. S.C. Coastal Council, 505 U.S. 1003, 1017 (1992). To answer this question, we must define the base against which the diminution of value can be measured.

277. Margaret Jane Radin coined the term "conceptual severance" to denote an illegitimate form of disaggregation that would define the property interest against which the government's incursion was made in terms of the regulatory incursion itself. *See* Margaret Jane Radin, *The Liberal Conception of Property: Cross Currents in the Jurisprudence of Takings*, 88 COLUM. L. REV. 1667, 1674-78 (1988). Conceptual severance has been explicitly rejected by the

the property interest against which the government's acts will be assessed. Widening the frame makes the government action appear comparatively less burdensome and tightening it does the opposite, producing opportunities for strategic behavior on both sides. Interesting questions abound, including what kinds of offsetting effects on the landowner should "count" in pulling the government action back from the "too far" brink.[278] These questions, although seemingly far removed from the aggregative issues that arise in the tort context, share common structural features that merit attention.

CONCLUSION

Tort law's focal point, the accident, is a discrete lump in a stream of risk-related behavior. At the same time, it represents a single instantiation of an interaction that may ultimately be repeated many times. The accident's scale and its singularity do not align especially well with the deterrence tasks that tort law seeks to pursue through the imposition of liability. This fact puts pressure on the aggregation choices that are implicated in the law's thresholding operations—the up-down judgments made at the accident level, which cumulatively shape the overall pattern of tort law.

I have focused here on three sets of aggregative choices: the amount of behavior to examine, the way that precautionary steps

---

Supreme Court. *See* Tahoe-Sierra Pres. Council, Inc. v. Tahoe Reg'l Planning Agency, 535 U.S. 302, 331 (2002) (rejecting "[p]etitioners' 'conceptual severance' argument" as inconsistent with "the parcel as a whole" analysis in *Penn Central* (quoting Penn Cent. Transp. Co. v. New York City, 438 U.S. 104, 130-31 (1978))). To rule out a circular definition of the property interest does not, however, tell us how the relevant denominator should be defined. *See* STEVEN J. EAGLE, REGULATORY TAKINGS § 7-7(b)(2) (5th ed. 2016) (suggesting an opposite risk of "conceptual agglomeration" in takings cases).

278. The transferrable development rights (TDRs) bestowed on the landmarked properties in the *Penn Central* case offer an interesting example of this question. Penn Cent. Transp. Co. v. New York City, 438 U.S. 104 (1978). The majority and dissent disagreed over whether these TDRs should enter the analysis in determining *whether* a taking had occurred (that is, whether they should be allowed to offset the negative impact of landmarking on the landowners) or whether they should be disregarded in that analysis and considered only as a form of compensation that should be assessed for its constitutional adequacy. *See id.* at 149-52 (Rehnquist, J., dissenting). The controversy is an artifact of the cliff effect built into takings doctrine, which raises the possibility that a governmental entity would try to pull back from the brink of a taking by offering something much less than full compensation as an offset. *See* Suitum v. Tahoe Reg'l Planning Agency, 520 U.S. 725, 747-48 (1997) (Scalia, J., concurring in part and concurring in the judgment).

are stacked together, and the conceptual replication of interactions in the analysis of risk. These aggregation choices carry important implications for tort law that have not been sufficiently recognized. Viewing these core theoretical puzzles of tort law through the lens of evaluative aggregation offers new insights on long-standing doctrinal questions as well as on emerging phenomena like autonomous vehicles. The analysis here also sheds light on similar aggregation issues that arise in other areas of law.