CONTINUOUS BURDENS OF PROOF

By Mark Spottswood[*]

--

ABSTRACT

The burden of proof is an essential mechanism to ensure that cases are decided fairly. Our existing form of proof burdens, unfortunately, has a number of surprising downsides. Conventional burdens of proof are mathematically discontinuous. This means that a tiny shift in a jury's confidence in guilt can lead to a dramatic change in consequences for the parties in a case.

This article explores the costs and benefits of an alternative approach to deciding cases, the continuous burden of proof, and concludes that adopting the right kind of continuous burden would help make our justice system fairer, more equitable, and more effective at deterring wrongdoing. Existing scholarship has analyzed one type of continuous burden of proof, the linear burden, which scales the amount of damages in direct proportion to a jury's confidence level that the defendant has engaged in the charged offense. Unfortunately, the choice between linear continuous burdens and traditional burdens involves a hard-to-reconcile tradeoff between the values of deterring wrongdoing and reducing the expected rate of errors at trial.

I offer two ways in which we might move beyond this apparent stalemate. First, I enumerate previously unexplored advantages of continuous burdens, including their ability to minimize the impacts of biases and other forms of structural unfairness at trial. Second, I offer the first exploration of the logistic continuous burden of proof. This novel means of deciding cases strikes a better balance between deterrence and expected errors than either a traditional burden or a linear continuous burden does. After making the case for the adoption of a logistic continuous burden of proof, this article will consider some challenges that would stand in its way and some reasonable steps we could take towards the continuous approach.

[*] Associate Professor, Florida State University College of Law. I am enormously grateful to all those who have offered their thoughts on the ideas developed in this paper, including Ramona Albin, Kelli Alces, Shawn Bayern, Ed Cheng, Avlana Eisenberg, David Kaye, Jay Kesten, Steve Kolbert, Jake Linford, Wayne Logan, Sandy Mayson, Sarah Mirkin, Mike Morley, Emily Murphy, Dale Nance, Mike Pardo, Federico Picinali, Annibal Rosario-Lebron, Mark Seidenfeld, Justin Sevier, George Siedel, Julia Simon-Kerr, Sandy Gurera Thompson, Hannah Wiseman, Sam Wiseman, Maggie Wittlin, and Heather Wolfe.

Electronic copy available at: https://ssrn.com/abstract=3348701

CONTINUOUS BURDENS OF PROOF

TABLE OF CONTENTS

Abstract 1

Table of Contents 2

Introduction 3

I. Choosing Between a Discontinuous Burden and a Linear Burden 7
- A. Deterrence 9
- B. Error Costs 12

II. Considering a Broader Set of Factors 15
- A. Reducing the Average Magnitude of Legal Errors 15
- B. Treating Like Cases Alike 23
- C. Disincentivizing Spoliation? 26

III. Dissolving the Dilemma: The Logistic Burden of Proof 27
- A. Understanding the Logistic Burden 28
- B. The Logistic Burden's Benefits 31

IV. Small Steps Towards a Continuous Burden 40
- A. Implementation Challenges 41
- B. Areas Where Continuous Burdens May Meet Less Resistance 43
- C. Areas Where Some of the Continuous Burden's Benefits May Already Be Seen in Practice 45

Conclusion 48

Electronic copy available at: https://ssrn.com/abstract=3348701

INTRODUCTION

Our legal system implicitly commits itself to the notion that judges and juries can, and do, determine the factual truth in each case that reaches trial. Such an assumption is baked into the system's nomenclature: we label juries as "fact-finders," and use the same term for judges who perform a similar role during bench trials.[1] This assumption is also implicit in the structure of conventional burdens of proof, which assign identical consequences to verdicts across wide ranges of variation in a decision-maker's confidence in guilt. Unfortunately, "fact finding" is a bad metaphor for what juries do. "Fact-finding" suggests that deciding a dispute is like looking for your keys; you go through a process of investigation, which (hopefully) ends with your keys being located. Once your keys are found, you have no further need to worry that you might be *wrong* about their location. If deciding legal cases was a simple enough process that we could truly "find" facts in this sense, we would almost never have to worry that juries have awarded victory to the wrong party.

Sadly, we do not live in such a simple world.[2] A better understanding of what juries do leads to the conclusion that they are engaged in an inherently uncertain *reconstruction* of past events. They may be more or less confident about their conclusions, but they should always acknowledge the possibility that their judgments might be wrong. Our system, however, acknowledges this risk of error in a very crude fashion. If a civil jury is 51% convinced, we dish out identical consequences as if they were 99% convinced, even though the risk of error in the first case is vastly higher. Conversely, we treat cases involving 49% confidence in liability very differently from cases involving 51%, even though the two states of information are much more similar. As we shall see, the artificial assignment of full blame once confidence surpasses 50% yields a number of troubling consequences.

To avoid these problems, we could switch to a very different kind of proof rule: the *continuous burden of proof*.[3] My goal in this article is to survey the existing literature concerning this alternative, suggest some novel reasons why it may be superior to our present practices, and then

---

[1] *See, e.g.* Reeves v. Sanderson Plumbing Prods., 530 U.S. 133, 141-49 (2000) (using the terms "trier of fact" and "factfinder" interchangeably in reference to a trial jury). *See generally* Black's Law Dictionary, entry for "Fact-Finder" (10th ed. 2014, Bryan A. Gardner, ed.) (discussing the use of the term more generally in legal discourse).

[2] *Cf.* Kansas v. Marsh, 548 U.S. 163, 207-08 (2006) (Souter, J., dissenting) (noting that "the period starting in1989 has seen repeated exonerations of convicts under death sentences, in numbers never imagined before the development of DNA tests"); BRANDON L. GARRETT, CONVICTING THE INNOCENT: WHERE CRIMINAL PROSECUTIONS GO WRONG 1-13 (2011) (describing the types of errors that gave rise to wrongful convictions, and noting that "similar errors" in cases that did not involve DNA evidence "may have convicted countless other innocent people and led to the guilty going free").

[3] Some previous articles have examined a particular type of continuous burden, which is linear in form, mostly focusing on the effects such burdens have on accuracy of outcomes and deterrence of wrongdoing. *Cf.* sources cited *infra* at Parts II-A and III. This article, however, builds on these foundational efforts in a number of ways. First, I try to ground our use of discontinuous burdens in the history of our trial process, as a means of deflating intuitions which assume that any rule that is long-standing is likely to be wise. Second, I focus more heavily on the specific dysfunctions that occur due to the *discontinuity* of conventional burdens, where large variations in consequences occur across a narrow variation in fact-finder confidence. *See generally* infra at Part II. Third, I consider the case of continuous burdens more generally, including a novel focus on the benefits of a continuous but non-linear liability function, the logistic burden of proof. *See infra* at Part III.

Electronic copy available at: https://ssrn.com/abstract=3348701

suggest some smaller steps we can take to obtain some of its benefits without radically altering existing practice.

To understand the notion of a continuous burden, it helps to first consider the structure of existing burdens of proof. As commonly understood, a burden of proof[4] instructs the fact-finder to declare that one party's case was true and the other's false once a specified level of confidence has been reached in liability or guilt. The amount of confidence needed will vary depending on the subject matter, and it is usually defined in vague terms. These levels vary from the low threshold of *reasonable suspicion* used for some preliminary inquiries, through the evenly balanced *preponderance* standard used to decide most civil cases, up to the stringent *beyond a reasonable doubt* standard required to obtain a criminal conviction. What is most important to observe for our present purposes is that these standards operate discontinuously, switching from no sanctions to full sanctions the moment that the specified level of confidence has been reached. In other words, if the fact-finder's confidence is just a bit below the threshold, the court will act as if the fact at issue is definitely *false*, at least for the purpose at hand.[5] Once it crosses that threshold, the court will act as if the fact has been conclusively shown to be *true*, and award a full victory to the party who bore the burden. In this respect, our typical burdens operate like a light switch with only two settings – on and off (or in the legal context, guilty or innocent).

A continuous burden of proof, by contrast, operates like a dimmer switch. As confidence increases in the likelihood of guilt or liability, the amount of damages or the length of incarceration increase gradually, starting from nothing and rising to the full measure of responsibility that the law can authorize. For instance, if we apply a continuous burden in a civil case, nothing dramatic happens when a jury's confidence in liability increases from a 49% chance to a 51% chance. Unlike the standard rule, which would suddenly shift from no damages to full damages given this miniscule shift, the continuous rule might increase the damages by only 2% of the full value.[6]

Although this approach might seem strange and unfamiliar to those used to existing burdens of proof, it faithfully represents the fact that 49% confidence and 51% confidence are nearly identical states of information about the world. To be sure, if we are *forced* to chose, it makes sense to chose a threshold of 50% in civil cases, as a means of minimizing our expected rate of

---

[4] For accessibility, I use the phrase "burden of proof" throughout this introduction. More technically minded readers will be aware that the burden of proof can be subdivided into burdens of persuasion (regarding the level at which a jury should be convinced before finding guilt or liability) and burdens of production (referring to the quantity of evidence needed before a court will allow a party to take a case to trial or ask a jury to render a verdict). *See* 2 Handbook of Fed. Evid. § 301:3 (8th ed.). Absent a clear statement to the contrary, my references to burden of proof in main text should be taken to refer only to burdens of persuasion.

[5] The only significant exception to this is judicial review of sufficiency of evidence, which can lead to some verdicts being set aside for being not just wrong, but unreasonable in light of the evidence presented. *See, e.g.*, Fed. R. Civ. P. 50; Fed. R. Crim. P. 29. Relief under these rules represents the exception that proves the rule: within broad bounds of reasonableness, juries are entitled to have their say, and their decisions are treated as unimpeachably correct even if they rest on obviously uncertain grounds such as credibility judgments.

[6] The above example uses a linear burden rule, with penalties increasing by a constant amount in proportion to a fact-finder's credence, in order to simplify the initial explanation of the nature of a continuous burden. In practice, it may be desirable to use a non-linear transformation from credence to consequences. See generally discussion, *infra* at Parts II and III.

Electronic copy available at: https://ssrn.com/abstract=3348701

error in a world of limited information.[7] But in the modern legal context, we have the capacity to fluidly vary many sanctions, so we no longer need to make an all-or-nothing decision.[8] Damages awards, levied fines, and the length of prison sentences can be adjusted upwards or downwards in response to varying levels of confidence, so there is no need to make a discontinuous choice between extreme options.[9]

The existing literature makes it seem that the choice between these two burdens requires us to make a difficult choice with respect to the competing values of optimizing the deterrence of wrongdoing and minimizing the extent to which we punish innocent people.[10] On the one hand, a number of scholars have illustrated the paradoxical effects that conventional burdens can create when acts spread a low risk of harm widely, or where an actor expects that their wrongdoing will be difficult to prove to a high level of confidence in court.[11] Such acts will be severely under-deterred by the existing rule, while a continuous rule will tend to strike a better balance. On the flip side, David Kaye has demonstrated, using the tools of decision theory, that in some simple cases a step-function liability rule will produce lower *expected* error costs than any alternative rule we might adopt.[12] In other words, if we scale punishment in proportion to proof, we will generally take more money from innocent defendants, and less from guilty defendants, as compared with the all-or-nothing rule. Since most of us care deeply about preventing wrongs from happening in the first place and accurately assigning blame when they do occur, this seems to present an irreconcilable stalemate.

I offer two different lines of argument to help us move beyond this impasse, both of which favor a continuous burden of proof over the traditional approach. *First*, I explore some underappreciated benefits of continuous burdens.[13] One additional benefit of continuous burdens is their ability to spread the risk of error across both parties in case. Because we only assign full liability or punishment in cases where we are nearly certain that the charged wrong has been committed, we expect to make far fewer very large errors, such as bankrupting a civil

---

[7] *See* Edward K Cheng & Michael S. Pardo, *Accuracy, optimality, and the preponderance standard*, 14 L. Probability & Risk 192, 194-201 (2015); David Kaye, The Limits of the Preponderance of the Evidence Standard: Justifiably Naked Statistical Evidence and Multiple Causation, 1982 Am. B. Found. Res. J. 487 (1982).
[8] There are, of course, some remaining cases in which a dichotomous choice is truly necessary. One cannot order 0.5 deportations, for instance. *See* Discussion, *infra* at Part V-A.
[9] *See* DENNIS V. LINDLEY, UNDERSTANDING UNCERTAINTY 103-105 (2006) (urging that all-or-nothing verdict decisions are inherently problematic since guilt is always uncertain).
[10] *See infra* at Part I.
[11] *See, e.g.*, Talia Fisher, *Conviction Without Conviction*, 96 Minn. L. Rev. 833, 857-59 (2012) (discussing under-deterrence in criminal cases); Herik Lando, *The Size of the Sanction Should Depend on the Weight of the Evidence*, 1 Rev. L. & Econ. 277, 281-83 (2005) (same); STEVEN SHAVELL, ECONOMIC ANALYSIS OF ACCIDENT LAW 115-17 (1987) (discussing over- and under-deterrence in civil tort cases); David Rosenberg, *The Causal Connection in Mass Exposure Cases: A Public Law Vision of the Tort System*, 97 Harv. L. Rev. 849, 861-67 (1984) (same).
[12] See generally David Kaye, The Limits of the Preponderance of the Evidence Standard: Justifiably Naked Statistical Evidence and Multiple Causation, 1982 Am. B. Found. Res. J. 487 (1982). Kaye focuses on expected error costs, which are determined based on the jury's level of confidence in guilt or liability, rather than the actual costs of error produced by the system, because true error costs depend on the distribution of merit across disputed cases, which is extremely difficult to measure or optimize. See David H. Kaye, Two Theories of the Civil Burden of Persuasion, 2 L. Probability & Risk 9, 10 (2003).
[13] *See infra* at Part II.

Electronic copy available at: https://ssrn.com/abstract=3348701

defendant or executing a criminal defendant.[14] Moreover, although some large errors will still occur, they will be reserved for cases in which guilt or innocence seemed to be nearly certain to the fact-finder, which feels less unjust than imposing a large risk of error on one party in a case that was close to a toss-up.

Continuous burdens also avoid the drastic difference that occurs when a jury is nudged across the line that separates guilt and innocence under a conventional burden. Since parties mostly take hard cases to trial, juries will have to make many close calls. Sadly, there are a number of ways that parties can obtain unfair advantages at trial. Some defendants may fall victim to a racial bias on the part of jury members, while others may find that their opponents can afford much more effective trial counsel. These kind of unfortunate influences are greatly magnified by a system that makes the difference between 51% confidence and 49% confidence the difference between full damages and no award at all. The continuous burden, by contrast, does not eliminate these influences, but it does greatly shrink their impact. Similarly, continuous burdens offer a more speculative benefit, which is their potential to reduce the incentives parties currently have to hide or destroy evidence. This benefit arises through the same mechanism: just as there is less to gain from having a better attorney in a close case, similarly there is less to gain by subtracting one piece of evidence from the jury's consideration.

I also offer a second line of argument, by devising a novel form of continuous proof burdens with particularly beneficial properties.[15] Existing work in this space has focused its attention solely on a linear burden when analyzing the costs and benefits of a continuous burden of proof, but other functional forms may be superior to both. I analyze one interesting alternative: a continuous burden of proof which takes the form of a logistic function,[16] mapping confidence levels onto damages amounts in a smooth but non-linear way. Logistic burdens can be designed with parameters that make them closely approximate either the traditional or the linear burden, but they can also be given a shape that is intermediate between the two extremes. This flexibility allows us to strike a balance between competing goals such as deterrence and reducing error rates, instead of optimizing one value at the expense of getting the worst outcome with respect to the other. Moreover, if we implemented a logistic burden of proof with the right parameters, we should make fewer large errors than are expected to arise under either the traditional rule or the more commonly discussed linear alternative.

---

[14] This idea was briefly explored, in the civil context, but subsequent work on burdens of proof has not developed it more fully. *See* Neil Orloff & Jerry Stedlinger, *A Framework for Evaluating the Preponderance of the Evidence Standard*, 131 U. Pa. L. Rev. 1159, 1165-68 (1968).

[15] *See infra* at Part III.

[16] Logistic burdens take the following functional form, using three scaling constants, r, C1 and C2, to allow the function's shape to be adjusted:

$$Award = \frac{C_1}{1 + e^{(-rp+\frac{r}{2})}} + C_2$$

The result is a sort of flattened S-shape, which rises slowly starting at a jury confidence level of 0, more quickly around 0.5, and then more slowly again as the jury's confidence approaches 1. As I shall discuss in Part III, neither the judge nor jury need manipulate this equation directly in order for us to implement a logistic burden of persuasion; instead, a rule-maker can set the appropriate parameters for the equation and then produce a table that translates confidence levels into appropriate percentages of damages.

Electronic copy available at: https://ssrn.com/abstract=3348701

The remainder of the paper considers issues of implementation.[17] Of course, in some areas existing laws and practices may make it very hard to implement a continuous burden of proof. Most notably, the United States Supreme Court has held that criminal defendants have a constitutional right not to be convicted unless their guilt is proved beyond a reasonable doubt, which would make it difficult to implement a fully continuous burden in criminal cases.[18] Nonetheless, there may be fewer obstacles in other contexts. Many decision-makers have large discretion over their mode of operation, with few constitutional or statutory constraints. Thus, it may be easier to implement continuous burdens in the context of internal corporate fact-finding, agency decision-making or arbitration. Some areas of law, such as tort law and capital sentencing, have already taken steps towards continuous burdens,[19] and the arguments I will develop below suggest that such efforts should be defended and extended. Likewise, some features of the trial process, such as civil settlements, may approximate the outcomes that a continuous burden would produce, so the analysis in this article provides a reason to view the rise of settlements (and the corresponding decline of trials) more favorably than we might otherwise be inclined to do.

I. CHOOSING BETWEEN A DISCONTINUOUS BURDEN AND A LINEAR BURDEN

This section will summarize the main debates that have occurred concerning the utility of either the traditional, discontinuous burden, and the alternative of a linear burden of proof rule. Before we begin, it may help to more clearly define each burden. Burdens of persuasion can be thought of as functions that assign a given fraction of the total possible damages or punishment,[20] based on the level of confidence that a jury has in guilt or liability. To make comparisons easy, it helps to express the burdens graphically, with the amount of damages awarded on the y-axis and the jury's confidence in liability on the x-axis. Using the language of probability theory, we will describe a jury who are certain of innocence at a level of 0. A jury that believes that guilt or innocence are equally likely have a confidence level of 0.5, and a jury that is certain of guilt has a confidence of 1.

Plotted in this fashion, conventional burdens take the form of a step-function, which assigns no liability from a confidence level of 0 up to a level of 0.5, and then assigns full damages from the point just above 0.5 up to 1. Such burdens are *discontinuous* in the mathematical sense, because they leap abruptly from one value to another value, rather than varying smoothly across their entire range.

---

[17] *See infra* at Part IV.

[18] In re Winship, 397 U.S. 358, 364 (1970). *See generally* Discussion, *infra* at Part V-A.

[19] *See* Discussion, *infra* at Part V-B.

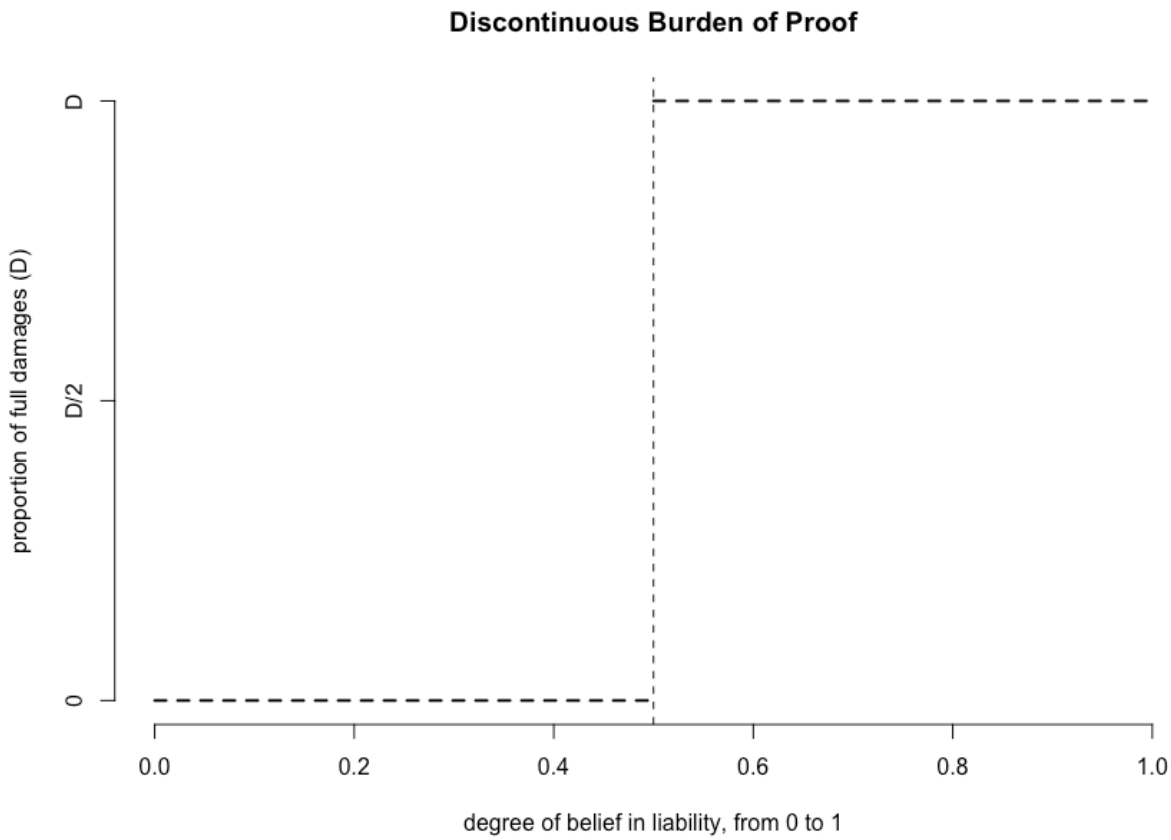
[20] Although judges typically instruct juries that the burden applies to the amount of damages as well, they typically have not articulated a method by which to do this. Damages, of course, can vary smoothly across a wide range of dollar values, so that any particular value may have a negligible likelihood of being correct. To the extent that they consider these complications, judges typically instruct the jury to award the mean of a range of possible damages, even though this may not lead to an award of damages that is more likely than not to be greater than or equal to the actual harm suffered by a plaintiff. For a thoughtful exposition of these issues, *see* Omer Y. Pelled, *All or Nothing or Something: Proportional Liability in Private Law* (working paper, 2020). For present purposes, I will assume (as most analysts do) that the two burdens are being applied to the non-damages elements, relative to a fixed total damages value that the jury will determine using a separate rule.

Electronic copy available at: https://ssrn.com/abstract=3348701

If we label the jury's confidence level in liability or guilt as p and the determined damages in the case as D, the rule can be expressed using the following equations:

$$Award = 0, when\ 0 \leq p \leq 0.5$$
$$Award = D, when\ 0.5 < p\ , \leq 1$$

And when plotted, the rule appears as the dashed line in the following graph, with the "jump" illustrated by the vertical line at the level of 0.5 jury confidence in liability:[21]



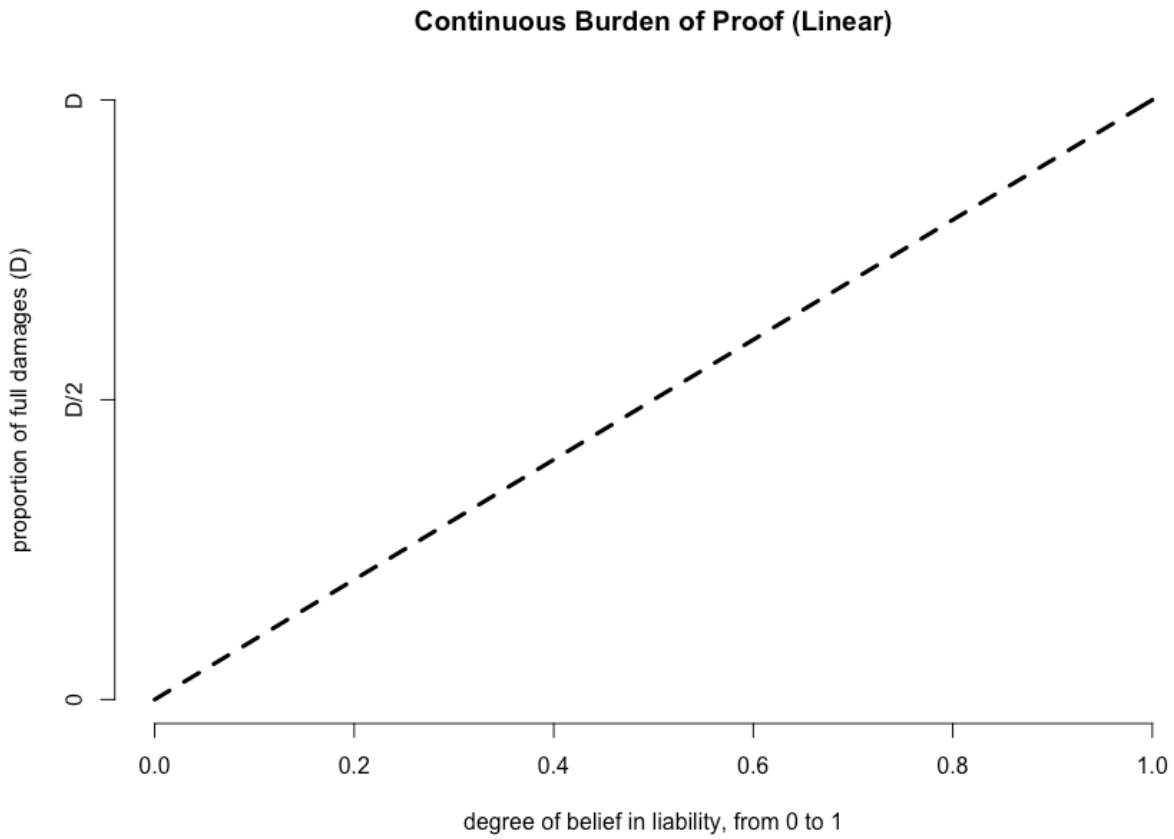
The main foil for this conventional burden in the literature has been a linear burden of proof, which has also been referred to as a proportional liability rule. Such a burden awards damages in proportion to the jury's level of confidence, so that as the confidence goes up, the fraction of damages increases as well. This sort of burden is *continuous*, in that it exhibits no jumps, instead increasing monotonically from 0 to 1. As an equation, it takes the following simple form:

$$Award = pD$$

[21] We could similarly express the criminal burden of proof "beyond a reasonable doubt," subject to uncertainty regarding the level of probability implied by that vague formulation. *See* LINDEY, *supra* note _, at 184; *see also* Discussion, *infra* at note _ (collecting and discussing prior literature concerning the level of probability that matches the criminal burden of persuasion standard).

Electronic copy available at: https://ssrn.com/abstract=3348701

When graphed, the linear burden appears as a simple line with no jumps from one value to another:



With these two kinds of proof burdens clearly in mind, let us now consider the existing debates concerning which is a better fit for our trial process.

A. Deterrence

It is widely agreed that a central benefit of imposing civil damages or criminal punishment is the deterrence of wrongdoing.[22] Unfortunately, scholars and courts have identified a number of ways in which discontinuous burdens of proof fail to optimize this deterrence function.[23] In cases

[22] *See, e.g.*, RICHARD A. POSNER, ECONOMIC ANALYSIS OF LAW 773 (9th ed. 2014); SHAVELL, *supra* note _, at 298.

[23] A recent paper in the law & economics tradition has sought to resolve this problem via step-function burden of proof rules with highly variable thresholds, in order to account for the need to better calibrate deterrence. *See* Louis Kaplow, *Burden of Proof*, 121 Yale L.J. 738 (2012); In addition to having some of the defects of step-function rules *in general*, which are discussed *infra* at subparts b-d of this section, this idea of variable discontinuous thresholds has further problems. In particular, such an approach makes highly unrealistic knowledge demands for the official who is responsible for adjusting the burdens, their unpredictable standards may be unfair to defendants who cannot predict the standards they will face in advance of trial, and they may struggle to obtain public legitimacy. *See generally* Edward K Cheng & Michael S. Pardo, *Accuracy, optimality, and the preponderance standard*, 14 L. Probability & Risk 192, 194-201 (2015). By contrast, the continuous burden does not vary case-to-case, so it is less likely to suffer these problems. Moreover, although one might imagine desiring a great deal of information in order to devise a continuous burden that optimally balances deterrence

Electronic copy available at: https://ssrn.com/abstract=3348701

where an actor can foresee that future cases will likely fail to exceed the proof threshold, the discontinuous rule will generally give little or no deterrence against wrongdoing.[24] Conversely, in cases where guilt is probable but not certain, the dichotomous rule will tend to provide excessive deterrence and thereby discourage economically valuable activity.[25] And these concerns may be magnified to the extent that potential wrongdoers are unduly optimistic risk-seekers who hope to gamble on the prospect of total exoneration, while being relatively insensitive to the potential costs of longer sentences if they are probabilistically uncertain to be imposed.[26]

Consider first the category of cases in which a party might cause some harm while knowing that it will be difficult for a victim to prove a case beyond the applicable threshold of proof. Such cases can take many forms. Perhaps the most discussed in the literature arises in the field of "toxic torts," in which a plaintiff alleges that they have fallen ill due to exposure to one of the defendant's products. In some such cases, uncertainty will arise because similar products are sold under generic branding to consumers by many companies, with long latent periods before any symptoms might arise, leading to great uncertainty about *which* company caused harm to any particular plaintiff. Under such conditions, a seller who knew they sold harmful products to less than half the market might act with impunity, knowing how hard it would be for future plaintiffs to hold them accountable. For this reason, some courts have experimented with probabilistically adjusted recovery of damages in some such cases.[27] But potential problems of this kind arise in other contexts that the courts have not yet addressed. For instance, the uncertainty might arise when a substance raises the likelihood of some illness which can also arise from other sources. So long as the majority of cases of illness arise from other causes, and so long as there are no distinctive markers that show which cases were caused by exposure to the substance, a defendant might once again act with impunity, confident in the knowledge that no individual plaintiff would be able to trace their harm back to the company's product.[28]

This problem is not confined to civil cases involving toxic exposures. Other types of civil cases, such as employment discrimination claims, may require a plaintiff to show that admitted actions were taken with improper intent. So long as there are other plausible reasons for firing someone, a supervisor might act for discriminatory reasons while remaining fairly confident that the plaintiff will be unable to prove the wrongdoing in court, due to the absence of a paper trail indicating the improper motivations.[29] In criminal law, cases of acquaintance rape regularly fail due to the lack of witnesses other than the victim and the defendant. In such cases, it will often

---

and error cost considerations, *see infra* at Part III, one need not perfect such a balance for it to be better than adopting a rule that maximizes one quantity at the other's expense. Moreover, many of the benefits reviewed in the later parts of this section do not require such information to be obtainable.

[24] *See* SHAVELL, *supra* note _, at 115.

[25] *Id*.

[26] Fisher, *supra* note _, at 857-58.

[27] *See* Matsuyama v. Birnbaum, 890 N.E.2d 819, 827–28 (Mass. 2008) (allowing proportional recovery for an illness based on the probability it was caused by an exposure, and surveying other cases taking a similar approach); Sindell v. Abott Labs, 26 Cal 3d 588 (1980 (market share liability)); *see also* Collins v. Eli Lilly & Co., 342 N.W.2d 37 (Wis. 1984); Copeland v. Celotex Corp., 447 So. 2d 908 (Fla. App.1984).

[28] *See, e.g.*, Daubert v. Merrell Dow Pharm., Inc., 43 F.3d 1311, 1321-22 & n.13 (1995) (noting that one must expect "unjust results" in applying the discontinuous burden to cases in which causation of disease must be proved statistically, whenever a substance increases a risk but does not double it).

[29] See Ivan E. Bodensteiner, The Implications of Psychological Research Related to Unconscious Discrimination and Implicit Bias in Proving Intentional Discrimination, 73 Mo. L. Rev. 83, 96 (2008).

Electronic copy available at: https://ssrn.com/abstract=3348701

be possible for the defendant to claim that an encounter was consensual. Even if the jury thinks the victim's story more likely, they may balk at concluding that the defendant's competing account does not create reasonable doubt.[30] Homicide cases may present similar problems when the defendant can argue that he acted in self-defense, or that the victim committed suicide.[31] Given the high standard of proof, criminals may often calculate that they can get away with serious crimes, even though they can be connected to the crime scene by physical evidence or motive, so long as the victim cannot testify or cannot corroborate their account.

Now consider how a continuous burden of proof alters the incentives in all these scenarios. In the toxic exposure cases, the problem is that a predictable percentage of the company's customers are being harmed, but any individual customer who falls ill cannot say that it was probably the company's fault. But with a proportional rule (assuming the other elements are conceded or proven to a very high probability), a company who caused 1 out of 3 of the resulting cases of illness would pay each plaintiff 33% of their damages, which correctly prices the harm the company caused. Thus, the perverse incentive created by a dichotomous burden disappears. Similarly, a company that causes 2 out of 3 cases of illness pays each plaintiff 67% of their damages, rather than 100%, and thus we avoid over-punishing the company. Similarly, a party will be less confident that they can get away with wrongdoing due to a shortage of evidence. The supervisor who dismisses an employee on the basis of race cannot get off scot-free just by suggesting another plausible reason; rather, they must demonstrate that the other reason is substantially more plausible than the improper motivation, or the company will end up paying a significant partial damages award. Likewise, the homicide and rape cases described above are much less likely to result in a full acquittal; rather, the defendants can expect at least some punishment, if their self-serving denials are not enough to convince fact-finders that they are probably innocent (even though they would raise a "reasonable" doubt.

At this point, no doubt, some readers are silently objecting that some of these additional defendants must be innocent. And indeed, an honest assessment of a continuous burden of proof must acknowledge that it will likely produce an increase in error costs. This issue will be considered more fully in Part III, where I will discuss the need for striking an appropriate *trade-off* between deterrence and error, rather than simply optimizing one at the expense of the other. But before we leave deterrence behind, we should examine one more way that a continuous burden can improve our ability to deter wrongdoing.

The above analysis applies most clearly if wrongdoers can easily predict the likely strength of the evidence against them in the future. This may sometimes be the case, but we might also expect some offenders to act carelessly without thinking things through so clearly. Talia Fisher analyzed the expected deterrence benefit of a continuous burden on ordinary criminal offenders in some detail. In her analysis, she noted that many criminal offenders may be risk seeking with respect to the potential possibility of a criminal conviction, which would raise the deterrence cost of

---

[30] *See* Deborah Tuerkheimer, *Incredible Women: Sexual Violence and the Credibility Discount*, 166 U. Pa. L. Rev. 1, 4–5 (2017) (noting the uphill battle that rape victims face in obtaining criminal convictions when their cases hinge on a credibility contest).

[31] *See e.g.*, Jensen v. Clements, 800 F.3d 892, 898, 906-08 (2015) (describing the difficulties in assessing the comparative strength of prosecution and defense evidence in a murder case where the defense claimed the alleged victim died by her own hand).

Electronic copy available at: https://ssrn.com/abstract=3348701

wrongful acquittals.[32] Moreover, such persons may also steeply discount their future utility, giving strong priority to the "here-and-now" without much regard for the consequences for their future self several decades down the line.[33] These speculations find support in the empirical literature on crime and deterrence. At one extreme, innovations focused on raising the probability that offenders will be apprehended and given even a small punishment seem to have large deterrent effects on wrongdoing. By contrast, interventions that focus on increasing the severity of sentences have smaller and less consistent effects.[34]

Thus, although in a rational actor model one can easily trade-off a low probability of achieving convictions by raising sentences in that subset of cases, in the real-world criminals may see little difference between a ten-year sentence and a thirty-year sentence, while being significantly more deterred by smaller sanctions that are more certain to occur. If these assumptions are correct, then the continuous burden yields even more benefits, because it significantly increases the likelihood of *some* punishment, while reducing the typical length of sentences.[35] Thus, both for those who can accurately predict the likely course of their trials, and for those who act recklessly with little regard for the future, the continuous burden of proof may provide a stronger deterrent than our existing approach to proving guilt and liability.

B. Error Costs

Most people do not care only about deterring wrongdoing at all costs; they also care about doing justice in individual cases. In a seminal paper, David Kaye showed that if we measure error costs in terms of the amount of dollars that are either wrongfully awarded to an undeserving plaintiff or withheld from a deserving plaintiff, then we should generally expect a higher quantity of error using a linear continuous burden. Kaye's analysis focuses on a hypothetical civil case where we value an error in favor of either party equally.[36] In its basic form, it addresses a case with a single plaintiff and a single defendant. Under our existing discontinuous liability rule, the defendant

---

[32] Fisher, *supra* note _, at 857-58. *See also* Doron Teichman, *Convicting with Reasonable Doubt*, 93 Notre Dame L. Rev. 757, 776 (2018); Gary S. Becker, *Crime and Punishment: An Economic Approach*, 76 J. Pol. Econ. 169, 207-08 (1968).

[33] Fisher, *supra* note _, at 857-58. *See also* Murat C. Mungan & Jonathan Klick, *Identifying Criminals' Risk Preferences*, 91 Ind. L.J. 791, 806–09 (2016).

[34] *See generally* Aaron Chalfin & Justin McCrary, *Criminal Deterrence: A Review of the Literature*, 55 J. ECON. LIT. 5 (2017); Daniel S. Nagin, *Deterrence in the Twenty-First Century*, 42 CRIME & JUSTICE 199 (2013). Two dramatic extremes help to illustrate this point. On one hand, multiple true experiments have offered high quality evidence that giving automatic jail sentences of 1-2 days for failing to pay fines or violating probation terms strongly deters those kinds of wrongdoing. *See* Nagin, *supra* this note, at 227-28 (collecting studies). Conversely, the literature on the relationship between the death penalty and crime has failed to yield evidence of a deterrent effect, with the highest quality studies suggesting no relationship between capital punishment and rates of offending. *See* Chalfin and McCrary, *supra* this note, at 28-29 (reviewing literature). Given that death sentences are imposed with extreme rarity even in the subset of cases where prosecutors initially seek them, *see* Nagin, *supra* this note, at 218, such findings suggest that potential offenders readily discount low probabilities of even extremely severe sanctions when deciding whether to commit crimes.

[35] These assumptions about defendants' risk preferences and future discounting also imply that prosecutors should be able to encourage more defendants to plead guilty in a world in which we expect more convictions, even if those convictions are generally for shorter sentences. Thus, the deterrent effects of a continuous burden may be magnified from what one might expect based only on the small share of cases that reach trial.

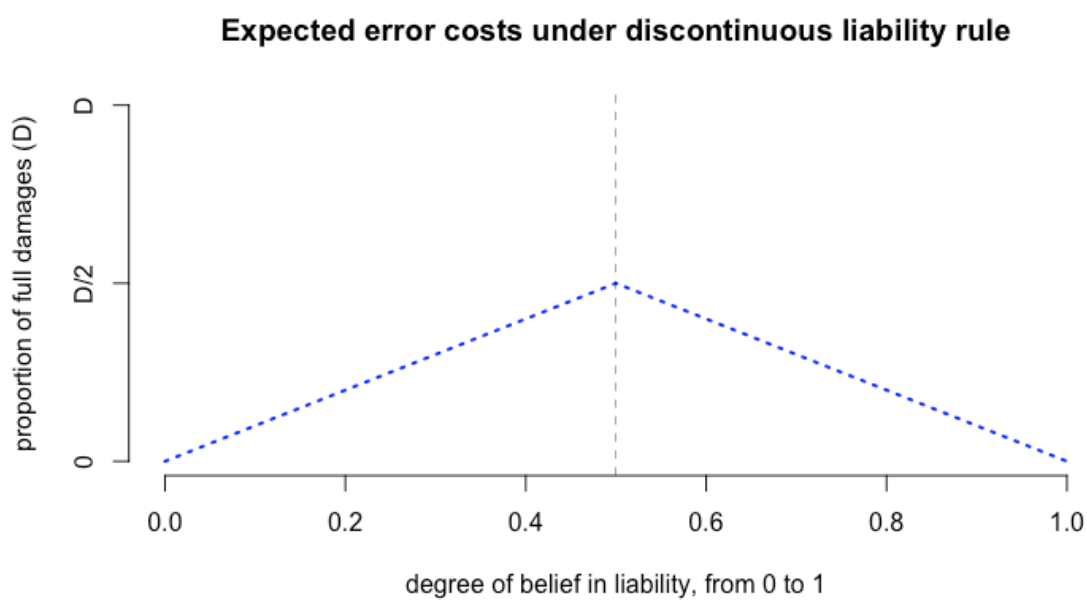
[36] Kaye, *supra* note _, at 496.

Electronic copy available at: https://ssrn.com/abstract=3348701

will either pay full damages *D* or not. If a deserving plaintiff receives no damages, that reflects an error of magnitude *D*, just as it would if a defendant who did nothing wrong pays the damages in full. Now assume that the jury has estimated some probability *p* that the defendant is actually liable. If we use the standard approach to determining liability, then the expected errors (E) are as follows:[37]

$$E = pD, \{0 \leq p \leq 0.5\}$$
$$E = (1-p)D, \{0.5 < p \leq 1\}$$

This arises because, when *p* is less than or equal to 0.5, no damages will be awarded. Since the probability that the plaintiff's claim is actually true in such cases is *p*, then for each case in that range the expected error is *p* multiplied by *D*. For instance, if there is a 25% chance of liability, but we award zero damages, then in 1 out of 4 cases a plaintiff who deserves to receive *D* gets nothing, so we expect an average error of ¼ times *D* in such cases. Alternatively, we can consider the case where *p* is greater than 0.5, which leads to the defendant paying full damages under the current rule. The likelihood that these cases would be decided erroneously is *1-p*, giving an expected error of *1-p* multipled by *D* (the full damages paid by the defendant). Ergo, in the case where there is a 75% chance of liability, 1 in 4 defendants (which is .25, or 1 - .75) are actually innocent of wrongdoing, but they pay *D* anyway, which yields an expected error of ¼ times D again. If we graph this across the full range of *p*, we see the following predicted errors, with the x-axis representing *p* and the y-axis representing expected errors as a percentage of *D*:



Kaye then considered the alternative of assigning liability as a linear function of the estimated likelihood of guilt, which has been recommended by many commentators as a means of ensuring optimal deterrence.[38] For ease of understanding it helps to first consider the expected errors that might be incurred by the parties separately. The plaintiff will be given the total damages, *D*, multiplied by *p*, the estimated likelihood that the defendant is actually liable for the claim. Each

[37] Id. at 498.
[38] *See* Discussion, *supra* at Part II-A, and accompanying footnotes.

Electronic copy available at: https://ssrn.com/abstract=3348701

time these damages are paid, there is a *1-p* chance that the defendant is innocent of the alleged wrongdoing, and the size of the error is the whole payment to the plaintiff, *pD*. Multiplying the probability of error times its size gives the expected error:

$$E_1(p) = (1-p)pD$$

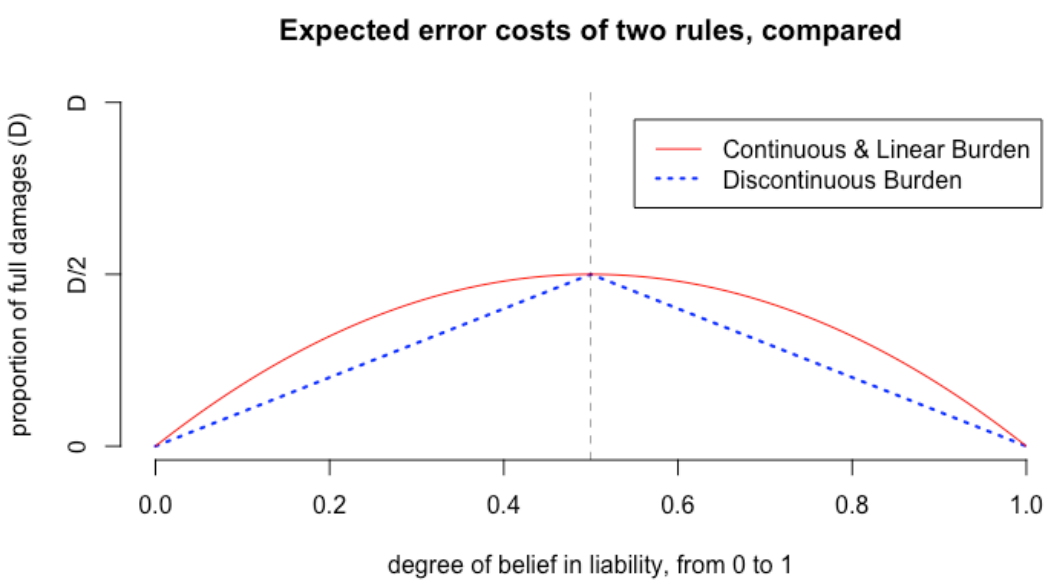
Conversely, each time that *pD* damages are paid, there is a *p* probability that the defendant did commit the alleged misconduct. Therefore, there is also an error favoring the defendant, because ideally the plaintiff should have received the full value of *D* every time they were wronged. The magnitude of such errors is the remaining portion of D that was not paid, *D-pD*, which can be simplified as *(1-p)D*. And because the probability that the defendant is actually liable is simply *p*, the overall expected magnitude of these errors is as follows:

$$E_2(p) = p(1-p)D$$

The overall expected error is simply the sum of the errors that would favor either party:[39]
15

$$E_{total}(p) = 2pD - 2p^2D$$

Plotting this error function alongside the previous one reveals a surprising implication:



As can be seen, the two rules have identical expected errors at three points: where the probability of liability is 0, 0.5, or 1. At all other points, we expect a somewhat greater error from the linear

[39] The above equation can be derived from the individual error terms, as shown below. Alert readers will have noticed that the separate error terms, $E_1(p)$ and $E_2(p)$, are equal to each other, greatly simplifying the analysis:
$E_{total}(p) = E1(p) + E2(p)$
$E_{total}(p) = (1-p)pD + p(1-p)D$
$E_{total}(p) = 2p(1-p)D$
$E_{total}(p) = 2pD - 2p^2D$

Electronic copy available at: https://ssrn.com/abstract=3348701

rule, due to the combined magnitude of smaller errors that affect either party,[40] with an average expected error over the whole range of probabilities of .25 of the damages for the discontinuous rule and 0.33 for the linear rule.[41] Based in part on this analysis, Kaye concludes[42] that, if we wish to minimize expected error costs in civil cases (given the above assumptions), the traditional discontinuous rule will best serve this purpose.

II. CONSIDERING A BROADER SET OF FACTORS

The discussion so far might have us think that the choice of a burden of proof rule is an intractable problem. After all, nearly everyone cares about both deterrence of wrongdoing and achieving justice in individual cases, so it seems very hard to make a choice between the two alternatives. One way out of this dilemma would be to find additional considerations that make the choice a clearer one. In this section, I will survey a few additional policy goals that have received far less attention, but which are inarguably relevant to any policy maker who must choose a burden of proof rule. As we shall see, continuous burdens have some additional advantages that the analysis so far has left out: First, they spread the risk of error more evenly across the parties, and are therefore better at avoiding the imposition of particularly large errors on individual parties. Second, they reduce the impacts of various biases and unfair advantageous in the trial process. Lastly and most speculatively, they may also decrease incentives that parties currently have to hide or destroy evidence.[43]

A. Reducing the Average Magnitude of Legal Errors

One way to reduce the apparent impasse between deterrence and error-minimization arguments is to consider, not just the absolute quantity of error, but the size of errors that are inflicted on each party separately. When we attend to the risk of error that each party will experience, we find that the current rule in fact concentrates its error risk on a single party, while the continuous

---

[40] Kaye, *supra* note _, at 499.

[41] This can be derived as follows. The average error of the discontinuous function is simply half its peak value of 0.5D, or 0.25D. For the average expected error of the linear continuous burden, one must integrate it over the range from 0-1:

$$Average\ error = \int_0^1 (2pD - 2p^2D)dp$$

$$Average\ error = \ [p^2D - \frac{2}{3}p^3D\ ]\ from\ \mathrm{p} = 0\ to\ \mathrm{p} = 1$$

$$Average\ error = \ D - \frac{2}{3}D = \frac{1}{3}D$$

Therefore, the average increase in expected error when moving from the discontinuous to the linear continuous rule is a rise from 0.25 to 0.33 repeating, or an overall increase of 8.33 repeating.

[42] His analysis goes further than just these two rules, and in fact includes a proof that the step function rule reduces errors better than any other means of transforming likelihood judgments into quantified proportions of damages. *See* Kaye, *supra* note _, at 499-500 n.42.

[43] One additional policy criterion that many readers may care about is the impact of either rule on the rate of settlements versus tried cases. This paper cannot include a full analysis of this criterion due to space constraints, but in other work I have shown that continuous burdens modestly increase the incentive of parties to settle rather than try cases, but at the expense of a decrease in the expected accuracy of those settlements. *See generally* Mark Spottswood, *Proof Discontinuities and Civil Settlements* (working paper 2020).

Electronic copy available at: https://ssrn.com/abstract=3348701

burden would spread errors more fairly. This has two implications for our choice between the two burdens discussed so far. First, the absolute size of errors inflicted by the discontinuous burden will be larger, and second, the risk of error that parties experience will be shared less evenly under that rule. As a result, we might find that the error minimization criterion provides weaker support for the existing rule than we might otherwise have assumed.

This point was first noted by Neil Orloff and Jerry Stedlinger, who extended Kaye's analysis by briefly considering how the two rules compared in terms of their ability to avoid inflicting particularly large errors on individual parties.[44] To see the way that the two rules differ in terms of error magnitudes, consider a simple example, in which a jury finds a defendant to be 51% likely to be liable. The existing rule pays the plaintiff her full damages, D. Since by assumption we should expect to be wrong in such cases 25% of the time, the expected error in such a circumstance is .49*D, and in all cases it is fully imposed on the defendant.

Now consider what happens under the linear rule. In such a system, the plaintiff should recover 51% of her damages at this level of proof. There is thus a 49% chance that this payment is wrongful because the defendant is innocent, as well as a 51% chance that this payment is inadequate by 49% of the damages because the defendant is in fact guilty. The expected absolute errors are therefore .49*.51 D (or approximately .25D) against the defendant + .51 * .49D (the same .25D) against the plaintiff. The discontinuous rule, in short, *concentrates* an especially large quantity of expected error entirely on the losing party in close cases, while the linear rule *spreads* the risk of error across both parties.[45]

Kaye's analysis would conclude that since the sum of these expected errors, 0.5D, is greater than .49D, the conventional burden is superior. But for a variety of reasons, we might think spreading errors is the better policy. First, as Orloff and Stedlinger note, large errors are "more jolting to a sense of fairness" than smaller errors. Additionally, people typically do not weigh losses linearly; we are particularly concerned to avoid very large losses, while being more willing to tolerate greater risks of smaller ones.[46] As a result, they suggest an alternative metric that penalizes a system for imposing large risks of losses on individual parties: the squared-error test.

Under the squared error test, we square the expected risk of error to each party before adding them together. Such tests are common statistical device used whenever larger errors seem

---

[44] *See* Neil Orloff & Jerry Stedlinger, *A Framework for Evaluating the Preponderance of the Evidence Standard*, 131 U. Pa. L. Rev. 1159, 1165-68 (1968).

[45] Kaye, *supra* note _, at 501-502. Kaye acknowledges this point, but offers a rejoinder. At least for parties who have frequent contact with the legal system, errors may average out over time. This would *not* be the case if parties face many cases with similar values of *p,* such as where a manufacturer exposes many potential plaintiffs to the same quantified risk of getting a disease. But more often, some cases will be stronger than others, and the variance in *p* may help to balance out the expected risks. *Id.* This observation, however, has limited force. Most people do not have a large number of interactions with the legal system over the course of their lives, and thus they may find that, in a singular encounter or a small number of encounters, they receive the losing end of the tradeoff more than they deserve. Some others, of course, will be overcompensated in such rare encounters, and still others will of course be treated fairly. We might fairly feel that minimizing errors by singling out some individuals to receive a larger share, while insulating others from the same burden, is less than ideally just, even if we will never know *which* individuals are bearing the brunt of this tradeoff.

[46] See id.

Electronic copy available at: https://ssrn.com/abstract=3348701

particularly problematic.[47] In the preceding example, the risk of error under each rule would be analyzed as follows. First, for the standard rule, we expect a squared error of .49*.49 * D * D, or about 24% of the squared damages. Next, for the linear rule, we expect a squared error of .25*.25 *D * D + .25*.25 *D *D = .0625 D^2 + 0625 D^2, for a total of about 13% of the squared damages. Such a test formalizes the intuition that a rule that imposes twice the risk of error on one party and none of the risk on the other party is worse than one that spreads half that risk evenly across both parties.

If we consider both rules in a more general way under the squared error test, we can quickly see that the linear rule has strong advantages in terms of minimizing the risk of larger errors imposed on individual parties. For the conventional rule, we can express the expected squared errors as follows:

$$Sum_of_E^2(p) \;=\; p^2D^2 \;\; when\; 0 \leq \; p \leq \; 0.5\}$$

$$Sum_of_E^2(p) \;=\; (1-p)^2D^2 \;\; when\; \{0.5 \; < \; p \leq \; 1\}$$

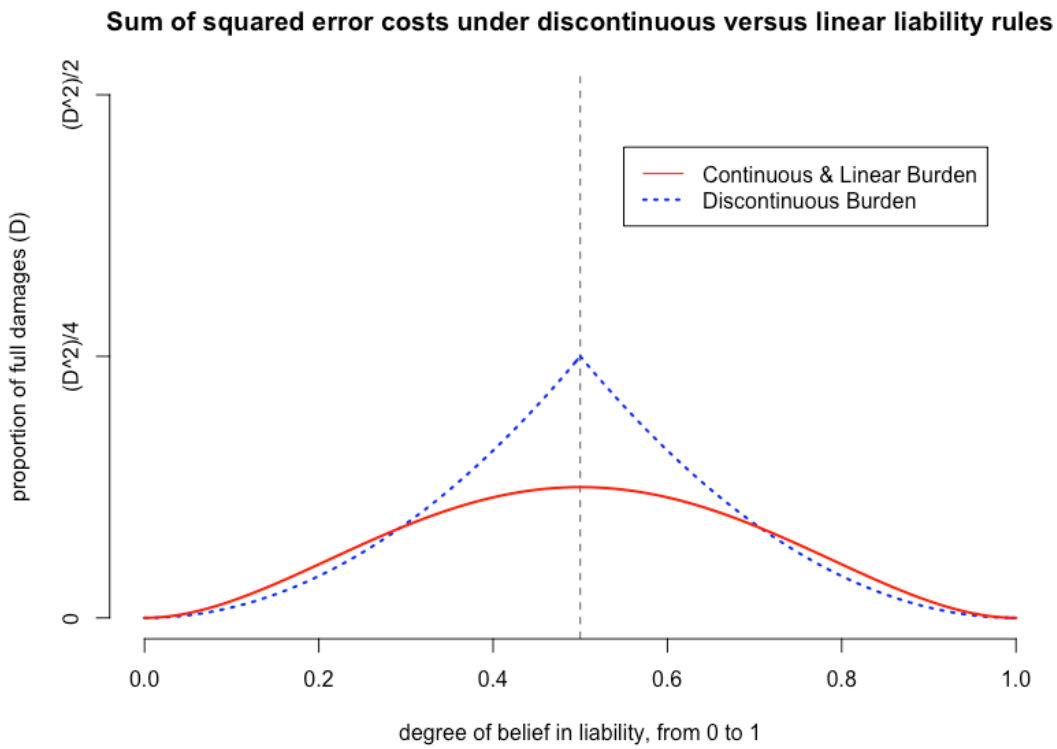
While the linear rule has the following squared-error expression:

$$Sum_of_E\text{^}2\,(p) \;=\; 2p\text{^}2 \;\; 〖(1-p)〗\text{^}2\; D\text{^}2$$

The following graph shows what we should expect if we sum the squared errors of each rule across the whole range of guilt probabilities. As one can see, the discontinuous rule performs quite poorly on the squared-error test, due to its tendency to concentrate all errors on the losing party:

[47] *Id.*; *see also* Joshua Davis, *Expected Value Arbitration*, 57 Okla. L. Rev. 47, 85-94 (2004); John S. Poole, *Improving the Reliability of Management Forecasts*, 14 J. Corp. L. 547, 635 n.182 (1989).

Electronic copy available at: https://ssrn.com/abstract=3348701



The average expected squared error across the whole range is 8.3% of the squared damages for the existing rule, while it is only 6.67% of the squared damages under the linear rule.[48] If we think large errors are particularly problematic, or simply value equality of treatment, this may lead us to think that the cost-of-error consideration weighs less heavily as a reason to favor the existing rule.

And in fact, there are reasons to think that the actual concentration of error under our existing rule is even more severe than the theory above would suggest. In theory, cases may be brought that will result in any level of confidence in guilt or liability. In practice, both party incentives and procedural limits will tend to produce a larger share of cases where the jury's ultimate confidence levels will be close to the proof threshold level, especially under our traditional burden of proof rule. For one thing, most cases where there is very little reason to think the defendant is guilty will probably never be brought in the first place, because plaintiffs and prosecutors usually have little to gain by bringing a sure loser. For another, pre-trial practice typically works to weed

[48] This can be derived by integrating both squared error functions over their range. Since the step-function's squared error term is symmetric around 0.5, this can be derived in the following simplified form.

$$\int_0^{0.5} p^2 d^2 dp + \int_{0.5}^1 (1-p)^2 d^2 dp =$$
$$d^2(\int_0^{0.5} p^2 dp + \int_{0.5}^1 (1-2p+p^2) dp) =$$
$$2d^2((0.5^3/3) = 0.0833d^2$$

While the average squared error term for the linear function is obtained with the following definite integral:

$$\int_0^1 2p^2(1-p)^2 d^2 dp = d^2 \int_0^1 2p^2 - 4p^3 + 2p^4 dp = d^2(2/3 - 1 + 2/5) = .067d^2$$

Electronic copy available at: https://ssrn.com/abstract=3348701

out such cases when they are brought.[49] On the other end of the scale, defendants are likely to seek to settle or plea bargain in most cases where losing at trial seems to be a sure thing, because they too have little to gain by incurring trial costs when the outcome is so clear.[50] As a result, cases actually brought to trial will not be uniformly distributed with respect to likely confidence levels, but should instead be clustered closer to the burden's threshold level.[51]

But as the graphs above have shown, it is precisely in these close cases that the present system will impose the greatest risk of large errors on individual parties. At the extreme, consider civil cases where a jury finds that the two parties' cases are truly in equipoise, so that they have a 50% confidence level in liability. In such cases, the law says that all plaintiffs should lose. But if the jury's confidence level is well-calibrated to the actual likelihoods of liability in these cases, that implies that half of these cases are being decided erroneously, in the sense that a deserving plaintiff is being given nothing, and a defendant who acted wrongfully will not have to account for the harm that they have caused. Of course, because the cases are so close, we have no way of untangling who is in which group. On the upside, half of the plaintiffs and defendants are being treated exactly correctly. But there is a serious downside, in that we have maximized the error in the remaining cases, giving a sizeable number of deserving plaintiffs nothing at all.[52]

Under these circumstances, we might well wonder if it would be fairer to balance the harms, by awarding only 50% of the damages in all these equally divided cases. Now, every case receives some amount of error, in that the undeserving plaintiffs receive a partial windfall from defendants who did nothing wrong. But the upside is that the other group is harmed only half as much. Given that there is no way to avoid causing harm in the aggregate in this situation, we might well conclude that it is better to spread it evenly rather than to concentrate it among a sub-group of cases.[53]

---

[49] *See, e.g.*, Fed. R. Civ. P. 12(b)(6) (permitting judges to dismiss cases that "fail to state a claim on which relief can be granted); Ashcroft v. Iqbal, 556 U.S. 662 (2009) (holding that judges should dismiss cases under this standard if the facts alleged in a complaint fail to make it plausible that the plaintiff will obtain relief on the merits).

[50] There will, of course, be some cases where the prosecutor seeks to make an example of a defendant by imposing a maximum penalty, and so will not accept any plea bargains. These cases, however, will be the exception rather than the rule in a busy criminal justice system.

[51] George L. Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 J. Legal Stud. 1, 3 (1984). *See also* Michael Heise & Martin T. Wells, *Revisiting Eisenberg and Plaintiff Success: State Court Civil Trial and Appellate Outcomes*, 13 J. Empirical Legal Stud. 516, 535 (2016) (finding an average plaintiff trial win rate of 54% across all civil case types, with some variation based on case types); Leandra Lederman, *Which Cases Go to Trial?: An Empirical Study of Predictors of Failure to Settle*, 49 Case W. Res. L. Rev. 315, 323–24 (1999) (noting that factors unrelated to the closeness of the case, such as one party's desire to set a precedent for future cases, may also influence the likelihood that cases will be tried). This distribution might shift a bit under the continuous rule, which produces slightly smaller incentives to settle hard cases (but greater incentives to settle easy ones). *See generally* Mark Spottswood, *Proof Discontinuities and Civil Settlements* (working paper, 2020). And in any event, the disparity in expected squared errors for the conventional rule is so extreme in the middle range that the shift in selection effects would have to be massively biased toward the extremes in order to make the linear rule disadvantages on this metric, and there is no theoretic reason to expect such an extreme shift.

[52] A method of formalizing this intuition, using a sum-of-squared errors test, will be illustrated in the next section.

[53] *See* John E. Coons, *Approaches to Court Imposed Compromise – The Uses of Doubt and Reason*, 58 Nw. U. L. Rev. 750, 756-57 (1963) (raising an early version of this argument, but limiting it to 50/50 cases).

Electronic copy available at: https://ssrn.com/abstract=3348701

Of course, cases involving a *perfect* equality of strength between the parties' cases will be rare. More often, one side or the other will have a small advantage. Still, many scholars have expressed concern that awarding total victory to one party when the evidence is close, but in their favor, can still lead to an outcome we view as unjust. Consider the classic "Gatecrasher Hypothetical," first formulated by L. Jonathan Cohen.[54] In this stylized example, 499 individuals paid for admission to a rodeo performance, but the manager discovered 1000 people sitting in the audience on the night of the show. Moreover, no physical ticket stubs were issued, so there is no way to sort out the paying patrons from the gate-crashers.[55] Assume that this is the sum total of evidence we can amass, and now imagine that the rodeo sues one of the patrons for the price of a ticket. Should it prevail?

The classic approach, under a discontinuous burden, is to say that yes, the probability of illegal entry is .501, so the rodeo should prevail. Of course, the implication of this is that the rodeo can recover the full price of a ticket from all 1000 of the persons seated that night, and thus recover quite a bit more than it is actually owed. Many scholars have seen this outcome as perverse, and have attempted to argue against it, usually by suggesting that there is something inadequate about proof via "naked statistical evidence" or (more fundamentally) the notion of probabilistic judgments of liability.[56] But such scholars must struggle to justify the opposite outcome, which is that the rodeo can recover nothing, because we do know in fact that the rodeo operators were substantially wronged on the night in question. Once again, under the existing rule we see that errors are concentrated in a way we find hard to stomach. Perhaps we allow everyone who can meet the burden of proof to recover, in which case many innocent patrons must pay the rodeo owners a windfall. Or perhaps we simply allow the first 501 customers to sue to recover, in which case the system concentrates its errors against the remaining 499.[57] Or we might reject such proof by "naked statistics" altogether, in which case the enterprise can receive no compensation at all for the torts committed against it. From the standpoint of fairness, none of these answers seem very satisfactory. The continuous burden, by contrast, does a better job: since we cannot distinguish between the patrons, we treat them equally, with some paying a bit more than they should and some paying a bit less. The rodeo, by this method, receives exactly[58]

---

[54] L. JONATHAN COHEN, THE PROBABLE AND THE PROVABLE 74-81 (1977).

[55] Or at least, let us assume this for the purpose of the example. In real life, testimonial evidence would no doubt raise and lower the probability of gate-crashing for particular patrons. Likewise, we must ignore reasons why the rodeo might bring a suit in the real world, such as litigation costs or bad publicity, or the possibility that a victory for the defense might incentivize future rodeo-operators to issue tickets. *Cf.* David A. Kaye, *The Paradox of the Gatecrasher and Other Stories*, 1978 Arizona State Law Journal 101, 106 (suggesting that a victory for the defense might be justifiable on this basis). The point is to isolate certain details in order to probe our intuitions about fairness, not to describe a perfectly realistic scenario.

[56] *Id.* at 75; *see also* Edward K. Cheng, Essay, *Reconceptualizing the Burden of Proof*, 122 Yale L.J. 1254, 1269-71 (2013); ALEX STEIN, FOUNDATIONS OF EVIDENCE LAW 82-83 (2005); Charles R. Nesson, *Reasonable Doubt and Permissive Inferences: The Value of Complexity*, 92 HARV. L. REV. 1187, 1192-94 (1979) (developing the similarly structured "Prison Yard" hypothetical); Tribe, *supra* note _, at 1340-41 (developing the similar "Blue Bus" hypothetical).

[57] Kaye has pointed out that the doctrine of unjust enrichment might be applied to forestall further recovery after the first 501 defendants have paid full damages, which would lead to this result. Kaye, *supra* note _, at 104.

[58] This is, again, assuming that we implement the continuous burden in its strongest, linear form. For reasons I will discuss *infra* in Part III, we might wish to balance this error-spreading function against other considerations, such as error-cost minimization.

Electronic copy available at: https://ssrn.com/abstract=3348701

as much payment as it is owed (since the over-compensation errors balance out the under-compensation errors).

The same considerations apply even more strongly when we consider criminal cases. Most commentators agree that wrongful convictions are a particularly grave type of legal error, given both the direct harm of imprisonment and the collateral consequences of becoming a convicted criminal.[59] What is less widely appreciated is that the risks of wrongful conviction vary across cases.[60] In some cases, we can be exceedingly confident that the accused has committed the crime charged. For instance, imagine a drug transaction in which the purchaser was an undercover officer who also recorded the exchange using a hidden video camera, and which was also witnessed by three bystanders with no apparent biases in favor of either the prosecution or the defense. Such considerations may not remove all *possible* doubt, but the risk of error in such a case is exceedingly remote. Now consider, instead, a case in which the key evidence against the accused consists of eyewitness testimony under challenging viewing conditions,[61] or the testimony of a former co-conspirator who gained substantial benefits from a prosecutor for cooperating.[62] Such situations should lead to lowered jury confidence precisely because we have good reason to think that some such accusations will turn out to be false. A continuous burden of proof can take this concern into account by offering a lowered punishment, and thus mitigate the potential harm of a wrongful conviction in precisely the cases where they are most likely to occur. The discontinuous burden, by contrast, will often impose the full sanction regardless of the increased likelihood of wrongful conviction.

Some readers may naturally worry, at this point, that the continuous burden creates an additional risk of wrongful convictions in cases that cannot meet the "beyond a reasonable doubt" threshold test. This is a legitimate concern – after all, we should expect even *more* errors in cases where the jury thinks guilt is 60% likely, for instance, than in the cases that barely satisfy the existing threshold. Imagine, for instance, a set of 10 defendants, each convicted of a crime that carries a maximum penalty of 10 years, based on a finding that guilt is 50% likely. It might seem cold comfort indeed to subject all 10 defendants to a 5-year prison term, knowing that half of them had been wrongfully convicted! But here we must clarify an aspect of the continuous burden that can easily be neglected for the sake of brevity at other points in this discussion: Just because a burden is *continuous*, does not imply that its range of possible sanctions must scale *linearly* with the jury's confidence in guilt. There are many possible ways we link confidence to guilt, and we can adjust the scale to reflect precisely the kind of concerns described above.

Thus, imagine that we believed that the appropriate punishment for a serious assault was a five-year prison sentence, assuming we knew with near-perfect confidence that the crime had truly been committed by the defendant. If we wished to avoid the under-deterrence and inequality of

---

[59] *See, e.g.*, DAN SIMON, IN DOUBT: THE PSYCHOLOGY OF THE CRIMINAL JUSTICE PROCESS 13 (2012); D. Michael Risinger, *Innocents Convicted: An Empirically Justified Factual Wrongful Conviction Rate*, 97 J. Crim. L. & Criminology 761, 789 (2007).

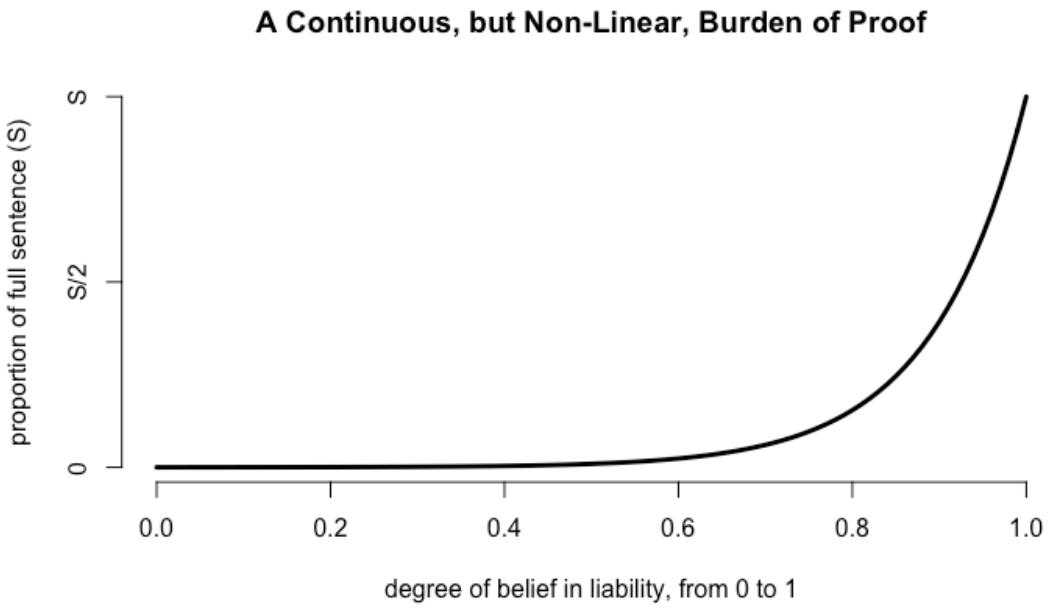
[60] *Cf.* Risinger, *supra* note _, at 783-88 (exploring some of the factors that might give rise to variance in the likelihood of wrongful convictions across different types of criminal cases).

[61] *See generally* SIMON, *supra* note _, at 50-89 (surveying weaknesses in eyewitness identification testimony).

[62] Eli Paul Mazur, *Rational Expectations of Leniency: Implicit Plea Agreements and the Prosecutor's Role As A Minister of Justice*, 51 Duke L.J. 1333, 1346 (2002) (discussing the serious reasons to doubt the reliability of accusations by former accomplices who testify in connection with generous plea agreements).

Electronic copy available at: https://ssrn.com/abstract=3348701

outcome considerations described above, we might choose a punishment scale that gives declining prison terms down to a 75% likelihood of guilt, which is towards the lower end of what judges or lay jurors might associate with proof beyond a reasonable doubt.[63] Below that level, but above the 50% level, we might assess a smoothly declining set of fines or jail stays and define the resulting offense as a misdemeanor to reduce its collateral consequences. Below the 50% level, the consequences might be even smaller fines, defined as civil penalties to further reduce the collateral consequences, but with the potential to increase future sanctions in the event of recidivism. And of course, we could implement such a system without the jury being required to understand the underlying mathematics, by having them report their confidence level, with the court then determining the amount of punishment by following a pre-determined schedule.



[63] There is some disagreement on exactly how one might quantify the notion of proof beyond a reasonable doubt, but commonly given figures tend to center in a range from 80 to 90% confidence. *Compare* Vargas v. Keane, 86 F.3d 1273, 1281-84 (1996) (Weinstein, J., concurring) (reporting a small survey given to sitting federal jurors, in which they were asked to quantify varying formulations of the BRD standard, and gave averaged answers ranging from 79-94%, depending on the formulation, with wide levels of individual variation among jurors, even when evaluating the same formulations); *and* United States v. Fatico, 458 F.Supp. 388, 410 (E.D.N.Y. 1978) (reporting the results of a survey of judges, in which estimates of the strength of the standard ranged from 76% to 95%). *with* C.M.A. McCauliff, *Burdens of Proof: Degrees of Belief, Quanta of Evidence, or Constitutional Guarantees*, 35 Vand. L. Rev. 1293, 1325 (1982) (reporting the results of a similar survey sent to judges, in which a majority clustered in the range given above, but in which there were a significant number of judges who felt that the standard required 98, 99, or even 100% certainty). Laypeople commonly give answers clustered in the 80-90% range when asked what the standard should be, but may use lower implicit thresholds when asked to recommend how a case should be resolved. *See generally* Svein Magnussen et al., *The Probability of Guilt in Criminal Cases: Are People Aware of Being 'Beyond Reasonable Doubt'*, 2013 Applied Cognitive Psychol. 1, 4-7 (reviewing literature and reporting additional experimental evidence that the standard in action is lower than the standard as stated by judges, with participants convicting at a 0.6 level when acting individually and at a 0.7 level following deliberations); David H. Kaye et al., *Statistics in the Jury Box: How Jurors Respond to Mitochondrial DNA Match Probabilities*, 4 J. Empirical Legal Stud. 797, 823 (2007) (finding that jurors would likely vote to convict based on pre-deliberation probabilities of .68, and post-deliberation probabilities of 0.76).

Electronic copy available at: https://ssrn.com/abstract=3348701

A straightforward mathematical way to devise such smooth gradations will be explored in more detail in Section III, below. What matters for the present discussion is that a continuous rule need not be linear in order to give some deterrent effect against harder to prove offenses nor to reduce the number of parties who experience the largest legal errors. And to the extent we think that wrongful convictions are especially costly, we can in fact design the scale to decay more rapidly as confidence declines, so that lower probability offenders receive *some* consequences, but not the severe ones that make wrongful convictions especially troubling.

B. Treating Like Cases Alike

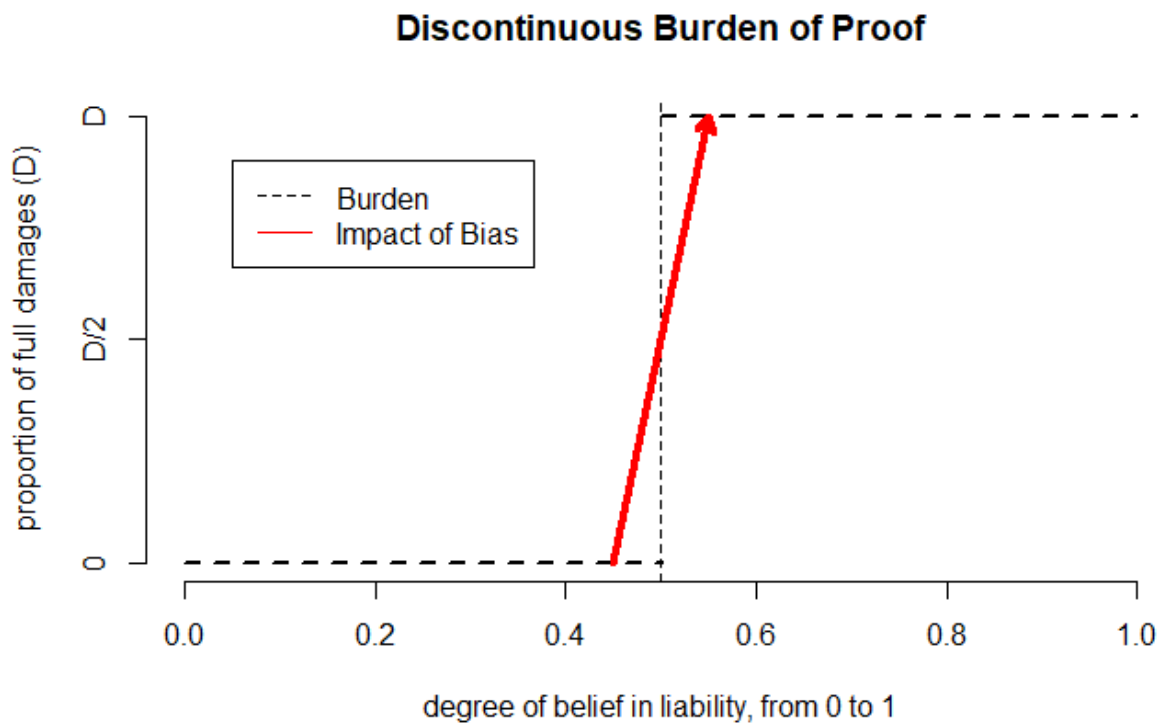
In addition to the error-spreading considerations we have just considered, a continuous burden may be preferable based on its ability to give more equal treatment to those who come before our courts. An under-appreciated cost of the discontinuous burden is that, in cases where a typical juror's confidence in guilt or liability will likely be close to the burden's decision threshold, small and irrelevant factors may tip the balance between one side and another. The factors which might play a dispositive role are numerous, and many will make us uncomfortable: wealth (in part through its ability to purchase more effective advocacy), subtle preferences for in-group members over social out-groups, halo effects, and evidence with excessive emotional weight may all tend to nudge jury members across the line from "not guilty" to "guilty." The continuous burden cannot take such effects away, of course, but as we shall see, it drastically reduces their impact, leading to minor variances in damages amounts or sentence lengths, rather than the difference between winning and losing everything that is at stake in a case.

Many worry that disparate wealth may lead to unfair outcomes in cases that go to trial. [64] Not all attorneys have equal skill, and better funded advocates can invest more heavily in investigation and trial preparation.[65] Of course, many cases have relatively clear outcomes, and even the best attorneys can only cloud the waters so much.[66] Ergo, we should expect advocacy quality (and therefore wealth) to make a significant difference only in cases where a typical jury would find the case difficult to decide because their confidence levels lay close to the existing proof threshold. This can be visualized on the chart below, which shows the current, discontinuous burden of proof in civil trials:

---

[64] *See, e.g.*, Albert Yoon, *The Importance of Litigant Wealth*, 59 DePaul L. Rev. 649, 660-69 (2010); Stanton Wheeler et al., *Do the "Haves" Come Out Ahead? Winning and Losing in State Supreme Courts, 1870-1970*, 21 Law & Soc'y Rev. 403, 408–09 (1987); Marc Galanter, *Why the "Haves" Come Out Ahead: Speculations on the Limits of Legal Change*, 9 Law & Soc'y Rev. 95, 103-04 (1974).
[65] Yoon, *supra* note _, at 660-69; Wheeler et al., *supra* note _, at 408-09.
[66] *See* Wheeler et al., *supra* note _, at 440-41 (finding that having access to stronger legal counsel provided a small but measurable advantage in State Supreme Court litigation).

Electronic copy available at: https://ssrn.com/abstract=3348701

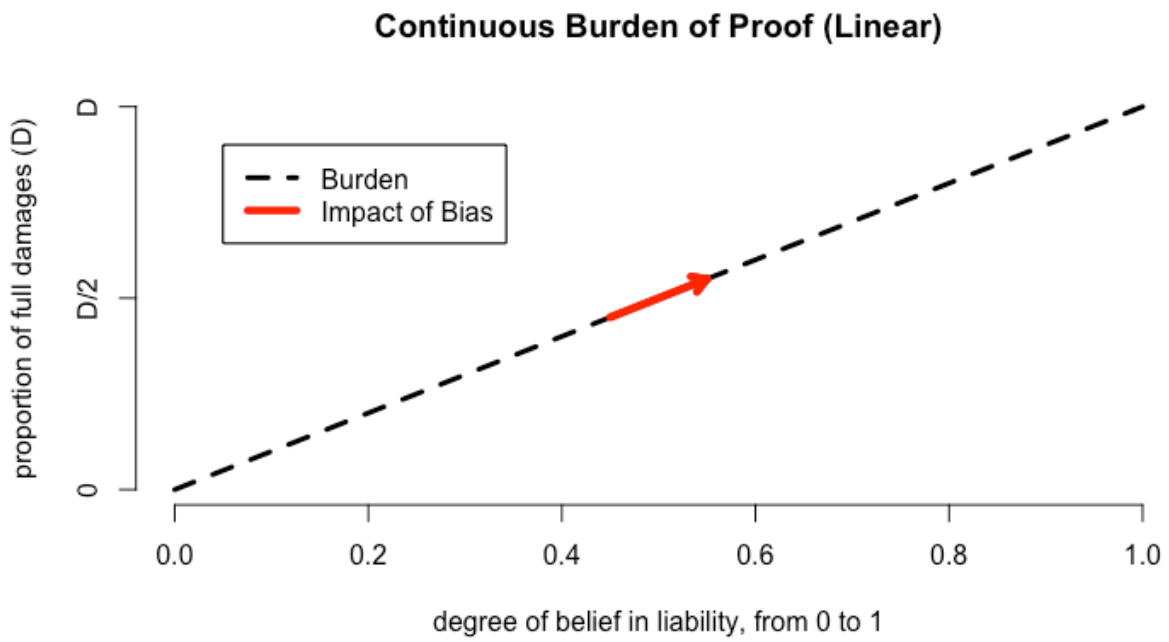


The red line shows what happens when a disparity in advocacy quality leads a jury to find liability 55% likely, when they would have found it only 45% likely given equal quality of advocacy. Given the discontinuous burden of proof, this minor change in confidence radically transforms the defendant's prospects, as she is now responsible for 100% of the damages. But if the same attorney was proceeding in a case where the jury would otherwise see liability as only 25% likely, the same level of skill will provide no practical benefit, because a defendant who is only 35% likely to be liable would still prevail under the existing rule. Thus, discontinuous burdens of proof can magnify small advantages, but this effect will primarily occur in cases close to their decision thresholds.

Now notice how differently the same effects would play out in a continuous regime. Moving the jury from 45% confidence to 55% confidence would not radically change the outcome. Rather, it might adjust the damages by as little as 10%.[67] Once again, this is easiest to visualize by plotting a simple version of a continuous burden of proof, in which the quantity of damages increase as a linear function of a jury's belief in liability, with a red line showing the impact of a small change in confidence in liability:

67 This example assumes that damages are computed as a linear function of the estimated likelihood of the plaintiff's claim (l) multiplied by the plaintiff's total compensable harm (h), such that D(l) = l*h. Such an approach would spread the impact of small variances in persuasion evenly across the whole range of estimated likelihoods, and thus would minimize the impacts of biased reasoning and other unfair advantages on the outcome of cases that reach trial. If we choose instead a different liability function as a means of balancing this with other considerations such as error costs, see infra at Part III, these effects would no longer be minimized, but they would still be much reduced from their impact under the currently dominant step-function liability rule.

Electronic copy available at: https://ssrn.com/abstract=3348701



Of course, the flip-side of this change is that the same 10% variance can occur wherever the jury's confidence levels are, so a greater number of cases could be impacted by advocacy effects. But I suspect that many readers may be less worried about wealth causing a 10% variance in sentence length or damages quantity, compared with a case where wealth enables guilty parties to escape consequences altogether.

Nor are differences in advocacy quality the only means by which a small advantage might lead to outsized harm by way of a discontinuous burden of proof. The tendency of jurors to trust members of in-groups slightly more than members of out-groups may likewise tip the scales in an otherwise close case.[68] It is important to note that such effects can occur even if the members of the jury would not consciously endorse a biased conclusion against a person or group. Rather, unconscious attitudes might shape the willingness of jury members to believe that a particular person would be likely to engage in wrongful conduct, making it easier to show that outgroup members are guilty or liable. Given the typical demographics of jury members, this may make it a bit harder for members of racial minorities, recent immigrants, or economically disadvantaged individuals to prevail in court, compared with white, native-born, and middle-class individuals. Once again, we should expect these tendencies to be modest compared with the impact of the strength of the evidence, but the threshold approach to proof can lead to dramatic disparities in close cases.

[68] *See* Mark Spottswood, *Live Hearings and Paper Trials,* 38 Fla. St. U.L. Rev. 827, 847-48 (2011) (reviewing studies relating to biases against out-group members).

Electronic copy available at: https://ssrn.com/abstract=3348701

This catalog could go further, encompassing such sources of error as halo effects[69] or emotionally biasing evidence,[70] but many readers may feel that this is belaboring the point. In an adversarial system, we normally expect to strike a rough balance across many potentially biasing features of a case, so that the strength of the evidence on either side can be the primary driver of legal results. Unfortunately, the discontinuous burden of proof makes small variances in confidence extremely consequential in close cases, even when they are produced by irrelevant or invidious factors. The continuous burden, by contrast, mitigates the damage that such factors can cause.

C. Disincentivizing Spoliation?

The advantages discussed in the preceding sections follow automatically from the nature of a continuous burden, along with a few other facts in which we can have high confidence. This section considers one other type of benefit that is more speculative, in that it depends on litigation dynamics that can be somewhat hard to predict. Still, there is reason to suppose that the existing structure of proof burdens magnifies the benefits that may flow from hiding or destroying evidence. The continuous burden, by contrast, would seem to give parties' a lessened incentive to destroy documents or silence witnesses who might testify against them.

Consider first the incentives that apply to spoliation under current law. On the one hand, destroying evidence is a high-risk strategy. If caught, parties might be subject to potentially damaging sanctions from a court,[71] tort liability,[72] or prosecution for obstruction of justice.[73] Additionally, a jury that is told about a party's attempts to hide the truth may infer the worst regarding the content of the evidence that is no longer available.[74] Unfortunately, there are strong incentives to spoliate that sometimes overcome these reasons for caution.[75] After all, if a

---

[69] Psychological "halo effects" cause us to unconsciously assume that persons with some positive qualities (like beauty or high social status) are more likely to carry other unrelated positive qualities, such as diligence or truthfulness. *Id.* at 845-47 (reviewing the literature on halo effects).

[70] See, e.g., Rebecca Hofstein Grady et al., Impact of Gruesome Photographic Evidence on Legal Decisions: A Meta-Analysis, 25 Psychiatry, Psychology & Law 503 (2018).

[71] *See, e.g.*, Silvestri v. Gen. Motors Corp., 271 F.3d 583, 593 (4th Cir. 2001) (noting that the district courts have the inherent power "to fashion an appropriate sanction for conduct which abuses the judicial process," and in some cases might properly go so far as to dismiss a claim entirely based on the willful destruction of material evidence).

[72] *See generally* Larison v. City of Trenton, 180 F.R.D. 261, 264 (D.N.J. 1998) (surveying states that have recognized tort liability for the intentional or negligent destruction of evidence).

[73] *See, e.g.,* United States v. Jahedi, 681 F. Supp. 2d 430, 439 (S.D.N.Y. 2009) (noting that courts have "consistently" held that the intentional destruction of documents with the intent to obstruct federal judicial proceedings is a violation of 18 U.S.C. § 1503).

[74] *See, e.g.*, Stevenson v. Union Pac. R. Co., 354 F.3d 739, 746 (8th Cir. 2004) (noting that adverse inference instructions are appropriate when a party intentionally spoliates inference); Finally, Part IV demonstrates how giving the instruction for less than bad faith spoliation *685 does more than provide a remedy and instead gives the nonspoliating party an unfair advantage. *Cf. Wm. Grayson Lambert, Keeping the Inference in the Adverse Inference Instruction: Ensuring the Instruction Is an Effective Sanction in Electronic Discovery Cases,* 64 S.C. L. Rev. 681, 681–85 (2013) (noting the varying standards that exist in the federal circuit courts for granting adverse inference instructions as a spoliation sanction).

[75] *See also* Dan H. Willoughby, Jr. et. al., *Sanctions for E-Discovery Violations: By the Numbers*, 60 Duke L.J. 789, 803 (2010) (noting that the failure to preserve electronically stored information is the single most common basis for discovery sanctions). The problem is almost certainly quite a bit worse than it would appear based on counting decisions in which courts punish spoliators, however. If sanctions were certain, the rewards of spoliation would

Electronic copy available at: https://ssrn.com/abstract=3348701

party can tip a case from one side of the proof threshold to the other, they may walk away unpunished. So if litigants anticipate losing under the existing pattern of evidence, think that they have a reasonable shot at prevailing once some evidence is destroyed, and believe they can do this without being discovered, they may think that the risk is worth it.

In a continuous regime, by contrast, a party is unlikely to achieve total victory by removing part of the evidence. After all, it is quite rare that the entire case against a party will arise from a single witness or document. Rather, a jury must infer guilt from many overlapping sources. Thus, in most cases, removing an important piece of evidence would not lead the jury to conclude that the defendant is probably innocent, even if it would make the case thin enough to raise a reasonable doubt about guilt. This greatly alters the risk reward trade-off. As discussed above, many defendants might discount their future utility rather steeply, and thus be relatively insensitive to changes in sentence length, even if they are highly motivated to obtain a not guilty verdict at trial.[76] Thus, they may find the risk of spoliation is no longer worthwhile when it involves trading off a risk of further prosecution or an adverse inference, on the one hand, against a mere sentence reduction if they are successful.

Of course, as before there is a trade-off involved in moving to the continuous regime. Parties who have very high confidence that they will win or lose a case would have little incentive to spoliate under existing rules, because they would not foresee that it would change the outcome in the case. But in a continuous regime, such conduct could sometimes be worth the risk, simply because it would change the jury's confidence level in a verdict and thus the level of damages or punishment imposed. Nonetheless, the amount of any such increase is probably small. After all, parties could still have incentives to spoliate evidence in easy cases under existing law, as a means of strengthening their negotiating position when hammering out a settlement or plea-bargain. The amount of *extra* inducement offered by the possibility of slightly changing a continuously variable trial verdict, compared with the already existing incentive to influence these bargaining outcomes, seems modest, given that trials are rare in comparison to settlements and plea bargains. Nonetheless, which of these two effects would dominate is ultimately an empirical question, which is why I have emphasized that the above analysis is tentative. As a result, we have reasonable grounds to think that a continuous regime would deter more spoliation than it would incentivize, even if this cannot be stated with certainty.

III. DISSOLVING THE DILEMMA: THE LOGISTIC BURDEN OF PROOF

Up until now, we have assumed (as do nearly all writers in the existing literature)[77] that the choice of burden is dichotomous: either we stick with the present rule or we use a linear burden of proof rule. Such a dichotomy, if it really existed, would be unfortunate. Most people think all of the policies that we have discussed above are valuable. Clearly, reducing errors is valuable,

---

be minimal, and few parties would ever risk it. Thus, we may reasonably infer that a significant fraction of the parties who hide or destroy evidence go unpunished and that the resulting errors go uncorrected.

[76] *See* Fisher, *supra* note _, at 857-58.

[77] The main counterexamples consist of occasional considerations of a discontinuous rule with an intermediate step. *See, e.g.*, Federico Picinali, *Do Theories of Punishment Necessarily Deliver a Binary System of Verdicts? An Exploratory Essay*, 12 CRIM. L. & PHIL. 555 (2018) (exploring arguments for and against a "many-valued" system, in which there are intermediate verdict levels between guilty and not-guilty.

Electronic copy available at: https://ssrn.com/abstract=3348701

and so is deterrence, equality of treatment, and the preservation of evidence. Sadly, some of these goals are at odds with one another. As we have seen, a linear burden gives us better deterrence of wrongdoing, reduces the risk that large errors will be imposed on individual parties, reduces the impacts of various forms of bias at trial, and may also reduce the risk of spoliation. But the rule also increases our expected overall error rate from 25% to 33% of the overall damages in civil cases. We might well worry that the benefits that we have discussed are not worth a price as steep as this. An ideal rule would be one that allows us to strike a reasonable balance between these competing considerations, rather than maximize some at the expense of others.

Happily, such a rule does exist. This section will first lay out an alternative form of continuous burden of persuasion, the logistic burden of proof. As we shall see, the logistic burden is flexible in ways that the other two standards are not. It is able to take many forms that are intermediate between the two extremes, and thereby enable us to strike the kind of balance discussed above. Moreover, on one measure (minimization of large losses) it actually outperforms both of the discontinuous and the linear burden. Thus, there are strong reasons to think that the ideal form of proof burden is logistic.

A. Understanding the Logistic Burden

Logistic functions let one construct a curve that takes a kind of flattened "S" shape, first rising slowly, then more quickly, then more slowly again. In its most basic form, the logistic function is defined by the following equation:

$$f(x) = \frac{1}{1 + e^{-x}}$$

The addition of constants allows us to define a logistic curve that lies neatly between the boundaries of a discontinuous step function, on one hand, and a continuous linear function, on the other.[78] Moreover, an intelligent choice of parameters will allow us to define logistic functions that rise more steeply or slowly, centered around any inflection point we like. The following examples will make use of the following modification of the basic logistic equation, with p once again representing the jury's confidence level with respect to guilt or liability. In addition, we add three constants: r, which varies the steepness with which the curve inflects towards its center, and C_1 and C_2, which are used to position it so that it observes certain basic features that are desirable in a civil burden of proof, such as awarding 0 damages at 0 confidence in guilt, 50% of the damages at 50% confidence, and full damages at 100% confidence.[79]

---

[79] If we wish to tune the function to observe the constraints derived in text, the values of C_1 and C_2 can be derived as follows:

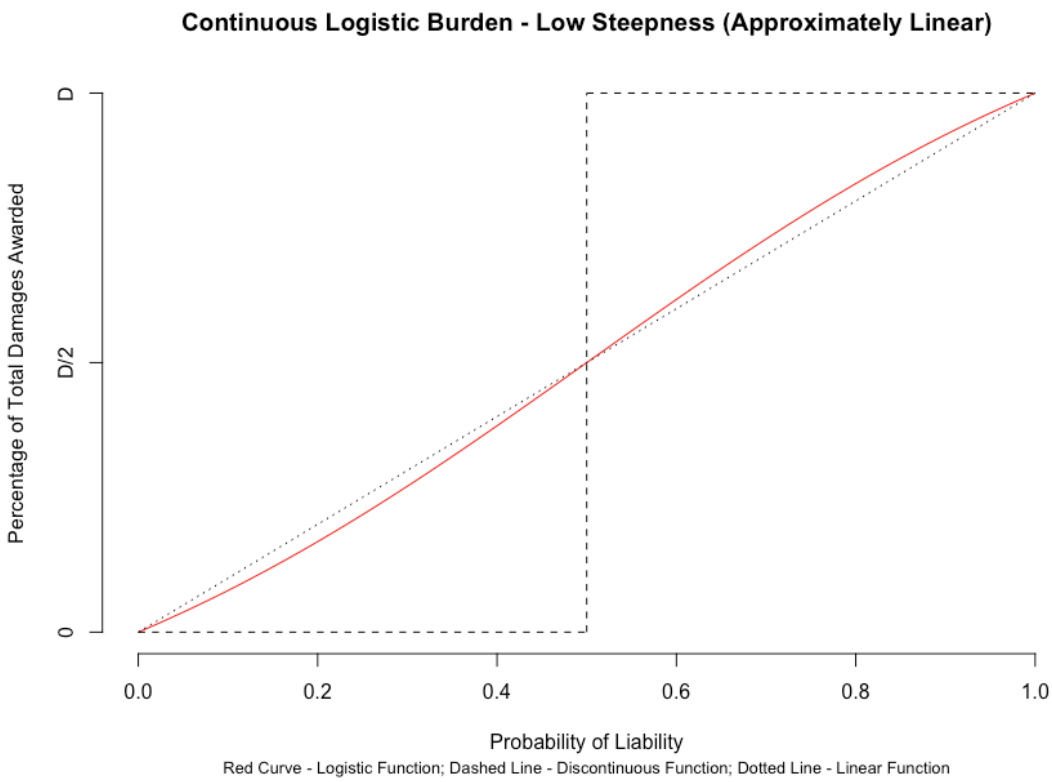
$$C_1 = \frac{1}{\frac{1}{1 + e^{-\frac{r}{2}}} - \frac{1}{1 + e^{\frac{r}{2}}}}$$

$$C_2 = -\frac{C_1}{1 + e^{\frac{r}{2}}}$$

Electronic copy available at: https://ssrn.com/abstract=3348701

$$Award = \frac{C_1}{1 + e^{(-rp+\frac{r}{2})}} + C_2$$

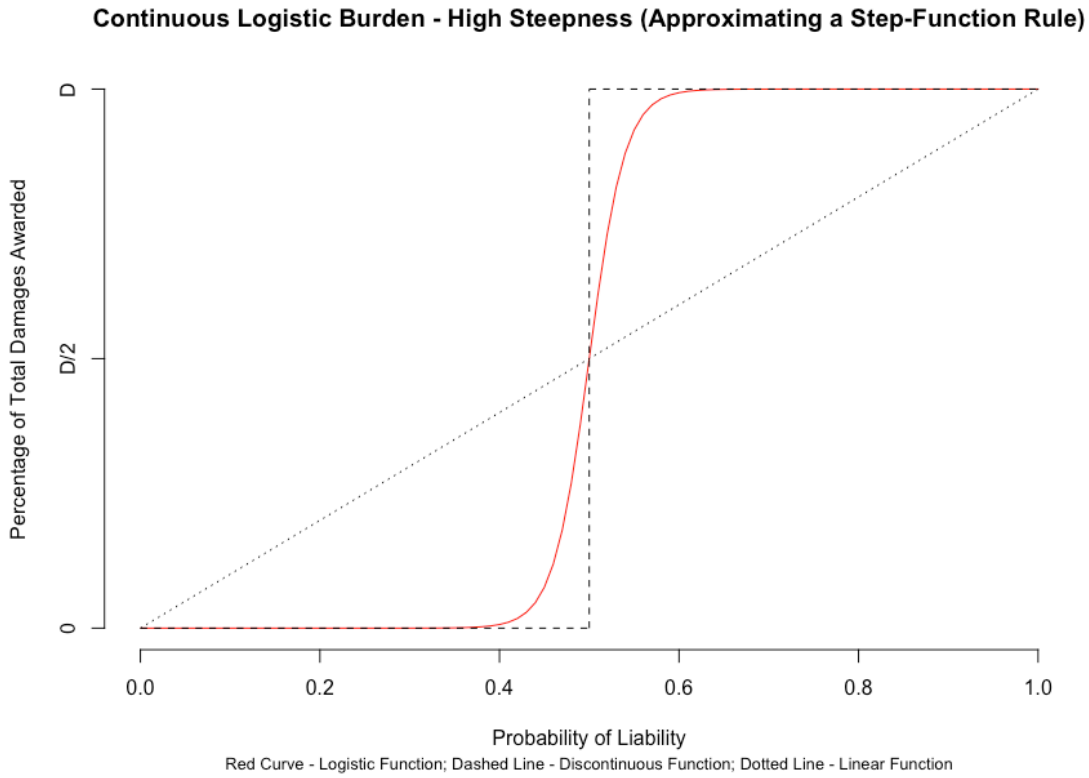
These features permit us to design a continuous burden of proof that can be arbitrarily close to either a straight line or a step function, as the following graphs will illustrate. At one extreme, the logistic function can take a nearly linear shape:[80]



And at the other extreme, it can be defined so that it is nearly identical to the discontinuous rule over most of its range, while still smoothing out the transition where the discontinuous rule's threshold would apply:[81]

[80] The low steepness graph is determined by r=3, C_1= 1.574, and C_2= -0.287. The function could in fact lie much closer to the line, and this amount of variation is used only for ease of visualization.
[81] The high steepness graph is given by r=50, C_1=1 and C_2=0.

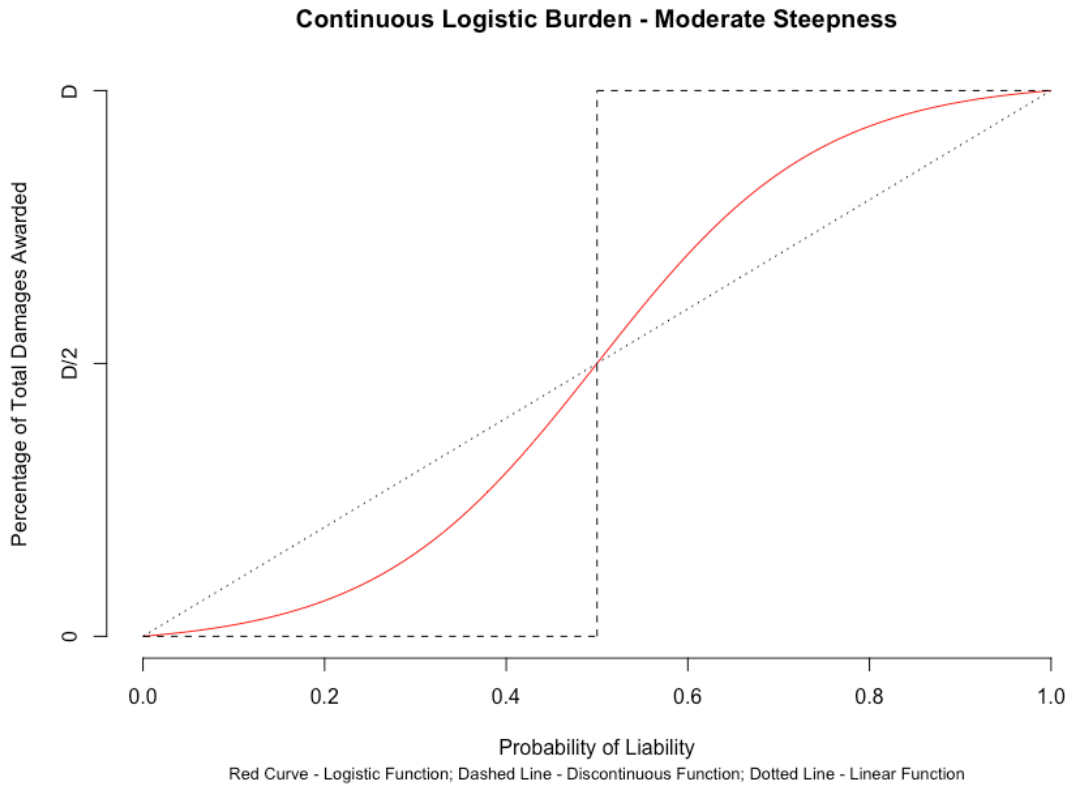
Electronic copy available at: https://ssrn.com/abstract=3348701



Most usefully, the logistic function can take a shape that is intermediate between these two extremes, allowing us to strike a true balance between objectives such as error minimization and deterrence:[82]

[82] The intermediate burden's parameters are r=8.155, C_1=1.03 and C_2=-0.07. Although these might seem somewhat arbitrary, the discussion *infra* will reveal that in addition to producing a logistic burden that lies neatly in the middle of the two extremes, they also optimize the sum of squared expected errors.

Electronic copy available at: https://ssrn.com/abstract=3348701



B. The Logistic Burden's Benefits

While this mathematical exercise may seem arcane to some legal readers, it has a very practical pay-off. First, note that all of these logistic functions, even the one that increases very steeply, assign 50% punishment at the 50% confidence mark. Thus, unlike a standard discontinuous burden rule, the logistic rule would always permit a civil jury to split the risk of error evenly when they feel the strength of the parties' cases is truly in equipoise.[83] Moreover, all of these rules would blunt the problematic effects of proof discontinuities on equality of outcomes, by reducing the amount of influence that small, legally irrelevant nudges towards one verdict or another can have on damages. And as the logistic burden rule grows closer to the linear one, we should expect to realize a more optimal quantity of deterrence, reduce the average magnitude of legal errors, and possibly decrease incentives towards spoliation of evidence.[84]

[83] *See* Coons, *supra* note _, at 756-57 (doubting that there is any equitable reason to award true ties to the defense, rather than splitting the verdict evenly in such cases).
[84] See Discussion, *supra* at Part III.

Electronic copy available at: https://ssrn.com/abstract=3348701

Best of all, the logistic burden lets us balance these advantages against the rise in error costs we can expect as we move towards a linear rule. We can construct expected error functions for the logistic curve as follows:[85]

$$E_{total}(p) = D\left(\frac{C_1}{1+e^{-rp+\frac{r}{2}}} + C_2\right) * (1-p) + D\left(\frac{C_1}{1+e^{-r(1-p)+\frac{r}{2}}} + C_2\right) * p$$

When we graph this error curve for the logistic liability rule side-by-side with the others that Kaye explored, we can see that the logistic rule always lies between the other two functions in terms of its overall expected rate of error.[86] Thus, we can choose parameters for it that strike whatever balance between the costs and benefits that we prefer. The low-steepness, nearly linear logistic burden we saw before yields an expected error curve that is just a bit better than the purely linear rule, therefore yielding excellent deterrence at the cost of higher overall error:

[85] Errors favoring the plaintiff will simply be the value given by the logistic function, multiplied by the expected probability that the defendant is not liable, which is 1-p:

$$E_1(p) = D\left(\frac{C_1}{1+e^{-rp+\frac{r}{2}}} + C_2\right) * (1-p)$$

Errors favoring the defendant will be the amount of the verdict *not* paid to the plaintiff under the logistic rule, multiplied by *p*, the probability of liability, as follows:

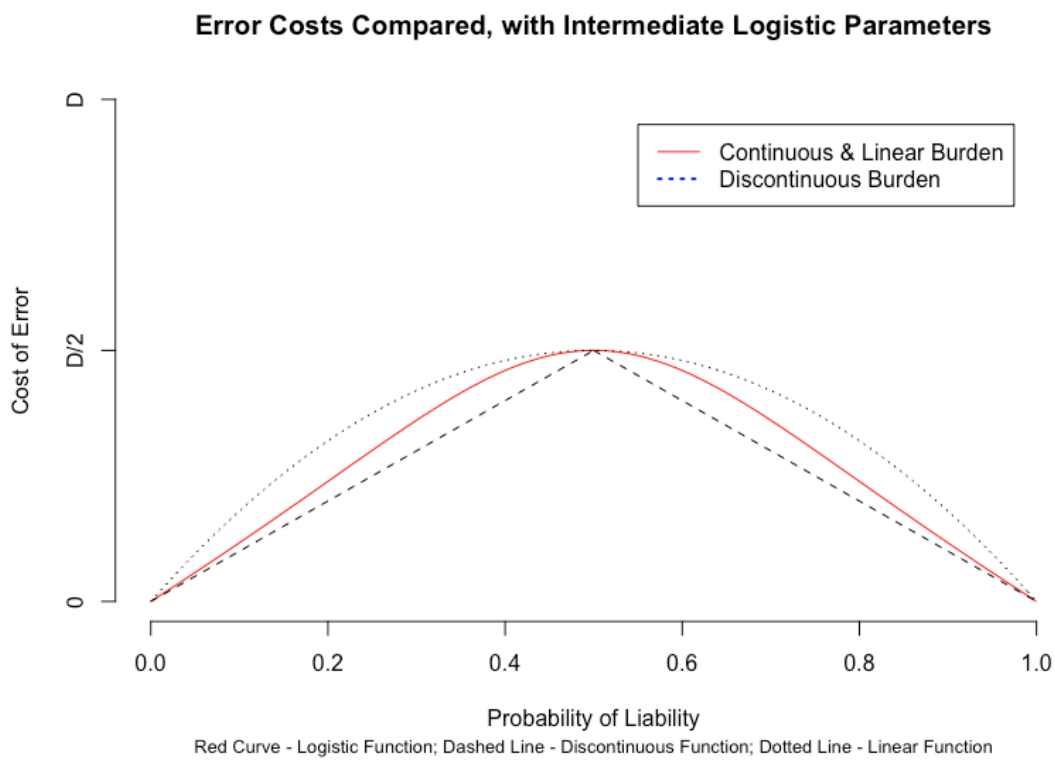
$$E_2(p) = D\left(\frac{C_1}{1+e^{-r(1-p)+\frac{r}{2}}} + C_2\right) * p$$

And the total error function, E_{total}, will simply be the sum of E_1 and E_2, which is expressed above in text.

[86] The following graphs use the same parameters for the logistic burden as the low, moderate, and high steepness versions illustrated above. The error curves for the linear and discontinuous rules are also the same as those illustrated above.

Electronic copy available at: https://ssrn.com/abstract=3348701

As we might expect, the parameters previously chosen for a moderate logistic burden yield an expected error curve that is intermediate between the two extremes. Its increase in expected error over the traditional rule is just less than half the size of what we expect with the linear rule (from 25% of the damages to 28.7% instead of the 33.3% we get with a linear rule).[87]



And as we would expect, the high-steepness logistic curve, which approximates the linear function, manages to smooth the central discontinuity while yielding a nearly identical rate of expected errors (and increase from 25% to only 25.1% of expected overall error):[88]

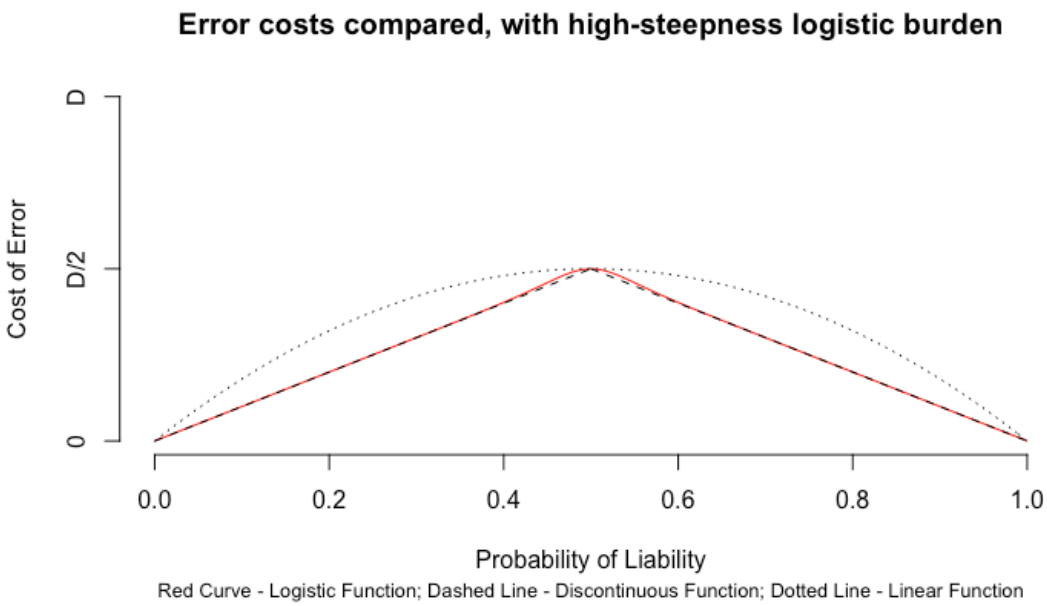
[87] This value is again obtained by integrating the expected error function for the logistic burden, with parameter values of r=8.155, C_1=1.03 and C_2=-0.07:

$$\int_0^1 ((\frac{C_1}{1+e^{-rp+r/2}} + C_2)(1-p) + (\frac{C_1}{1+e^{-r(1-p)+r/2}} + C_2)p)dp$$

Given the complexity of the integrand, I computed the value reported above.

[88] This was also computed using the above integral, with the values of r=50, C_1=1 and C_2=0.

Electronic copy available at: https://ssrn.com/abstract=3348701



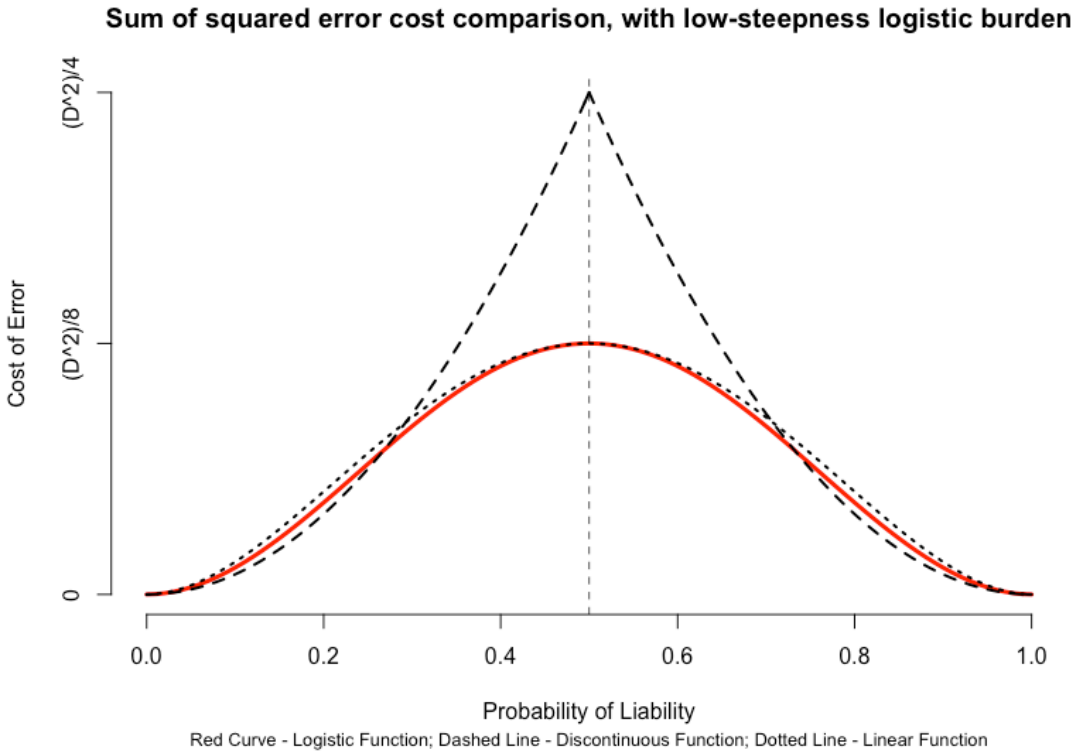
Finally, it is also illuminating to compare the squared-error curves of the three rules, to see how the logistic burden[89] stacks up in terms of its ability to avoid imposing particularly large errors on individual parties. On this metric, the logistic curve actually outperforms both of the other rules, across a wide range of parameter values. First, we can see that the low-steepness, approximately linear version lies has a lower expected squared error value than both rules in its central range, while slightly exceeding the linear rule's values at high and low values of jury confidence:

[89] The sum-of-squared-errors function for the logistic burden is as follows:

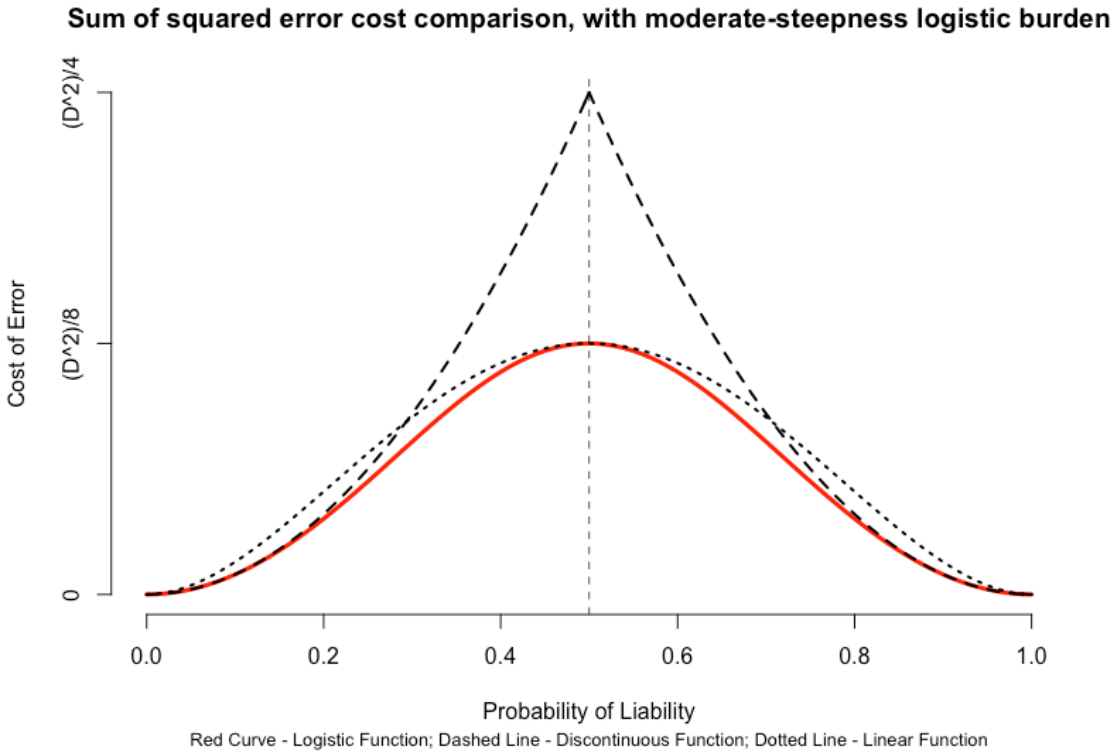
$$\text{Sum_of_}E^2(\text{p}) = \text{D}\left(\frac{\text{A}}{1+\text{e}^{-\text{rp}+\frac{\text{r}}{2}}} + \text{B}\right)^2 (1-p)^2 + \text{D}\left(\frac{\text{A}}{1+\text{e}^{-\text{r}(1-\text{p})+\frac{\text{r}}{2}}} + \text{B}\right)^2 p^2$$

The values of the C_1, C_2, and r constants for the low-, moderate-, and high-steepness logistic burdens are as given above.

Electronic copy available at: https://ssrn.com/abstract=3348701

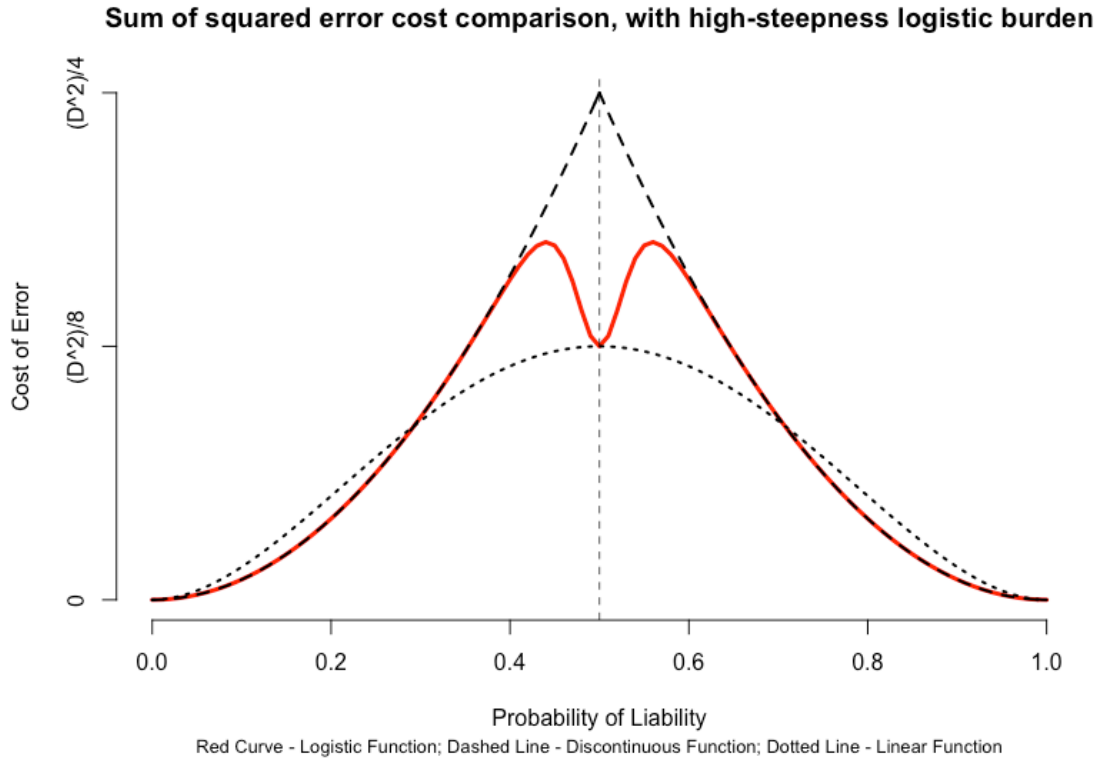


The logistic burden with moderate parameters does an even better job, managing to match or outperform *both* rules across the entire range of jury confidence levels:



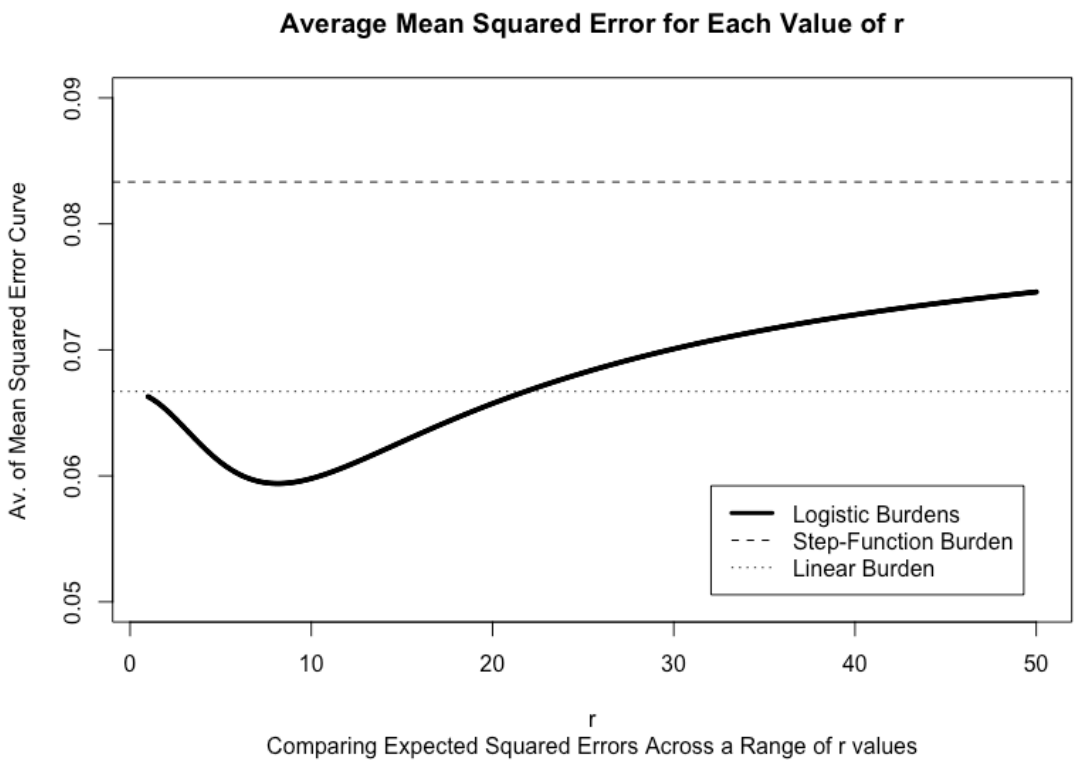
Electronic copy available at: https://ssrn.com/abstract=3348701

Finally, the "smoothed step-function," high-steepness version of the logistic burden performs admirably in the high and low ranges of jury confidence levels, but shows a tendency to magnify errors in its central range (although not to the full extent of the traditional rule):



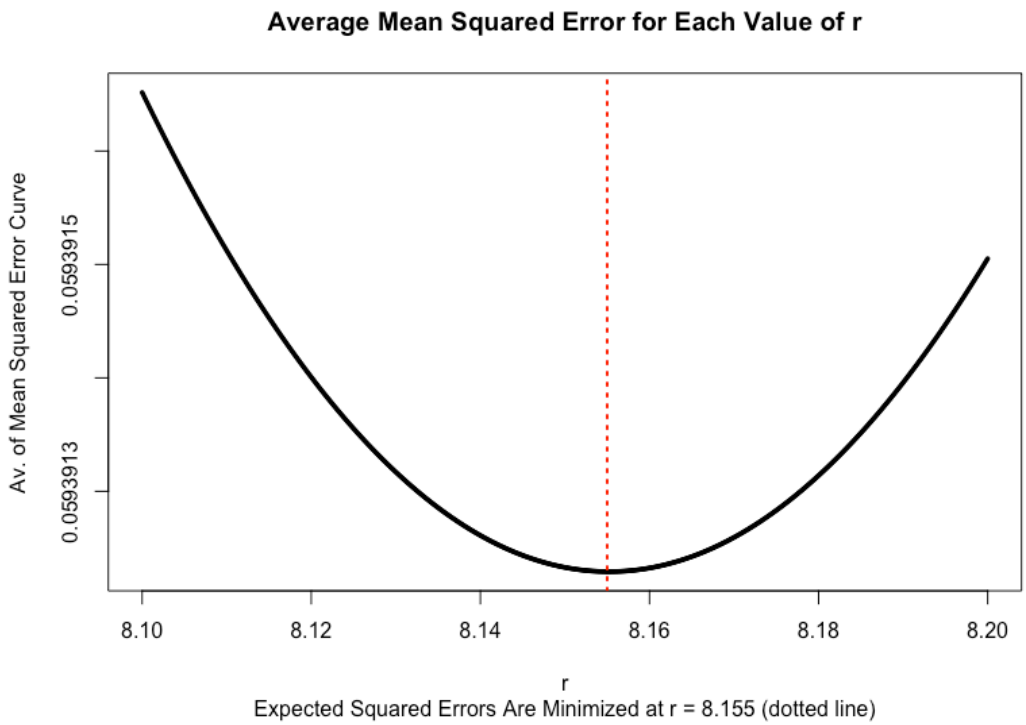
As these graphs illustrate, if we wish to obtain almost all of the error reduction of a step-function rule, we can do so with a logistic burden of proof that still eases the sharpness of the transition near the step-function's discontinuity, allowing some of the benefits discussed above. Moreover, one can easily survey the range of logistic parameters to find the version of the rule that minimizes overall expected squared errors. The following chart shows the overall expected error rate for each value of r between 0 and 50, with the scaling constants adjusted so that the burden always imposes 0 liability at 0 jury confidence, .5 liability at .5 confidence, and full liability at 100% confidence.

Electronic copy available at: https://ssrn.com/abstract=3348701



As can be seen, the logistic rule dominates both other rules across a very wide range of shapes on this metric. Zooming in, we can locate the optimal parameters from the standpoint of minimizing squared errors:

Electronic copy available at: https://ssrn.com/abstract=3348701



As can be seen in this last chart, the logistic burden has a minimized value of expected squared error (and thus the expected magnitude of errors that will be inflicted on individual parties) at the values of r=8.155, C_1=1.03 and C_2=-0.07. This is the "moderate" version of the rule I have illustrated above.

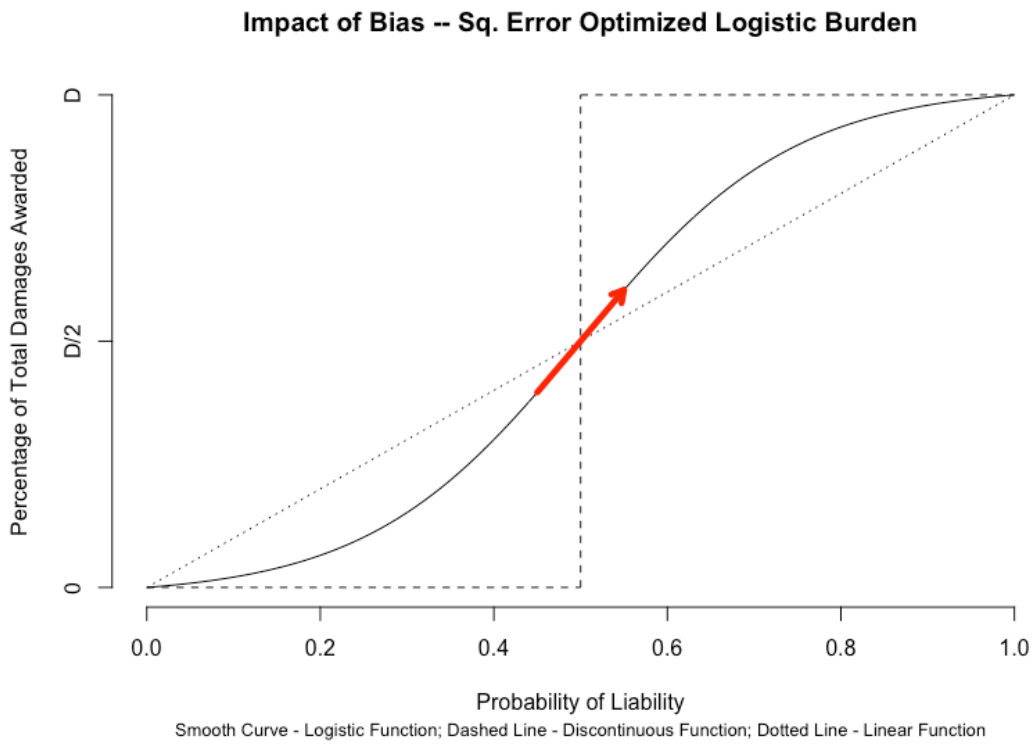
Several things are therefore notable about this intermediate formulation of the logistic burden of proof. First, as already shown, it minimizes the risk of very large errors inflicted on any single party in a case. Second, although it increases the absolute risk of overall error beyond what the traditional rule does, it does so by less than half of what the linear burden of proof would do. Third, by imposing some penalties in cases below the 0.5 threshold, and less than full penalties for cases that lie above that threshold but below full confidence, it is likely to give more optimal deterrence of wrongdoing than the traditional rule. Similarly, because the gains of destroying evidence are less likely to move a party from full liability to zero liability, it should reduce incentives towards spoliation.

This rule also goes a long way towards reducing the impact of unfair influences at trials. As discussed above, an influence such as racial bias or the ability to afford more effective advocacy is generally likely to shift overall jury confidence levels by only modest amounts. But since the traditional rule filters out "easy" cases through settlements while encouraging parties to take closer cases to trial, it can lead to such small influences radically changing a party's outcome, leading to substantial concerns about inequality of treatment of litigants.

This may be clearest if we revisit our earlier example, in which an improper influence on the jury shifted their confidence levels in liability from 45% to 55%. Under the traditional rule, this results in a change in the outcome of 100% of the damages, while under the linear rule, it results

Electronic copy available at: https://ssrn.com/abstract=3348701

in only a 10% increase in damages. Now consider the impact under the square-error-optimized logistic burden of proof:



Instead of a 10% increase, this version of the logistic burden increases the share of total damages awarded from 40% to 60%. While this does represent a doubled increase from the linear rule, it is only 20% of the change that our existing rule would yield, which is still a large improvement. So to sum up, this variant of the logistic burden minimizes large errors, cuts the impact of biases and similar influences by 80% in close cases, and otherwise strikes a balance between the worthy goals of deterring wrongdoing and minimizing total error. All things considered, this form of the rule clearly merits serious consideration for those who wish to improve the overall functioning of our trial process.[90]

Finally, although it might seem challenging to implement a logistic burden in an ordinary courtroom with lay jurors, the obstacles are almost certainly smaller than one might initially assume. A jury need not be instructed on the math included in the footnotes of this section, in order for a court to implement a logistic burden. Rather, a jury could merely be instructed to report their degree of belief in liability, represented on a scale from 0 to 100% confidence, in

[90] Of course, it is quite possible that some readers will prefer to strike a different trade-off among these competing values. If we think that the values of deterrence, equality of treatment, and spoliation reduction warrant a larger deviation from error minimization, we can choose a rule (such as the low steepness logistic burden) that gives these concerns greater weight. Or conversely, those who care the most about reducing overall errors can obtain almost full realization of the step-function's benefits on that score while smoothing its discontinuity by choosing a high-steepness logistic burden.

Electronic copy available at: https://ssrn.com/abstract=3348701

place of their current, dichotomous verdict decisions. Nor would judges be required to understand the intricacies of logistic functions, as they could be provided with software that seamlessly translates the jury's confidence level and damages determination into a specific verdict amount, or even a simple table that matches jury confidence levels with the damages that should be awarded under the optimized logistic burden. The only persons who would ever need to grapple with the details of the logistic function would be rule-makers, who would necessarily have to decide what balance to strike between the competing goals of error-reduction, on the one hand, and deterrence, equality of treatment, error-spreading, and spoliation reduction, on the other.[91]

IV. SMALL STEPS TOWARDS A CONTINUOUS BURDEN

In the preceding sections, I have set forth some reasons why adopting continuous burdens of proof would be beneficial for all kinds of cases within our legal system. Of course, anyone who has studied the history of trial by jury knows that it is an institution that is slow to evolve. Moreover, when major changes happen, it is usually in the wake of a highly visible crisis in the court system, rather than as a response to considered reflection.[92] In addition, legal change typically exhibits some amount of path dependency, and some existing aspects of American law would make it very hard to adopt a continuous burden across the board within our legal system. Such obstacles do not make analysis of the foregoing kind useless, however. It is always hard to predict when crises may come and what shape they may take, so there is value in having thought through issues in advance. Moreover, obstacles that exist in one country may be absent in another, and even seemingly unshakeable doctrines sometimes give way over time.[93] Still, some readers may yearn for more readily implementable suggestions. Accordingly, after reviewing some of the obstacles to widespread implementation of the continuous burden, I will explore some small steps we might take, consistent with existing American practices, towards realizing some of its benefits.

---

[91] One interesting aspect of such an analysis, which I do not have space to fully explore in this article, is the possibility that the appropriate parameters might be adjusted to counteract certain biases in estimating probabilistic likelihoods of guilt. One potentially relevant bias that might be taken into account is the tendency to give evidence greater or lesser weight depending on whether it is coherent with a juror's existing intuitive view of the case. Such tendencies can result in estimates of guilt that are systematically pulled towards extreme values, leading jurors in close cases to think that one result is obviously more defensible than the other. *See generally* Dan Simon, *A Third View of the Black Box: Cognitive Coherence in Legal Decisionmaking*, 71 U. Chi. L. Rev. 511 (2004). Such a tendency will naturally tend to pull estimations of guilt probability away from the mean, so that a given punishment function will resemble a somewhat steeper logistic function in its actual impact. Accordingly, if we could obtain consistent and reliable measures of this tendency across likely jurors, we might wish to adopt a logistic curve of lower steepness than would otherwise make sense on policy grounds, knowing that the coherence bias would lead to more steeply diverging assignments of damages or sentences than the function would assign if given unbiased inputs.

[92] *See, e.g.*, Mark Spottswood, *Truth, Lies, and the Confrontation Clause*, 89 U. Colo. L. Rev. 566, 570-78 (2018) (describing the crisis of widely acknowledged perjuries that led to the development of the law of hearsay and the criminal defendant's right to confront adverse witnesses).

[93] *See, e.g.*, Fisher, *supra* note _, at 656-96 (describing the downfall of the witness competency rules that provided the primary system of evidentiary control prior to the 1800s).

Electronic copy available at: https://ssrn.com/abstract=3348701

A. Implementation Challenges

If we were designing a legal system from scratch we might wish to implement continuous burdens across the board. But our existing legal system has a few constraining features that would make it hard to adopt such a rule so broadly, particularly in criminal cases. First, there is the rule of *In re Winship*,[94] which gives constitutional mandate to the beyond a reasonable doubt threshold of proof in criminal cases. Next, most modern jurisdictions have adopted discretionary modes of criminal sentencing, which would complicate the simple transformation from confidence in guilt to sentencing outcomes described above. Finally, there is the inherent conservativism that our society has long adopted when considering reforms to the trial process.[95]

In re Winship held that the "beyond a reasonable doubt standard" is required by the Due Process Clause in all ordinary criminal cases, so that neither the states nor the federal government may impose criminal punishment based on a lower level of confidence.[96] *Winship*'s rule presents a serious obstacle to the implementation of a continuous burden in criminal cases. If we cannot convict persons below the threshold of reasonable doubt, we can only maintain a continuous burden by beginning with zero punishment at the lowest level of confidence that meets the standard, and rising upwards from there. But under such a rule, the very low levels of punishment that would be required at low levels of confidence *above* the threshold would seem far too low to offer adequate deterrence or punishment, especially for crimes like murder or serious assaults. We could still obtain some benefits in terms of reducing the maximum expected errors by implementing an escalating punishment function in the range from reasonable doubt to certainty of guilt.[97] Under some assumptions, this might also yield an additional benefit to deterrence.[98] But closely tied to the *Winship* rule is the widespread hostility of many courts to any attempts to define the reasonable doubt threshold in probabilistic terms.[99] In the absence of such a definition, we would be hard pressed to articulate the range within which the remaining flexibility could operate. And of course, the other benefits of a continuous burden, such as mitigating unequal treatment and disincentivizing spoliation, would be much reduced once we reintroduce a large discontinuity at the level of reasonable doubt.[100]

In addition to this well-rooted constitutional doctrine, those seeking to introduce continuous burdens in American criminal trials would face a second major obstacle, in the form of modern

---

[94] 397 U.S. 358 (1970).

[95] In addition to these considerations, the varying modes of criminal sanction, and the impacts of collateral consequences, would require careful consideration if we wished to design a continuous burden of proof for criminal cases. *See generally* Avlana K. Eisenberg, *Discontinuity in Criminal Law* (working paper 2020).

[96] *Id.* at 364, 368.

[97] *Cf.* Lando, *supra* note _, at 285 (arguing for a regime of criminal penalties that vary based on the jury's confidence in guilt, but only above a threshold level of confidence).

[98] *Id.* at 281-83 (arguing that escalating punishments based on confidence levels allows us to distribute limited punishment resources in a way that is more likely to deter future wrongdoing).

[99] *See, e.g.*, McCullough v. State, 657 P.2d 1157, 1159 (Nev. 1983) (stating that "[a]ny attempt to quantify" the standard "may impermissibly lower the prosecution's burden of proof, and is likely to confuse rather than clarify"). *See generally* Peter Tillers & Jonathan Gottfried, *Case Comment,* United States v. Copeland, *369 F. Supp. 2d 275 (E.D.N.Y. 2005): A Collateral Attack on the Legal Maxim that Proof Beyond a Reasonable Doubt Is Unquantifiable?*, 5 Law *P*robability & Risk 135, 135-36 (2006) (collecting cases).

[100] For this reason, I have not attempted to specify a precise mathematical form for criminal burdens in the analysis above.

Electronic copy available at: https://ssrn.com/abstract=3348701

sentencing doctrine. Historically, many states and the federal government had given judges broad discretion concerning the length of a sentence that a convicted offender would receive, so long as it fell within the overall statutory range for an offense.[101] In *Apprendi v. Washington*, however, the Supreme Court placed a clear boundary on that flexibility: judges could not rely on additional findings to levy sentences *outside* of the normal range applicable to the crime for which a jury had convicted a defendant, unless those findings had also been made by a jury, subject to the beyond-a-reasonable-doubt standard.[102] Meanwhile the federal government, and several states,[103] has moved to rein in judges' sentencing discretion even further through the use of mandatory sentencing guidelines, out of concern that the discretion produced indefensible variation among punishments given in factually similar cases. In response to this trend, the Court held that such innovations also violated the principle in *Apprendi*, at least so long as the judges were mandatorily bound to make upward adjustments in sentences on the basis of facts not found by a jury.[104] The only way that sentencing guidelines can continue in force is if they are merely advisory in nature, preserving to the trial judge the power to depart from their recommendations.[105]

Thus, the sentencing regime would further complicate the institution of a continuous burden in criminal cases. At best, juries might announce a confidence level in guilt, which would then form the center of a statutory range, within which the court's sentencing discretion would be preserved. Or alternatively, a judge's sentencing discretion might be *expanded* to permit the judge's own confidence levels to be used as a reason for varying sentences within the legislatively imposed range. Some have even speculated that judges may already be doing this privately, whether consciously or otherwise.[106]

So *Winship* stands firmly in the way of implementing continuous burdens in criminal cases, and the nature of modern sentencing adds further complications. The final obstacle, although entirely informal, may be the biggest roadblock of all: the innate conservatism of lawyers when considering reforms to trial procedure. This conservatism, more than anything else, leads me to doubt that any jurisdiction will fully adopt a continuous burden of proof anytime soon. Since most of the harms wrought by a dichotomous burden are subtle and hard to detect in any individual case, it seems unlikely that a crisis might arise that would force policymakers to reconsider such a long-standing part of the trial ritual. Nonetheless, given the strong arguments

---

[101] *See* Williams v. New York, 337 U.S. 241, 246 (1949). This was itself an evolution away from earlier common law practice, in which judges were bound to issue fixed and determinate sentences for particular offenses. *See* Apprendi v. New Jersey, 530 U.S. 466, 479-80 (2000). The seeming severity of *that* rule was mediated in practice by the frequent recommendation by judges that selected defendants receive royal reprieves or pardons. *See* J.M. BEATTIE, CRIME AND THE COURTS IN ENGLAND: 1660-1800, at 409.
[102] *Apprendi*, 530 U.S. at 490.
[103] *See* RICHARD S. FRASE, JUST SENTENCING: PRINCIPLES AND PROCEDURES FOR A WORKABLE SYSTEM 4-5 (2013) (describing the spread of guidelines-based models among the states
[104] Blakely v. Washington, 542 U.S. 296, 305 (2004).
[105] *See* United States v. Booker, 543 U.S. 220, 245 (2005) (severing the portion of the federal sentencing guidelines that made them mandatory in application, as a means of preserving the constitutionality of the broader scheme).
[106] *See* Lando, *supra* note _, at 285; *see also* Fisher, *supra* note _, at 847-48 (suggesting that judges enhance sentences following a jury verdict because it gives them greater certainty in factual guilt).

Electronic copy available at: https://ssrn.com/abstract=3348701

in favor of the continuous burden, it makes sense to consider some more modest ways in which it could be partially implemented, with less likelihood of resistance.

B. Areas Where Continuous Burdens May Meet Less Resistance

Despite the concerns raised in the previous section, there are a number of settings in which a continuous burden might be adopted with far less resistance. First, a great deal of fact-finding occurs outside of courts, typically with more flexibility than we see in a traditional trial. Second, there are some categories of civil cases where the dichotomous burden seems particularly ill-suited to do justice, such as the area of toxic torts, which might naturally be a good starting point for experimenting with a continuous burden. Adoption in any of these areas might permit judges and lawyers to gain familiarity applying a continuous burden, give society a chance to grow accustomed to it, and allow scholars to assess its costs and benefits in operation.

To begin with, there is obvious potential for experimenting with a continuous burden of proof in settings where fact-finding is done outside of the courts. Most major employers must make factual determinations when deciding whether to discipline employees for workplace infractions, among other matters.[107] Likewise, universities must inquire into facts before deciding whether students should be disciplined.[108] Many disputes that could go to court are instead handled by arbitrators, whether through a pre-dispute agreement or a decision to avoid the high costs and delays of court proceedings once a claim has been brought.[109] And finally, many administrative agencies, at both the state and the federal level, operate their own internal dispute resolution systems, in which administrative law judges hold hearings to determine such matters as benefit determinations, worker's compensation payments, or the deportability of aliens.

Two factors make instituting continuous burdens more feasible. First, they all involve a fair amount of flexibility as to the form that a fact-finding inquiry should take. Employers can choose whatever manner of internal investigation best suits their goals, for instance, and arbitrators can use any procedures that do not violate the terms of the parties' arbitration agreements. Second, the public devotes less attention to studying the details of such procedures, and they rarely form the basis of stirring TV dramas. As a result, decision-makers in these settings may be more willing to experiment, with less worry about "corrupting" ancient institutions and less fear of public outcry.

Still, not all determinations made in less-formal contexts would be a good fit for continuous burdens. The key factor that has made the continuous burden a theoretically attractive option in so many modern court cases is the fact that modern remedies and punishments are generally easy to vary on a smooth scale. Both damages payments, and prison sentences, can easily be made longer or shorter in order to correspond with varying levels of confidence. And some less-formal fact-finding is used in service of similarly flexible outcomes as well. An administrative law judge

---

[107] See generally Theodore R. Lotchin, No Good Deed Goes Unpunished? Establishing A Self-Evaluative Privilege for Corporate Internal Investigations, 46 Wm. & Mary L. Rev. 1137, 1138 (2004).
[108] See Tamara Rice Lave, Campus Sexual Assault Adjudication: Why Universities Should Reject the Dear Colleague Letter, 64 U. Kan. L. Rev. 915, 923 (2016).
[109] See generally Amy J. Schmitz, *Legislating in the Light: Considering Empirical Data in Crafting Arbitration Reforms*, 15 Harv. Negot. L. Rev. 115, 120-28, 136-43 (2010).

Electronic copy available at: https://ssrn.com/abstract=3348701

determining an appropriate level of compensation for a past harm, for instance, could operate much like a civil jury would under a continuous burden rule.[110] But some consequences are, in their nature, discontinuous. We cannot order 75% of a deportation, for instance.

Nonetheless, to the extent that less severe remedies can be ordered on a scale below a more serious and irrevocable outcome, we can imagine implementing partially continuous standard up to a reasonably high level of confidence, at which point the dichotomous decision is issued. Consider, for example, the ongoing controversy regarding standards of proof in campus investigations into student-on-student sexual misconduct. On the one hand, victim's advocates rightfully worry that many offenders could escape all consequences if we must prove their guilt beyond a reasonable doubt, as in a criminal case. Many campus rapes will occur under situations where the accused perpetrator can plausibly claim consent, making the outcome turn on hard-to-unravel questions of credibility.[111] As a result, it will be hard for most decision-makers to reach a near-certain level of belief in guilt even when wrongdoing has truly occurred. On the other hand, under a preponderance rule, wrongfully accused persons might also reasonably fear wrongful expulsion from campus, with serious long-term collateral consequences, even in cases where fact-finders feel the claimant's case is just stronger than a coin-toss. The result seems to be a zero-sum conflict, where concern for one group's interests must necessarily lead one to be dismissive of the other group.

Happily, the implementation of a continuous burden might ease both sides' concerns to a significant extent. We might imagine a college's Title IX officer defining a range of punishments, from a no-contact order of varying severity and length at low levels of confidence in guilt, through suspensions of varying length in the middle portions of the range, with expulsions limited to cases in which proof had been proved to a very high confidence level, such as 90% or above. Under such a rule, victims' ability to obtain *some* redress for their wrong, and especially to avoid further contact with their rapists, would be expanded, while accused persons could be more confident that they would not receive truly life-altering consequences unless university officials were strongly persuaded of their guilt.

Beyond these less formal sorts of fact-finding, let us next consider one other arena in which continuous burdens might be reasonably attainable: civil cases involving liability arising from exposure to allegedly toxic substances. Such cases present a number of recurring problems that make them especially ill-suited to a step-function liability rule, only some of which have been addressed by existing doctrines. First, there are often multiple producers of the same harmful substance operating in the same market. Since many toxic exposures cause disease only after long latency periods, many plaintiffs will be unable to remember which producer's product they ingested.[112] Second, even where plaintiffs can, in fact, remember whose products they were exposed to, there may be overlapping exposures to similar products. And since diseases like cancer may be caused by quite small quantities of a toxic substance, figuring out *which* product injured a plaintiff requires a statistical guess. Finally, most diseases have more than one potential

---

[110] *Cf.* Robert L. Ramsey & Robert S. Habermann, *The Federal Black Lung Program-the View from the Top*, 87 W. Va. L. Rev. 575 (1985) (describing the administrative process by which claimants can recover benefits for pneumoconiosis disabilities arising out of coal mine employment).
[111] Naomi M. Mann, *Taming Title IX Tensions*, 20 U. Pa. J. Const. L. 631, 651 (2018)
[112] *See* Sindell v. Abott Labs, 26 Cal 3d 588 (1980).

Electronic copy available at: https://ssrn.com/abstract=3348701

cause, with the result that even a certain exposure only leads to an increased risk of toxic injury.[113] Therefore, the fact that an exposed person later suffered a particular disease does not mean that their disease was caused by the exposure. Unfortunately for the law, the most that scientific inquiry can usually tell us in such cases is that the exposure raised the "relative risk" of disease by a given amount.[114]

As other scholars have noted, all of these factors make a discontinuous burden of proof a particularly poor fit for toxic torts.[115] Consider a company who know that their drug raises the risk of a particular cancer which is ordinarily found in 10 out of every 100,000 people, with the result that now 15 out of 100,000 people will get it. The preponderance standard will prevent this company for every being liable for harm, given that in any individual case, it will be more likely than not that the person would have become ill even if they had *not* taken the drug.[116] As a result, this company has minimal incentives to warn consumers of the risk or find ways to further mitigate it, even when the costs of doing so are small in relation to the benefit to society at large. The same result could arise if the disease was caused more often, but the market was divided between many producers of a single generic drug, none of whom had a market share above 50%. A few courts have already begun to consider using continuous burdens to deal with the market share cases.[117] But since the problems sweep more broadly, it seems that this might be an area where the continuous burden might have more appeal as a *general* rule. Indeed, doing so would forestall one problem that will inevitably arise if courts try to solve these problems in a piecemeal fashion: what happens when plaintiffs present claims that include multiple forms of irreducible uncertainty, such as probabilistic proof of disease causation and a fragmented market of similar products? Citing to prior precedents where an *ad hoc* adjustment was made to deal with one of these forms of uncertainty in isolation will not provide a procedure for addressing them simultaneously in a fair way. The continuous burden rule, by contrast, would allow a jury to take each form of uncertainty into account (ideally with guidance from expert witnesses or an appointed master), in order to reach a decision that optimizes deterrence and fairly distributes the risk of errors.

C. Areas Where Some of the Continuous Burden's Benefits May Already Be Seen in Practice

Having considered some small steps we might take towards adopting a continuous burden, it is also worthwhile to note some ways in which our existing system has subtly embraced its underlying logic. Across both civil and criminal cases, there are a number of extant doctrines that allow sanctions to vary in response to varying levels of confidence in culpability. I have already referred to the most obvious such area above, which is the doctrine of market share liability in toxic tort cases. In the criminal context, the most obvious and visible erosion of the discontinuous approach to assigning guilt is the doctrine of residual doubt, which applies in the

---

[113] *See* Matsuyama v. Birnbaum, 890 N.E.2d 819, 827–28 (Mass. 2008) (collecting cases that allowed proportional recovery for an increased risk of disease due to an exposure).
[114] *See* Daubert v. Merrell Dow Pharm., Inc., 43 F.3d 1311, 1321-22 & n.13 (1995).
[115] *See* Discussion, *supra* Parts II-A and III.
[116] *Daubert*, 43 F.3d at 1321-22 & n.13.
[117] *See* Matsuyama, 890 N.E.2d at 827-28; Sindell 607 P.2d at 936-38; In re 'Agent Orange' Prod. Liab. Litig., 597 F. Supp. 740 (E.D.N.Y. 1984).

Electronic copy available at: https://ssrn.com/abstract=3348701

context of capital sentencing. Where applicable, the doctrine permits a jury that has convicted a person beyond a *reasonable* doubt to consider any remaining, smaller doubts as a mitigating factor in favor of a non-capital sentence.[118] This approach sensibly acknowledges that a jury might still think there was a significant chance that their verdict might be shown to be mistaken, even after convicting a person of a capital crime. Given the irrevocability of a death sentence, such concerns provide one very reasonable basis for deciding on a sentence of life in prison instead.

There has been great controversy over the applicability and extent of the residual doubt doctrine. The Supreme Court has rejected the idea that criminal defendants are constitutionally entitled to have the sentencing judge instruct their jury to consider residual doubt as a mitigating factor,[119] but it has not forbidden the states from including it as part of their own approach to capital sentencing. Accordingly, some federal and state courts permit defendants to argue residual doubt to a jury, some mandate that it be considered, and some bar defendants from raising the issue at all.[120]

Given the strong arguments in favor of the continuous burden of proof, it seems sensible to embrace the doctrine of residual doubt as a partial implementation of it. After all, if we are interested in reducing the magnitude of our errors, there seem to be few better places to start than by lowering the likelihood that we will execute factually innocent persons. To be sure, adopting the doctrine does not dispose of the discontinuity at the initial level of reasonable doubt, but it does reduce some of its perverse impact in the higher ranges of jury confidence. Accordingly, it would appear that more states should permit residual doubt to be argued, or even go further and require its consideration.[121] Conversely, it has been suggested that whether or not they are instructed or encouraged to do so, most juries will, in fact, accord significant weight to residual doubt in determining whether a death sentence is appropriate.[122] At the very least, the arguments above suggest that this reality should not concern us, and that we would do more harm than good if we attempted to override this natural aspect of jury reasoning with instructions to the contrary.

We might even imagine building upon the residual doubt doctrine to realize similar benefits in non-capital cases.[123] Even if the details of guidelines-driven sentencing currently make it difficult to realize a pure continuous burden rule, they do provide a means by which judges could adjust sentences above the reasonable doubt threshold based on varying levels of confidence in guilt. Given the Supreme Court's requirement that any guidelines be advisory in nature, a sentencing judge might reasonably depart downwards from the guidelines range if the jury's verdict (or a defendant's plea) seems substantially doubtful, even though it rests on legally sufficient evidence. Meanwhile, the judge might use their own sense of the probability of guilt to move their discretionary sentence towards the bottom of a guidelines range when guilt seemed doubtful, or

---

[118] *See* Fisher *supra* note _, at 838-40.
[119] *See* Franklin v. Lynaugh, 487 US 164 (1988).
[120] Fisher, *supra* note _, at 840-41 (surveying cases).
[121] *See, e.g.*, State v. Hartman, 42 S.W. 3d 44 (Tenn. 2001) (noting that "Tennessee law requires that a defendant be allowed to present evidence at a re-sentencing hearing to establish residual doubt as a nonstatutory mitigating circumstance").
[122] *See* Bruce A. Antkowiak, *Judicial Nullification*, 38 Creighton L. Rev. 545, 582-83 (2005).
[123] *See also id.* at 549-50 (making a pre-*Booker* argument that the residual doubt doctrine should be extended and used by sentencing judges in non-capital cases).

Electronic copy available at: https://ssrn.com/abstract=3348701

towards the top when it seems nearly certain. Such a thoughtful use of discretion might well work to smooth the proof discontinuity a bit, ensuring that we reserve our most serious punishments for offenders whose guilt is particularly likely, and mitigating the risk of imposing punishment on those more likely to be innocent.

One other way in which continuous burdens may be partially realized in existing practice is in the rising trend towards negotiated resolution of cases. Parties who are attempting to settle civil cases or reach plea bargains in criminal ones naturally take the strength of anticipated evidence into account when formulating their bargaining positions.[124] As a result, parties will tend to reach agreements that levy larger sentences or payments when the evidence against a defendant is strong, then they would in cases where the evidence is weaker. This variation, in turn, may serve to roughly approximate the valuation that a fact-finder would levy under a continuous burden of proof.[125] To be clear, I do not mean to suggest that the correspondence is ideal or close to ideal; disparities in bargaining power or access to evidence may lead parties to mis-value their cases, especially if little discovery has been had at the point when a bargain is reached. Moreover, the estimated probability of victory at trial is connected to the probability that a party committed an offense, but the two concepts are not identical, and in practice they may diverge to some extent.[126] And of course, we might have other reasons to be concerned about high rates of settlement beyond a worry that they unfairly deprive deserving plaintiffs of the full value of damages, such as a foregone opportunity for courts to publicly declare what the law is or for the public to learn about a defendant's harmful ongoing conduct.[127] Still, to the extent that a continuous burden provides the benefits I have articulated above, we might be less concerned about the trend towards more settlements and plea bargains, given that they may result in better deterrence, more equal treatment among potential offenders, and less risk of imposing large errors on undeserving parties.

---

[124] *See* Stewart J. Schwab & Michael Heise, *Splitting Logs: An Empirical Perspective on Employment Discrimination Settlements,* 96 Cornell L. Rev. 931, 937 (2011) (noting that most economic models of settlement "share common conceptual ground insofar as litigants' assessments of their legal exposure inform their bargaining positions").

[125] *See* Mark Spottswood, *Proof Discontinuity and Civil Settlements* (working paper 2020) (using data simulation to assess the effects that different burdens of proof are likely to have with respect to settlement behavior, and observing that under all forms of burden assessed in this paper, the expected outcome of settlements closely approximates the result that the continuous burden would yield at trial, albeit with wide variance in individual outcomes).

[126] *See id.* Most obviously, parties may mis-forecast their chances of success due to optimistic bias. *Cf.* Stephanos Bibas, *Plea Bargaining Outside the Shadow of Trial*, 117 Harv. L. Rev. 2463, 2498–99 (2004). Less obviously, a given estimated probability of liability is consistent with a surprising range of probabilities of victory, depending on how much variance there will be among potential decision-makers in their own subjective estimations. Thus, in a case where a typical juror would think the defendant is 60% likely to have committed the alleged wrong, there might be a near-certain probability of victory if nearly every potential fact-finder would agree that the probability was within 5% of that central figure. Conversely, a party might think that the estimation was much more noisy, with some fact-finders being likely to think the appropriate probability was 80% and other fact-finders being likely to think it was 40%. Despite having the same central tendency, such a distribution of possible subjective estimations would yield a much lower predicted likelihood of success at trial. *See* Spottswood, *supra* n. 117. *Cf.* Alex Stein, *Inefficient Evidence*, 66 Alabama L. Rev. 423, 432-33 (2015) (conceptualizing this kind of variation as the *noisiness* of potential evidence).

[127] *See* Marc Galanter & Mia Cahill, "Most Cases Settle": Judicial Promotion and Regulation of Settlements, 46 Stan. L. Rev. 1339, 1378-87 (1994) (surveying a range of concerns regarding the public impacts of settlements). *Cf.* Spottswood, *supra* n. 117 (finding that the continuous burden yielded an increase in the rate of settled cases, but that the resulting settlements involved a higher rate of expected error overall).

Electronic copy available at: https://ssrn.com/abstract=3348701

CONCLUSION

A practice as ancient and familiar as our discontinuous burden of proof can easily go unexamined. But when we look closely at our existing burden of proof and compare it with potential alternatives, it turns out to have a number of significant downsides. The discontinuous liability rule has one major strength, which is its tendency to minimize the total value of expected errors in simple cases. Unfortunately, it has many weaknesses as well. Compared with a continuous burden of proof, it will often provide suboptimal deterrence to wrongdoers. It will exacerbate inequalities of outcome due to wealth, status, race, and other invidious factors in legal decision-making. It will concentrate all risk of error on a single party, and thus produce a greater share of truly distressing outcomes, such as death sentences given to factually innocent persons or full damages extracted from a company that committed no legal wrong. And it may also give parties stronger incentives to spoliate evidence than would exist in a continuous regime.

The best form of burden of proof would strike a reasonable balance between these competing concerns. As I outlined above, the logistic burden of proof has a strong claim to strike the best balance in civil cases. Such a burden of proof can minimize the risk of large errors by optimally spreading the risk of error across both parties. It can strike a trade-off between the otherwise hard-to-reconcile goals of deterring wrongdoing and minimizing the total quantity of expected error. And it can go a long way towards reducing the impacts of unfair biases and the incentive to hide or destroy important evidence.

Of course, the burdens of history and tradition loom large over those who would seek to optimize our trial processes in novel ways. However, there are some ways in which we might begin to experiment with explicitly continuous burdens of proof. At the same time, I have sought to illustrate that our system, in action, may assign liability and punishment in ways that are closer to the continuous ideal than our officially discontinuous rule. The doctrine of residual doubt, the flexibility of judicial sentencing, the possibility of jury compromise, and the regular resort to settlements and plea bargains provide many ways in which doubts about guilt or liability can reduce the assessed level of punishment for a potential wrong-doer. To the extent that readers are convinced by the normative claims I have urged above, these practices may be more valuable than might otherwise be supposed. Either way, I hope that my observations may stimulate further discussion on the optimal design of burdens of proof.

Electronic copy available at: https://ssrn.com/abstract=3348701