Article

Conviction Without Conviction

Talia Fisher[†]

INTRODUCTION

Under Romano-Canon law and, later, under medieval English and Continental law, "full proof" requirements for serious crimes entailed the incriminating testimony of at least two witnesses.[1] However, the testimony of one witness or the testimony of two witnesses who were insufficiently credible could amount to "half proof," which justified the imposition of a lenient sanction.[2] Partial certainty of guilt resulted in partial punishment, reflecting what Michel Foucault termed the "continuous gradation" principle:

1.　*See* Charles L. Barzun, *Rules of Weight*, 83 NOTRE DAME L. REV. 1957, 1964 (2008) ("The number of witnesses required to prove various acts varied, but for many crimes, such as murder, the testimony of at least two witnesses was required for the 'full proof' necessary to sustain a conviction."); Barbara J. Shapiro, *"Fact" and the Proof of Fact in Anglo-American Law*, *in* HOW LAW KNOWS 25, 30 (Austin Sarat et al. eds., 2007) (describing the sources of the two-witnesses requirement and its adoption by various systems of law); John H. Wigmore, *Required Numbers of Witnesses; A Brief History of the Numerical System in England*, 15 HARV. L. REV. 83, 84 (1901) (describing the two-witnesses prerequisite in Roman law, early English law, and Continental civil law). The "full proof" requirement could also be fulfilled when the accused confessed to the alleged crime.

2.　JOHN H. LANGBEIN, TORTURE AND THE LAW OF PROOF: EUROPE AND ENGLAND IN THE ANCIENT REGIME 47 (1977) (describing the development of the system of full-proof and half-proof and discussing the practice of *Verdachtsstrafe*, "punishment for suspicion," partial punishment upon the court's belief in a defendant's guilt without full Romano-Canonical proof).

> Guilt did not begin when all the evidence was gathered together; piece by piece, it was constituted by each of the elements that made it possible to recognize a guilty person. Thus a semi-proof did not leave the suspect innocent until such time as it was completed; it made him semi-guilty; slight evidence of a serious crime marked someone as slightly criminal. In short, penal demonstration did not obey a dualistic system: true or false; but a principle of continuous gradation; a degree reached in the demonstration already formed a degree of guilt and consequently involved a degree of punishment.[3]

Unlike the linear (continuous gradation) conceptualization of criminal guilt and punishment in the ancient juridical worlds, today's criminal judgments are construed in a binary manner.[4] The decision-making processes underlying the determination of guilt and punishment in criminal trials are currently governed by what will hereinafter be termed the "threshold model." The threshold model construes conviction as an on-off decision, leading to all-or-nothing sentencing. According to this decision-making structure, when the incriminating evidence fails to cross the beyond-a-reasonable-doubt threshold, the defendant must be categorically acquitted.[5] The court can-

---

3. MICHEL FOUCAULT, DISCIPLINE AND PUNISH 42 (Alan Sheridan trans., Vintage Books 2d ed. 1995); *see also* EDWARD PETERS, TORTURE 84 (expanded ed. 1996) (claiming that in the Romano-Canonical systems of the ancient world, it took evidence to acquit as well as to convict, and when evidence for either was lacking, the imposition of partial punishments filled the epistemic gap).

4. *See* Samuel Bray, *Not Proven: Introducing a Third Verdict*, 72 U. CHI. L. REV. 1299, 1308 (2005) (analyzing the binary nature of the criminal verdict).

5. The United States Supreme Court has repeatedly held that the due process clause prohibits the conviction of a defendant except upon proof beyond-a-reasonable doubt of every fact necessary to constitute the crime with which he is charged. *See, e.g.*, *In re* Winship, 397 U.S. 358, 364 (1970). *In re Winship* was a juvenile delinquency case, in which the defendant was found guilty of what would have constituted a criminal act for an adult, by the preponderance-of-the-evidence standard of proof then required by New York law. *Id.* at 360. The judge conceded that under a "beyond a reasonable doubt standard" the verdict might have been different. *Id.* Following his conviction, the defendant challenged the constitutionality of the "preponderance of the evidence" standard. *Id.* The New York appellate courts affirmed the constitutionality of the law. *Id.* This United States Supreme Court reversed, holding that in criminal cases, proof of guilt "beyond a reasonable doubt" was a constitutional requirement of the Due Process Clause. *Id.* at 364; *see also* United States v. Gaudin, 515 U.S. 506, 509–11 (1995) (noting that Constitution give a criminal defendant the right to demand a jury find him guilty of all elements of a crime); Jackson v. Virginia, 443 U.S. 307, 326 (1979) (extending the constitutional requirement of the beyond-a-reasonable-doubt standard to appellate review).

The standard applies both with respect to misdemeanors and to felonies, at all degrees of offense, and even when the case is tried without a jury. *See, e.g.*, United States v. Collazo, 117 F.3d 793, 795 (5th Cir. 1997); United States v. Randolph, 93 F.3d 656, 660 (9th Cir. 1996). For a discussion of the origins of

not choose to convict the defendant to a lesser degree or express its epistemic doubt in the mitigation of the sentence. Satisfaction of the beyond-a-reasonable-doubt standard leads to the polar opposite result—of categorical conviction—and to the imposition of punishment whose severity is detached from the degree of persuasion as to the defendant's guilt.[6] Under the threshold model, once the defendant is found guilty beyond a reasonable doubt, any residual doubt as to the defendant's culpability is not included among the sentencing considerations and does not affect sentence severity.[7] The prevailing threshold model facilitates an all-or-nothing sentencing regime, where no punishment is imposed in the epistemic space below the beyond-a-reasonable-doubt threshold, while from that point upwards the severity of punishment is fixed in that it is disconnected from the weight of the evidence regarding culpability.

The threshold model, with its categorical formulation of criminal conviction and all-or-nothing sentencing configuration, has been established to the point of being considered axiomatic. The purpose of this Article is to call into question the inevitability of the threshold model. More specifically, this Article reassesses the idea of guilt as a purely binary phenomenon that is limited to an on-off configuration (either guilty or not guilty). It will also reconsider the derivative distribution of punishment, whereby no punishment is imposed in the epistemic space below the beyond-a-reasonable-doubt threshold, while from that threshold upwards the severity of punishment is disconnected from any remaining doubt as to guilt.

---

the beyond-a-reasonable-doubt constitutional requirement, see Dale A. Nance, *Civility and the Burden of Proof*, 17 HARV. J.L. & PUB. POL'Y 647, 656–68 (1994). For an elaborate historical account, see generally JAMES Q. WHITMAN, THE ORIGINS OF REASONABLE DOUBT: THEOLOGICAL ROOTS OF THE CRIMINAL TRIAL (2008) (claiming that the beyond-a-reasonable-doubt rule is not an epistemological rule but a moral rule, which was aimed at defending the morality of the jurors who sit in judgment); *see also* BARBARA J. SHAPIRO, "BEYOND REASONABLE DOUBT" AND "PROBABLE CAUSE": HISTORICAL PERSPECTIVE ON THE ANGLO-AMERICAN LAW OF EVIDENCE (1991) (exploring legal standards of proof in western law).

6. *See* Elizabeth R. Jungman, *Beyond All Doubt,* 91 GEO. L.J. 1065, 1089 (2003) (claiming that all questions of guilt are understood to be resolved at the conviction phase of trial).

7. The severity of the criminal sanction is supposed to be based solely on variables relating either to the gravity of the particular offense or to the characteristics of the convicted defendant herself. *See* Elizabeth T. Lear, *Is Conviction Irrelevant?*, 40 UCLA L. REV. 1179, 1179 (1993) (claiming that the sentencing guidelines are infused with considerations relating to the circumstances surrounding the commission of the crime).

The threshold model will be evaluated by contrasting it with the alternative regime of "probabilistic decision making." Under the probabilistic model, criminal guilt and punishment are construed in a linear manner: a plurality of conviction options exists along the evidentiary spectrum and punishment gravity is correlated with certainty of guilt. From a practical standpoint, this means a decision regime that recognizes different classes of convictions, including intermediate categories, such as conviction on guilt beyond a reasonable doubt, conviction on guilt by clear and convincing evidence, or conviction by preponderance of the evidence. Punishment severity is adjusted to the corresponding conviction category. For example, for a given offense, conviction on guilt beyond all residual doubt would yield maximal punishment (capital punishment, for instance), whereas conviction on guilt beyond a reasonable doubt would render near-maximal sanctions (allowing for life imprisonment but precluding capital punishment). Conviction by clear and convincing evidence for the same crime would lead to an intermediate level of punishment (less than a life sentence), while conviction on guilt by a preponderance of the evidence would entail only the lowest of the possible sanction alternatives (such as a fine).[8]

---

8. Of course, implementing a probabilistic regime would likely open the door to a continuous spectrum of intermediate convictions of exact statistical rates with corresponding degrees of punishment, such as, conviction at an 86 percent degree of certainty or 73 percent degree of certainty. However, in practice, the spectrum of choice can be expected to be more limited, for in typical cases, the court lacks the tools and information necessary for making an exact calculation of the probability of guilt. Consequently, the court draws its conclusions from categorical generalizations, making the judicial decision a crude assessment that is not amenable to precise statistical quantification. Irrespective of the practical hurdles, there are those who object in principle to turning the judicial decision into a precise statistical determination, as I will discuss further on in this article.

See infra Part III.C for a discussion of the expressive component of Charles R. Nesson's criticism. Nesson expressed the opinion that an exact quantification of the reasonable-doubt standard would impair its functional purpose, which is subjective in nature. *See* Charles R. Nesson, *Reasonable Doubt and Permissive Inferences: The Value of Complexity*, 92 HARV. L. REV. 1187, 1197 (1979); *see also* A.A.S. ZUCKERMAN, THE PRINCIPLES OF CRIMINAL EVIDENCE 132 (1989) (noting that people almost never have sufficient data to make precise measures of guilt); Lawrence M. Solan, *Refocusing the Burden of Proof in Criminal Cases: Some Doubt About Reasonable Doubt,* 78 TEX. L. REV. 105, 125–26 (1999) (noting that instructions focusing on doubt increase conviction rates); Laurence H. Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329, 1372 (1971) (noting problems with quantifying guilt).

The probabilistic model poses both a descriptive and a normative challenge to the threshold model. On the descriptive front I will claim that, contrary to common wisdom and to the theoretical understanding of decision making in the criminal arena, closer scrutiny reveals that central criminal law doctrines effectively deviate from the threshold model and exhibit the logic of the probabilistic model. The residual-doubt doctrine, the recidivist premium, and the jury trial penalty have established a de facto correlation between certainty of guilt and severity of punishment, infusing the criminal trial realm with probabilistic logic. In light of these instances of probabilistic decision-making, which have already taken root in the criminal arena, there is room to contest the inexorableness of the threshold model.

On the normative plane, this Article raises some of the principal arguments in favor of a linear conceptualization of criminal conviction and point to the utilitarian, expressive, re-tributivist, and political legitimacy bases for probabilistic sentencing. This Article shows that a sliding scale punishment, correlated with the certainty of guilt, is preferable to uniform punishment in the epistemic space above the beyond-a-reasonable-doubt threshold. It also demonstrates that when the level of certainty as to the defendant's guilt does not meet the beyond-a-reasonable-doubt standard of proof, conviction under a lower evidentiary standard and the imposition of partial punishment, reflecting the epistemic doubt, can also lead—in certain circumstances—to better outcomes than the existing alternative of acquittal without punishment.

This Article proceeds as follows: Part I will discuss prevailing criminal law doctrines and practices that express probabilistic decision making logic. Part II will briefly present the deterrence-based case for probabilistic sentencing. Part III will contend with the expressive dimensions of probabilistic decision making in criminal trials. Part IV will elucidate the re-

---

Supporting the normative difficulties in making a precise statistical quantification of judicial decisions is the fact that even "reasonable doubt" remains vague from a statistical perspective. Solan, *supra* at 126. Indeed**,** courts have systematically objected to a statistical quantification of this evidence standard. *See, e.g.*, State v. Cruz, 639 A.2d 534, 537–38 (Conn. App. Ct. 1994). It should be noted from a practical perspective that the majority of judges surveyed in one empirical study assessed the reasonable-doubt standard of proof as standing at 90 percent. Solan, *supra* at 126. Regardless, under a probabilistic punishment regime, judges can be expected to resort to general evidence standards that have already been recognized in the case-law, such as the preponderance-of-evidence or clear-and-convincing-evidence standards of proof.

tributivist and mixed retributivist-utilitarian considerations relating to the probabilistic model. Part V will examine the probabilistic model from a political legitimacy perspective. The final Part will offer concluding remarks.

I. PROBABILISTIC DECISION-MAKING UNDER PREVAILING LAW

Probabilistic decision making in criminal law is not just part of a distant past. Closer scrutiny reveals that central doctrines and practices, currently prevailing in the area of criminal law, exhibit the logic of probabilistic decision making. Some of these doctrines, such as residual doubt, create a direct and explicit correlativity between certainty of guilt and severity of punishment. Other legal practices, such as the recidivist sentencing premium or jury trial penalty, forge an indirect reciprocity. In this Part, I survey a number of the prevailing criminal law doctrines that embody probabilistic decision-making logic. I also offer preliminary support for the claim that regardless of the formal legal regime, decisions regarding sentence severity are effectively influenced by the extent of certainty regarding the defendant's guilt.

A. THE RESIDUAL DOUBT DOCTRINE

Residual doubt relates to the epistemic space that stretches between the beyond-a-reasonable-doubt standard of proof and the point of absolute certainty.[9] It creates a correlation between gravity of punishment and certainty of guilt when it acts as a mitigating factor—namely, "when juries decide not to impose a death sentence because they are not absolutely certain of the defendant's guilt."[10] To explain in brief, the fact that the jury was persuaded of the defendant's guilt beyond a reasonable doubt does not—in and of itself—preclude any remaining doubt as to culpability.[11] Even if such epistemic uncertainties do not

---

9. Franklin v. Lynaugh, 487 U.S. 164, 188 (1988) (O'Connor, J., concurring) (describing residual doubt as "a lingering uncertainty about facts, a state of mind that exists somewhere between 'beyond a reasonable doubt' and 'absolute certainty'").

10. Christina S. Pignatelli, *Residual Doubt: It's a Life Saver*, 13 CAP. DEF. J. 307, 314 (2001).

11. *See* Margery Malkin Koosed, *Averting Mistaken Executions by Adopting the Model Penal Code's Exclusion of Death in the Presence of Lingering Doubt,* 21 N. ILL. U. L. REV. 41, 55 (2001) (defining residual doubt as "doubt that is experienced, discussed, and ultimately remains, though it does not yield an acquittal of the crime or a negative finding as to the aggravating fac-

amount to reasonable doubt, they cannot be completely dismissed.[12] Under prevailing law, the existence of residual doubt as to a defendant's guilt in a capital offense could, in some cases, constitute a (non statutory) consideration for commuting the death sentence to life imprisonment.[13] Federal courts applied the doctrine sweepingly until it was redefined by the Supreme Court in *Franklin v. Lynaugh,*[14] where the Court held that a defendant does not have a constitutional right to have the judge instruct the jury that it should consider residual doubt about guilt in its sentencing decision.[15] The *Franklin* ruling did not

tor. It may not be a 'reasonable doubt' but it is a real doubt nonetheless"); Pignatelli, *supra* note 10, at 308 (defining residual doubt as "(1) actual, reasonable doubt about guilt of any crime; (2) actual, reasonable doubt that defendant was guilty of a capital offense, as opposed to other offenses; (3) a small degree of doubt about (1) or (2), sufficient to cause the juror not to want to foreclose (by execution) the possibility that new evidence might appear in the future").

   12.   As the Fifth Circuit said,
   The fact that jurors have determined guilt beyond a reasonable doubt does not necessarily mean that no juror entertained *any* doubt whatsoever. There may be no *reasonable* doubt—doubt based on reason— and yet some *genuine* doubt exists. It may reflect a mere possibility; it may be but the whimsy of one juror or several. Yet this whimsical doubt—this absence of absolute certainty—can be real.
Smith v. Balkcom, 660 F.2d 573, 580 (5th Cir. 1981).

   13.   *See* Jennifer R. Treadway, Note, *"Residual Doubt" in Capital Sentencing: No Doubt It Is an Appropriate Mitigating Factor,* 43 CASE W. RES. L. REV. 215, 215–28 (1992) (examining the treatment of the residual doubt doctrine by the Supreme Court and by the different states). Justice Thurgood Marshall has commented on the justification of the residual doubt doctrine:
   [O]ne of the most fearful aspects of the death penalty is its finality . . . . The belief that such an ultimate and final penalty is inappropriate where there are doubts as to guilt, even if they do not rise to the level necessary for acquittal, is a feeling that stems from common sense and fundamental notions of justice.
Heiney v. Florida, 469 U.S. 920, 921–22 (1984) (Marshall, J., dissenting from denial of certiorari).

   14.   487 U.S. 164, 183 (1988) (rejecting the petitioner's claim that the trial court's jury instructions—which did not inform the jury of the possibility of commuting the death sentence to life imprisonment due to residual doubt— were in violation of the Constitution). For further description of *Franklin v. Lynaugh*, see Jungman, *supra* note 6, at 1087.

   15.   487 U.S. at 175; *see also* Penry v. Lynaugh, 492 U.S. 302, 320 (1989), *abrogated on other grounds by* Atkins v. Virginia, 536 U.S. 304, 321 (2002) (characterizing *Franklin v. Lynaugh* as a case in which the majority agreed that "residual doubt" is not a constitutionally mandated mitigating factor). Moreover, it arises from the majority opinion in *Franklin* that even if it were within the ambit of constitutionally protected mitigating factors, there still is no corresponding right to jury instruction. *See* 487 U.S. at 174–75 (noting that the particular defendant had not been deprived of any constitutional right, as

address the broader and more fundamental question of whether, aside from the ensuing jury instructions, residual doubt regarding guilt can serve as a mitigating factor in sentencing. In its 2006 *Oregon v. Guzek* decision, the Supreme Court reiterated its ruling in *Franklin*, holding that a judge is not obligated to direct the jury with regard to the residual doubt doctrine, but that this does not preclude evidence regarding residual doubt from being presented and the jury from considering it.[16] The question regarding the defendant's right to present evidence of residual doubt to the jury was also deliberated in *United States v. Davis*, where it was held that in the federal legal system, defendants are permitted to raise a residual doubt claim at their sentencing hearing.[17] Moreover, the court held that when the defense makes such a claim, the jury is required to weigh it in determining the sentence.[18]

At the state level, there is great variance among the different jurisdictions on the issue of inclusion of residual doubt in the category of mitigating circumstances,[19] for *Franklin* left it for each state to decide as to the application of the doctrine in its justice system.[20] Some states, such as New Jersey, require both oral and written presentation of the residual doubt doctrine before the jury at the sentencing hearing.[21] In Tennessee, the defendant's attorney has the option of bringing a residual doubt claim before the jury.[22] By contrast, in Florida, the doctrine is not recognized and defendants are precluded from raising residual doubt as a mitigating factor in sentencing.[23] In

the trial court had not prevented the defendant or his counsel from presenting the residual doubt doctrine before the jury).

16. 546 U.S. 517, 525–26 (2006).

17. 132 F. Supp. 2d 455, 468 (E.D. La. 2001) (holding that in the federal system, defendants are permitted to raise residual doubt arguments under 18 U.S.C. § 3592 (2006).

18. *Id.* (ruling that residual doubt arguments must be considered by the jury if raised by the defense).

19. *See* Louis D. Bilonis, *Legitimating Death,* 91 MICH. L. REV. 1643, 1695 (1993) (surveying the different approaches state systems have adopted on the issue of residual doubt).

20. *Franklin*, 487 U.S. at 173; *see also* Treadway, *supra* note 13, at 222 n.58 (claiming that this approach is consistent with the Supreme Court's general unwillingness to impose constitutional controls with respect to the procedures that states use in their criminal justice systems).

21. *See* State v. Biegenwald, 594 A.2d 172, 197 (N.J. 1991) ("Jurors must be permitted to consider and give effect to any mitigating evidence.").

22. *See* State v. Hartman, 42 S.W.3d 44, 55 (Tenn. 2001).

23. *See* King v. State, 514 So. 2d 354, 358 (Fla. 1987); *see also* Jungman, *supra* note 6, at 1088.

Virginia, too, the doctrine is not applied.[24] The Virginia Supreme Court held in *Stockton v. Commonwealth* that defendants are not entitled to challenge a determination of criminal liability at the sentencing hearing, for the issue of guilt has already been resolved in the first stage of trial.[25] In *Atkins v. Commonwealth*, the defendant contested the bifurcated decision-making system per se, claiming that it violates his constitutional rights in preventing him from arguing residual doubt as to his guilt at sentencing.[26] The Virginia Supreme Court rejected this claim, holding that defendants cannot make residual doubt claims at sentencing.[27]

These formal legal regimes notwithstanding, the prevailing practice that has emerged among juries is to mitigate the punishment and impose life imprisonment, rather than the death penalty, when there are residual doubts as to a defendant's guilt.[28]

> How the law treats the issue of a jury's consideration of residual doubt is one thing . . . but how juries treat it is reasonably clear. Various studies have concluded that residual doubt about guilt is the single most persuasive aspect of a case leading to the imposition of a life sentence.[29]

For jurors, then, residual doubt has become a de facto mitigating factor, including in those states where the doctrine does not apply de jure.[30] For example, in Florida, 69 percent of jurors

---

24. *See* Pignatelli, *supra* note 10, at 312 ("While some states have given residual doubt a prominent role in their sentencing schemes, Virginia has declined to do so.").

25. 402 S.E.2d 196, 207 (Va. 1991); *see also* Frye v. Commonwealth, 345 S.E.2d 267, 283 (Va. 1986) (ruling that the defendant was not entitled to challenge the determination regarding his criminal responsibility at the sentencing hearing, for "the issue of guilt had been resolved in the first phase of the trial and could not properly be raised again in the penalty phase").

26. 534 S.E.2d 312, 314 (Va. 2000).

27. *Id.* at 315.

28. *See* Treadway, *supra* note 13, at 231–32 (noting that residual doubt acts as a de facto mitigating factor).

29. Bruce A. Antkowiak, *Judicial Nullification*, 38 CREIGHTON L. REV. 545, 582 (2005).

30. *See* Scott E. Sundby, *The Capital Jury and Empathy: The Problem of Worthy and Unworthy Victims*, 88 CORNELL L. REV. 343, 371 (2003) (suggesting that juries are impacted by their impression from the guilt stage of the trial at sentencing even without the formal application of the residual doubt doctrine); *see also* Stephen P. Garvey, *Aggravation and Mitigation in Capital Cases: What Do Jurors Think?*, 98 COLUM. L. REV. 1538, 1563 (1998) ("[T]he best thing a capital defendant can do to improve his chances of receiving a life sentence . . . is to raise doubt about his guilt."); Susan D. Rozelle, *The Principled Executioner: Capital Juries' Bias and the Benefits of True Bifurcation*, 38

who had decided on a life sentence admitted in retrospect that their choice of the reduced penalty had stemmed from residual doubt as to the defendant's guilt.[31] A similar pattern is apparent amongst jurors in Virginia, where the residual doubt doctrine also does not officially apply.[32] In a comprehensive study that examined 600 homicide cases in Georgia on a variety of parameters, it was found that residual doubt about guilt effectively influenced the severity of the penalty imposed and led to the commuting of death sentences to life sentences.[33]

In sum, juries over the years have integrated uncertainty regarding guilt into their sentencing calculations, turning residual doubt into a central mitigating factor for conversion of death sentences to life imprisonment.[34] The residual doubt doctrine can, thus, be construed as an instance of probabilistic decision making: On the sentencing plane, it facilitates correlativity between punishment severity and certainty of guilt. On the culpability level, it leads to the effective fragmentation of the category of conviction into two (informal) sub-categories of conviction on guilt beyond residual doubt and conviction on guilt beyond a reasonable doubt.

B. THE RECIDIVIST SENTENCING PREMIUM

Enhancement of sentences due to recidivism is another example of how probabilistic logic has infiltrated criminal proceedings. Repeat offenders are punished more harshly than first-time offenders.[35] In fact, a defendant's criminal record

---

ARIZ. ST. L.J. 769, 775 (2006) ("Residual doubt, or the jurors' lingering fear that the defendant may not be guilty after all, is the most potent mitigator in capital cases . . . .").

31. William S. Geimer & Jonathan Amsterdam, *Why Jurors Vote Life or Death: Operative Factors in Ten Florida Death Penalty Cases,* 15 AM. J. CRIM. L. 1, 28 (1988) (interviewing and examining jury members from capital trials in Florida).

32. *See* Pignatelli, *supra* note 10 ("[C]omments by capital case jurors in Virginia indicate that despite Virginia's prohibition on argument about residual doubt at sentencing, their own residual doubt impacted their decision to recommend a life sentence.").

33. *See* Arnold Barnett, *Some Distribution Patterns for the Georgia Death Sentence*, 18 U.C. DAVIS L. REV. 1327, 1343 (1985).

34. *See* Welsh S. White, *Effective Assistance of Counsel in Capital Cases: The Evolving Standard of Care,* 1993 U. ILL. L. REV. 323, 357 n.236 (claiming that residual doubt about a defendant's guilt is a major factor in a jury's decision to impose life sentences and may even amount to the "strongest possible mitigating evidence").

35. *See* C.Y. Cyrus Chu et al., *Punishing Repeat Offenders More Severely,* 20 INT'L REV. L. & ECON. 127, 127 (2000) (arguing that punishing repeat of-

(alongside offense severity) is the weightiest determinative factor in sentence gravity.[36] The imposition of a "recidivist premium" is entrenched in formal legal doctrine on the federal, state, and administrative levels.[37] The federal sentencing guidelines, for instance, use two sets of numerical scores to determine a convicted offender's sentence: offense level, which represents the seriousness of the crime, and criminal history. Under these guidelines, the richer an offender's criminal record, the more determinative its effect on the applicable range of punishment for the most recent offense.[38] State guidelines likewise incorporate recidivist provisions.[39] Noteworthy examples of recidivism-based penalty escalation statutes include the "Three-Strikes-and-You're-Out" laws adopted by various state jurisdictions.[40] Courts—both state and federal—have also effectively endorsed the policy of recidivist premiums. There are unequivocal indications that first-time offenders receive lighter sentences than repeat offenders, and that offenders with a long history of criminality are more severely punished than those with few prior

fenders more harshly than first-time offenders is a generally accepted practice and characteristic of the vast majority of penal codes and sentencing guidelines); *see also* Moshe Burnovski & Zvi Safra, *Deterrence Effects of Sequential Punishment Policies: Should Repeat Offenders Be More Severely Punished?*, 14 INT'L REV. L. & ECON. 341, 341 (1994); Julian V. Roberts, *The Role of Criminal Record in the Sentencing Process*, 22 CRIME & JUST. 303, 342 (1997).

36. Youngjae Lee, *Recidivism as Omission: A Relational Account,* 87 TEX. L. REV. 571, 571 (2009) ("In the United States, the most important determinant of punishment for a crime, other than the seriousness of the crime itself, is the offender's criminal history.").

37. David A. Dana, *Rethinking the Puzzle of Escalating Penalties for Repeat Offenders,* 110 YALE L.J. 733, 735 (2001) ("The principle is embedded in formal federal, state, and administrative codes . . . .").

38. U.S. SENTENCING GUIDELINES MANUAL ch. 4, pt. A (2010); *see also* Donna H. Lee, *Resuscitating Proportionality in Noncapital Criminal Sentencing*, 40 ARIZ. ST. L.J. 527, 572 (2008); Michael Edmund O'Neill, *Abraham's Legacy: An Empirical Assessment of (Nearly) First-Time Offenders in the Federal System*, 42 B.C. L. REV. 291, 305–06 (2001).

39. Richard S. Frase, *Is Guided Discretion Sufficient? Overview of State Sentencing Guidelines,* 44 ST. LOUIS U. L.J. 425, 434 (2000) ("[N]ewer guidelines systems are likely to incorporate recidivist and 'dangerous offender' provisions . . . .").

40. John F. Pfaff, *The Continued Vitality of Structured Sentencing Following* Blakely: *The Effectiveness of Voluntary Guidelines*, 54 UCLA L. REV. 235, 245 n.27 (2006) (specifying the twenty-five jurisdictions to adopt a three-strikes policy). For a survey of three-strikes laws, see Marc Mauer, *Why Are Tough on Crime Policies So Popular?*, 11 STAN. L. & POL'Y REV. 9 (1999). For a more elaborate discussion of recidivist statutes, see 39 AM. JUR. 2D *Habitual Criminals* §§ 1–55 (2008).

convictions.[41] In addition, the U.S. Supreme Court has held that recidivist premiums do not violate the double jeopardy prohibition.[42] Complementing the formal legal doctrine and case law are prosecutors' enforcement norms that also embrace the notion of increased sentences for recidivists.[43] The principle of enhancement for re-offenders has become so broadly accepted and popular "that it strikes most people as simple common sense."[44]

This widespread practical acceptance of recidivist premiums, notwithstanding the question of whether an offender's criminal history ought to affect her penalty for a current offense, has been the subject of heated theoretical debate dating back to Plato.[45] Utilitarian justifications for the recidivist premium include deterrence considerations that "dictate[] that a clear message be sent to society that repeated criminal behavior will aggravate the need for punishment with each recurrence."[46] It has been suggested in this context that individuals who have been convicted of one crime need enhanced penalties to be optimally deterred from re-offending, for by engaging in criminal behavior in the past, such individuals have revealed their proclivity for criminal activity.[47] Moreover, one must bear in mind that when offenders have been subjected to prior criminal punishment, the formal sanction is eroded, resulting in a weaker deterrent effect than for first-time offenders. Individuals with criminal records have lower opportunity costs; the marginal cost of the first years behind bars is lower than the marginal cost of subsequent imprisonment years, and the additional reputation costs entailed in a greater number of convic-

---

41. Dana, *supra* note 37.

42. Witte v. United States, 515 U.S. 389, 400 (1995) (holding use of a previous offense in the sentencing of a current offense does not constitute a double jeopardy violation).

43. Dana, *supra* note 37 (claiming that the principle of increased penalties due to recidivism has become an enforcement norm "of prosecutors and government officials at all levels of government"); *see also* A. Mitchell Polinsky & Steven Shavell, *On Offense History and the Theory of Deterrence*, 18 INT'L REV. L. & ECON. 305, 305 (1998) ("[T]he law often sanctions repeat offenders more severely than first-time offenders.").

44. Dana, *supra* note 37.

45. O'Neill, *supra* note 38, at 296 (arguing that the preoccupation with a defendant's criminal past has extensive theoretical foundations that can be traced back to Aristotle and Plato).

46. *See* U.S. SENTENCING GUIDELINES MANUAL, *supra* note 38.

47. Katherine J. Strandburg, *Deterrence and the Conviction of Innocents,* 35 CONN. L. REV. 1321, 1338 (2003) ("[T]hose who have previously committed crimes have revealed a preference for criminal activity . . . ."); *see also* Lee, *supra* note 36, at 573.

tions decrease as the number of convictions rises.[48] From a re-habilitation perspective, prior convictions are considered an important proxy for low rehabilitative potential.[49] Other conse-quentialist rationales for the recidivist doctrine relate to the need to preserve public trust in the justice system and to the damage that may result from the "revolving door" phenomenon, where offenders repeatedly exit and reenter jail in short peri-ods of time.[50] Retributivists have traditionally been more criti-cal of sentencing enhancement based on prior convictions, their common objection being that the punishment should fit the current crime and not past behavior.[51] However, even propo-nents of just desert have defended recidivist doctrines on the grounds that the moral culpability of past offenders is graver in light of their prior acts, for which they were punished and from which they failed to learn. Indeed, Andrew von Hirsch main-tains that the later sentences can bring to bear the full weight of the law because the offender has been forewarned of the ille-gitimacy of his behavior (by way of the previous sentence).[52]

It is my contention that alongside these various rationales, the recidivist premium can also be construed as a manifesta-tion of the link between certainty of guilt and severity of pun-ishment. There could be room to claim that the additional in-formation submitted post-conviction—regarding the defendant's prior convictions—updates and increases the like-lihood of her involvement in the latest alleged offense. For ex-ample, statistics show that a person who has committed past offenses is more likely to be involved in subsequent criminal ac-tivity.[53] Submission of information regarding past convictions

48. Strandburg, *supra* note 47 ("[T]he additional stigma due to additional convictions for similar offenses surely decreases rapidly as the number of con-victions rises.").

49. JULIAN V. ROBERTS, PUNISHING PERSISTENT OFFENDERS: EXPLORING COMMUNITY AND OFFENDER PERSPECTIVES 35–36 (2008) (discussing the sig-nificance of recidivism for estimating rehabilitation capacity).

50. *See* Markus Dirk Dubber, *The Unprincipled Punishment of Repeat Of-fenders: A Critique of California's Habitual Criminal Statute*, 43 STAN. L. REV. 193, 194 (1990) (terming this phenomenon "revolving door justice").

51. *See* Lee, *supra* note 36, at 575 (describing the argument that the se-verity of a crime is unrelated to the offender's personal criminal history).

52. ANDREW VON HIRSCH, DOING JUSTICE: THE CHOICE OF PUNISHMENTS 85 (1976).

53. *See* Alon Harel & Ariel Porat, *Aggregating Probabilities Across Cases: Criminal Responsibility for Unspecified Offenses*, 94 MINN. L. REV. 261, 280 (2009) ("It is the interdependence between the past offense and the present alleged offense that provides the grounds for conviction."); *see also* A. Rubin-stein, *An Optimal Conviction Policy for Offenses that May Have Been Commit-*

thus reinforces the first-stage convicting verdict and pushes the probability of guilt, which has already crossed the reasonable doubt threshold (as inferred from the fact of conviction) to a point that comes even closer to absolute certainty. The recidivist premium both reflects and is a consequence of the greater certainty level regarding a repeat offender's guilt and involvement in the current offense. This also explains why the recidivist premium for offenders with numerous prior convictions is more substantial than that imposed on repeat offenders with a more modest criminal record: the richer a convicted defendant's criminal record, the greater the added certainty of her culpability in the most recent offense. To conclude, recidivism doctrines can be understood as an expression of probabilistic decision making in criminal trials. On the sentencing plane, the recidivist premium is a context in which severity of punishment is enhanced to correlate with updated probability of guilt. On the conviction plane, the updating process effectively institutes another category of conviction in the epistemic space above the beyond-a-reasonable-doubt threshold—one that is in greater proximity to the endpoint of absolute certainty.

C. THE JURY TRIAL PENALTY

The imposition of harsher sentences on convicted defendants who chose to assert their constitutionally protected procedural rights—most notably, the right to trial by jury—has become routine practice in many American courtrooms.[54] Data collected nationwide over the past four decades indicate that when the offense-type variable is constant, bench trial sentences are, on average, more lenient than jury trial sentences.[55]

---

ted by Accident, APPLIED GAME THEORY 406, 407 (1979) (arguing that in a repeat game scenario, the most just outcome is reached by being lenient to offenders with a light criminal record but being harsh to repeat offenders). For an empirical examination of the link between exposure to information of prior convictions and probability of conviction on current charges, see Theodore Eisenberg & Valerie P. Hans, *Taking a Stand on Taking the Stand: The Effect of a Prior Criminal Record on the Decision to Testify and on Trial Outcomes* 30 (Cornell Legal Studies Research Article No. 07-012, Aug. 8, 2007), *available at* http://ssrn.com/abstract=998529.

54. Nancy J. King et al., *When Process Affects Punishment: Differences in Sentences After Guilty Plea, Bench Trial, and Jury Trial in Five Guidelines State*s, 105 COLUM. L. REV. 959, 962–63 (2005) (claiming that the practice of exchanging punishment concessions for waiver of procedural safeguards is widespread in the United States and elsewhere).

55. *See, e.g., id.* at 962 n.7 (surveying the literature on the trial penalty phenomenon); Gary D. LaFree, *Adversarial and Nonadversarial Justice: A Comparison of Guilty Pleas and Trials*, 23 CRIMINOLOGY 289, 300 (1985)

This discrepancy would imply that courts impose a "jury trial penalty" on defendants who invoke their right to a full jury trial.[56] The most prevalent justification for this phenomenon is that it is intended to provide defendants with an incentive to waive expensive procedural safeguards, so as to reduce the costs entailed in the administration of criminal justice.[57] Another justification for the jury trial penalty, most particularly imposed on those defendants who plead not guilty, relates to the lack of remorse that the insistence on innocence expresses.[58] Other explanations point to the greater public scrutiny of jury trials, which may discourage elected trial judges from imposing lenient sentences in such cases.[59]

These considerations notwithstanding, it can also be argued that the imposition of higher sentences in jury trials can be understood as an expression of the link between certainty of guilt and severity of punishment. The decision of twelve jurors to convict may embody a greater measure of certainty as to the defendant's guilt than the verdict of a lone judge. Not only is a

---

(showing sentence discrepancies between jury and bench trials across state courts); Lawrence P. Tiffany et al., *A Statistical Analysis of Sentencing in Federal Courts: Defendants Convicted After Trial 1967–1968*, 4 J. LEGAL STUD. 369, 379 (1975) (demonstrating that conviction by jury leads to more severe sentences than bench trial conviction); Thomas M. Uhlman & N. Darlene Walker, *"He Takes Some of My Time; I Take Some of His": An Analysis of Judicial Sentencing Patterns in Jury Cases*, 14 LAW & SOC'Y REV. 323, 335 (1980) (finding jury trial sentences harsher than bench trial sentences); Jeffery T. Ulmer & Mindy S. Bradley, *Variation in Trial Penalties Among Serious Violent Offenses*, 44 CRIMINOLOGY 631, 653–54 (2006).

56. *See* DAVID W. NEUBAUER, AMERICA'S COURTS & THE CRIMINAL JUSTICE SYSTEM 287 (1988) (claiming that the penalty reflects the philosophy that the defendant must "pay" for taking up the jury's time); Stephen J. Schulhofer, *Is Plea Bargaining Inevitable?*, 97 HARV. L. REV. 1037, 1062–63 (1984) (arguing that defendants waive procedural rights for leniency in punishment); Thomas Weigend, *Is the Criminal Process About Truth?: A German Perspective*, 26 HARV. J.L. & PUB. POL'Y 157, 167 (2003) (arguing that "defendants pay a heavy price for the 'second bite' of a jury trial if they are convicted"). For further discussion of the size of the penalty, see Ronald F. Wright, *Trial Distortion and the End of Innocence in Federal Criminal Justice*, 154 U. PA. L. REV. 79, 113 n.96 (2005) (noting that a guilty plea can sometimes cut a defendant's sentence in half).

57. King et al., *supra* note 54, at 964 (describing the efficiency rationale for the trial penalty and quoting one judge's famous remark in this regard: "He takes some of my time, I take some of his").

58. *Id.* at 963 n.12.

59. *Id.* at 964 ("[P]ublic scrutiny of sentences may be highest in cases that go to jury trial. This may contribute to the reluctance of elected trial judges and prosecutors to select and recommend lenient sentences after a jury has returned a guilty verdict.").

jury conviction based on the prosecution's ability to convince a greater number of fact-finders that the defendant is guilty as charged, but juries, it has been asserted, tend to a priori favor the defense on issues of reasonable doubt.[60] As Schulhofer has stated, "a bench trial is seen as an option that reduces the chances for acquittal . . . ."[61] It can thus be claimed that the relatively higher degree of certainty as to a defendant's guilt in cases where she is found guilty by a jury—as compared to bench trials—is another contributing factor in the discrepancy in sentences. The relative gravity of sentences in jury trials is reflective of the elevated certainty as to guilt in the wake of a jury verdict, whereas the relatively lenient sentences in bench trials are due, at least in part, to a lower degree of epistemic confidence in the conviction. Analogous claims can be made with respect to the exercise of other procedural rights, including the right to cross examine witnesses or the right to assistance of counsel.

D. THE CRIMINAL-CIVIL INTERPLAY AND PLEA BARGAINS

Correlativity between the certainty of guilt and the severity of punishment, as well as the fragmentation of the category of conviction into sub classes, can also be detected outside of the criminal trial arena—as exemplified by the interplay between criminal trials and civil trials and by the practice of plea bargaining.

Starting with the former, as suggested to me by David Sklansky, the interplay between civil and criminal trials can be construed as another way in which our justice system accommodates probabilistic decision making. When a defendant is acquitted in criminal trial, but is then held liable in a tort suit on parallel grounds, the civil judgment effectively functions as a halfway criminal conviction—reflecting satisfaction of the less demanding civil standard of proof as to the involvement in the alleged conduct, and leading to the imposition of a partial sanction in the form of monetary compensation. An illuminating example of this function of the civil judgment is the O.J.

---

60. Schulhofer, *supra* note 56, at 1063 ("[D]efenders . . . believe that juries are often more favorable to the defense on reasonable doubt questions . . . .").

61. *Id.* (arguing that bench trials reduce the probability of acquittal, while increasing the chance for leniency in sentencing); *see also* Weigend, *supra* note 56, at 164–67 (claiming that a number of factors unrelated to actual guilt, such as emotion, instinct, and the lack of a written opinion can lead a jury to acquit a defendant when a judge would be less likely to do so).

Simpson case.[62] Following Simpson's acquittal in a criminal trial for the murders of Ron Goldman and Nicole Brown Simpson, the Goldman and Brown families filed wrongful death and survival suits in civil court. On February 4, 1997, a civil jury found Simpson liable for the wrongful death of Ron Goldman.[63] Fred Goldman, the victim's father, stated repeatedly that the civil suit was not financially motivated, but rather that he had filed it so that the court would establish through the civil trial what the criminal trial had failed to substantiate—namely, that O.J. Simpson was responsible for his son's death.[64] This is also how the public perceived the role of the civil judgment in the O.J. Simpson case.[65]

The practice of plea bargaining is another arena of probabilistic decision making.[66] Plea bargains have come to dominate the American criminal justice system.[67] About 95 percent of individuals indicted for felony crimes in the United States resolve their cases by way of a guilty plea.[68] Plea bargains are struck in

---

62. For the criminal trial, see Transcript of Delivery of Verdict, People v. Simpson, 1995 WL 704381 (Cal. Super. Ct. 1995) (No. BA097211). For the civil trial, see Amended Joint Trial Statement, Rufo v. Simpson, 1996 WL 530984 (Cal. Super. Ct. 1996) (Nos. SC031947, SC036340, and SC036876).

63. Laura Hock, *What's in a Name? Fred Goldman's Quest to Acquire O.J. Simpson's Right of Publicity and the Suit's Implications for Celebrities*, 35 PEPP. L. REV. 347, 352–53 (2008).

64. Allen M. Linden, *Acceptance Speech: The Fleming Award 2004*, 33 PEPP. L. REV. 929, 931 (2006).

65. In fact, Goldman announced that he would renounce the financial award in the event that Simpson admits the crime. Max Bolstad, *Learning from Japan: The Case for Increased Use of Apology in Mediation*, 48 CLEV. ST. L. REV. 545, 567 (2000).

66. *See* Henrik Lando, *The Size of the Sanction Should Depend on the Weight of the Evidence*, 1 REV. L. & ECON. 277, 286 (2005) ("[L]egal systems that rely on plea bargaining (mainly the American) are likely to produce sanctions that are weighted by the probability of conviction . . . .").

67. *See* Jason Mazzone, *The Waiver Paradox*, 97 NW. U. L. REV. 801, 801 (2003) (claiming that the right to a jury trial and the right against self-incrimination are routinely bargained away in the criminal arena, in exchange for sentence reduction).

68. Michael M. O'Hear, *Plea Bargaining and Procedural Justice*, 42 GA. L. REV. 407, 409 (2008) ("[In felony cases filed in May 2002 in the nation's seventy-five largest counties g]uilty pleas accounted for 95% of the estimated 31,772 convictions obtained within 1 year of arrest . . . ." (citing BUREAU OF JUSTICE STATISTICS, U.S. DEP'T OF JUSTICE, FELONY DEFENDANTS IN LARGE URBAN COUNTIES, 2002, at 28 (2006))); *see also* Jacqueline E. Ross, *The Entrenched Position of Plea Bargaining in United States Legal Practice*, 54 AM. J. COMP. L. 717, 717 (2006) ("In the criminal justice systems of the 50 states, over 95 percent of all criminal cases are disposed of without a trial, through the entry of a guilty plea. In the federal system the percentage of bargained-for convictions

the shadow of the criminal trial and are reflective of the likelihood of conviction in court (and of the post conviction sentence).[69] The likelihood of conviction, in turn, is a function of the strength of the prosecution's case. In this way, plea bargains calibrate sentences to the probative value of the incriminating evidence and form a link between the severities of the (negotiated) sentences imposed upon defendants and the level of proof gathered against them. Obviously, the view that plea bargaining is an instance of partial guilt being translated into partial sentencing requires a skeptical leap with regard to the probative weight of the confession underlying it. But, the fact that the courts or juries do not independently evaluate the actual evidence of the crime could be interpreted as providing the necessary indication that plea bargain convictions do not substantiate guilt beyond a reasonable doubt. This view was also expressed by Albert W. Alschuler, who claimed that "[t]he plea bargaining system effectively substitutes a concept of partial guilt for the requirement of proof of guilt beyond a reasonable doubt."[70]

Indeed, plea bargains are not perfect instances of probabilistic decision making; the correlativity between severity of punishment and probability of guilt generated by plea bargains is far from precise. Plea bargain outcomes are affected by a host of exogenous factors other than the expected outcomes at trial—including agency costs, the defendant's attitude to risk, information asymmetries between the prosecution and defense, the quality of legal counsel, and various cognitive biases.[71] But

---

is even higher.").

69. For an elaborate description of the shadow-of-trial plea bargaining model, see Stephanos Bibas, *Plea Bargaining Outside the Shadow of Trial*, 117 HARV. L. REV. 2464, 2465–67 (2004).

70. Albert W. Alschuler, *Straining at Gnats and Swallowing Camels: The Selective Morality of Professor Bibas*, 88 CORNELL L. REV. 1412, 1413–14 (2003).

71. *See id.* at 1413–15 (discussing how structural distortions emanating from agency costs, attorney competence, workloads resources, sentencing and bail rules, information gaps, psychological biases, and socio-economic variables distort plea outcomes); Frank H. Easterbrook, *Criminal Procedure as a Market System,* 12 J. LEGAL STUD. 289, 309 (discussing the effects of agency costs and conflicts of interest between the defendant and her attorney); Robert E. Scott & William J. Stuntz, *Plea Bargaining as Contract*, 101 YALE L.J. 1909, 1925 (1992) (discussing the effects of attitude toward risk). For a parallel discussion in the context of the civil sphere, see Carrie Menkel-Meadow, *Whose Dispute Is It Anyway?: A Philosophical and Democratic Defense of Settlement (In Some Cases)*, 83 GEO. L.J. 2663, 2663–71 (1995).

Plea bargaining outcomes may deviate from the shadow of the criminal trial for yet another reason: the prosecution's primary interest in reaching

despite the absence of a perfect correlativity between the certainty of guilt and the negotiated sentence, plea bargains, by and large, adjust sentencing to the level of proof regarding culpability.[72] Due to their prominent role,[73] plea bargains have entrenched a sweeping regime of probabilistic decision making in the criminal justice system.

Thus far, I have sought to ground the claim that a probabilistic decision making model is consistent with a wide range of current criminal law doctrines and practices that tie severity of punishment to certainty of guilt and that effectively disintegrate the category of criminal conviction into a plurality of subcategories. As shown above, these probabilistic practices are prevalent in the framework of the criminal court, in the interplay between criminal and civil trials, as well as in the plea bargaining arena. They relate both to the epistemic space stretching above the reasonable doubt standard (as in the case of residual doubt) as well as to the space below it (as in plea bargaining).[74] In light of the dissonance between the threshold ideal and the reality of the criminal justice system, there is room to contest the axiomatic nature of this regime. This Article now turns to formulate the normative case for probabilistic decision making in criminal trials. The claim in this regard will be that deterrence considerations, expressive issues, and ex post fairness requirements support the widespread and explicit

---

plea agreements is the reduction of enforcement costs—i.e., the "release" from having to continue investing resources in the gathering of more proof once the plea bargain has been struck. Without the plea bargain, enforcement officials may be induced to persist in their search for additional evidence to be presented at trial, and thus the probative weight of the incriminating evidence could end up being greater than at the point at which the plea bargain is made.

72. Another reason why plea bargains are not perfect examples of probabilistic decision making emanates from the fact that they are effectively struck outside the criminal trial arena. Although these agreements are subject to the court's formal stamp of approval, from a conceptual perspective, they belong to the "contract world," reflecting a contractual ordering of the criminal case. Yet, given that the vast majority of criminal cases are concluded in various types of plea bargains, the practice cannot be regarded as merely anecdotal. *See* Strandburg, *supra* note 47, at 1331 (describing plea bargaining as the most widely used method of criminal conviction).

73. *See* Daniel P. Blank, *Plea Bargain Waivers Reconsidered: A Legal Pragmatist's Guide to Loss, Abandonment and Alienation*, 68 FORDHAM L. REV. 2011, 2016 (2000) (noting that "plea bargaining has long been a fixture of criminal adjudication in the United States").

74. *See* Patricia M. Wald, *Guilty Beyond a Reasonable Doubt: A Norm Gives Way to the Numbers*, 1993 U. CHI. LEGAL F. 101, 108 (claiming that plea bargaining effectively lower the standard of proof).

exercise of probabilistic decision making in the criminal arena, beyond the parameters of specific legal doctrines and practices.

E. THE NORMATIVE UNDERPINNINGS OF THE PROBABILISTIC MODEL

The question of the normative desirability of probabilistic decision making in the criminal arena can be understood as an offshoot of the proverbial tort law debate over causal apportionment of liability and probabilistic recovery. John Coons's 1964 article challenged civil procedure's winner-takes-all decision rule and inquired into remedy-splitting based on certainty of liability.[75] Almost two decades passed before the issue of probabilistic recoveries was raised again with the publication of Kaye's article in 1982.[76] There, Kaye asserted that the civil all-or-nothing decision rule was preferable to an expected value rule.[77] The former system, he argued, minimizes the overall costs of error, whereas the latter enhances the social costs of error.[78] The central development in the literature on probabilistic remedies in tort law emerged in subsequent years with a string of publications that abandoned the ex post error minimization approach and turned instead to the issue of probabilistic liability from an ex ante deterrence perspective.[79] These studies

---

75.   John E. Coons, *Approaches to Court Imposed Compromise—The Uses of Doubt and Reason*, 58 NW. U. L. REV. 750, 754–55 (1964). Coons defined four classes of cases in which courts should be allowed residual authority to issue compromise verdicts and to split the remedy between the parties, in deviation from the binary regime of full remedy awards: compromise of doubt; compromise of policy; compromise of discretion; and compromise due to community interests. *Id.* at 754. The pertinent category for our purposes is compromise of doubt cases, in which Coons advocated remedy-splitting due to factual or legal doubts regarding legal liability. *Id.* at 755–60. In this context, he allowed only a narrow opening (which was widened by subsequent authors), restricting remedy-splitting to a 50%-50% ratio and to instances in which the court finds the parties' versions equally likely. *Id.* at 778.

Nonetheless, Coons' pioneering approach was broader than later work by other authors, in that it forged a link between the allocation of the remedy and the level of certainty of liability not only on the factual front but also on the legal front. Only in 2001 did Michael Abramowicz revisit the possibility of allocating the remedy due to uncertainty on legal issues. *See* Michael Abramowicz, *A Compromise Approach to Compromise Verdicts*, 89 CALIF. L. REV. 231, 299–300 (2001).

76.   David Kaye, *The Limits of the Preponderance of the Evidence Standard: Justifiably Naked Statistical Evidence and Multiple Causation*, 1982 AM. BAR FOUND. RES. J. 487.

77.   *Id. passim.*

78.   *Id.* at 491.

79.   Examples of works making the claim that proportional liability would

showed that the shift to a regime of probabilistic remedies would advance deterrence objectives in instances of systematic failure in proving causation.[80]

Contrary to the decades-long scholarly interest in the civil context, the issue of probabilistic decision making in the criminal sphere was essentially overlooked by researchers. The first sign of change came only in 2005, with the publication of a novel article by Henrik Lando, in which he showed that for incarcerable offenses, graduating sanctions with the probability of guilt to reflect remaining doubt—even once the beyond-a-reasonable-doubt level of certainty has been attained—would

---

promote deterrence include ARIEL PORAT & ALEX STEIN, TORT LIABILITY UNDER UNCERTAINTY 57–83 (2001); Saul Levmore, *Probabilistic Recoveries, Restitution, and Recurring Wrongs*, 19 J. LEGAL STUD. 691 (1992); Glen O. Robinson, *Multiple Causation in Tort Law: Reflections on the DES Cases*, 68 VA. L. REV. 713 (1982); David Rosenberg, *The Causal Connection in Mass Exposure Cases: A "Public Law" Vision of the Tort System*, 97 HARV. L. REV. 849 (1984); and Steven Shavell, *Uncertainty over Causation and the Determination of Civil Liability*, 28 J.L. & ECON. 587 (1985). Counter-criticism can be found in Charles Nesson's seminal article. Charles Nesson, *The Evidence or the Event? On Judicial Proof and the Acceptability of Verdicts*, 98 HARV. L. REV. 1357 (1985).

80. *See* Mario J. Rizzo & Frank S. Arnold, *Causal Apportionment in the Law of Torts: An Economic Theory*, 80 COLUM. L. REV. 1399 (1980). The failures discussed in the literature can be divided into two central types of cases. The first category encompasses multiple causation cases, in which there is structural uncertainty as to whether, in the concrete circumstances, the damage was caused as a result of the defendant's behavior or due to exogenous background factors. A typical example is the factory that emits pollution known to produce 20 percent of the cancer morbidity in a given area, but it cannot be proven in each individual case that the factory's polluting is the cause of illness. In such circumstances, holding the factory liable for 20 percent of the damage for each plaintiff will ensure internalization of the damage caused due to the pollution, and will generate a set of optimal incentives for potential wrongdoers.

The second class of cases relates to mass torts and comprises instances of uncertainty with regard to the wrongdoer's identity: singling out from among a number of potential wrongdoers the actual wrongdoer who caused the plaintiff's injury. These are situations in which it is not possible to determine by a preponderance of the evidence that the particular defendant caused the damage to the plaintiff, but it can be proven that he belongs to the group that includes the actual wrongdoer. In this context, too, the argument has been made that a shift to a Market Share Liability decision regime, under which each of the potential wrongdoers would bear a proportion of the damage relative to its share of the market (which reflects the probability of individual responsibility for the specific plaintiff's damage), would enable optimal deterrence and ameliorate the inherent failures of the all-or-nothing decision regime, where each and every wrongdoer is exempt from legal liability. *See* PORAT & STEIN, *supra* note 79, at 58.

promote deterrence objectives.[81] The normative discussion in this Article follows in Lando's path and expands it to the study of a full-fledged probabilistic regime.[82] Unlike Lando's focus on probabilistic sentencing and on the deterrence-based case for this practice, the analysis presented in this Article will examine the question of probabilistic decision making, both in the context of sentencing and in the context of the configuration of guilt, enhancing the spectrum of normative perspectives to include expressive, retributive, and political legitimacy considerations. Moreover, the discussion will not be limited to challenging the uniformity of punishment in the epistemic space above the beyond-a-reasonable-doubt threshold. Rather, it will challenge the threshold model in its entirety, addressing the possibility of correcting systematic evidentiary failures by applying a probabilistic sentencing model in the epistemic expanse stretching below the beyond-a-reasonable-doubt point.

---

81.   Lando, *supra* note 66, at 281. Another milestone was the 2009 Harel & Porat article, which applies the conjunction principle to cases in which multiple unrelated charges are made against a defendant. Harel & Porat, *supra* note 53 *passim*. Under the current legal regime, if a defendant is accused of committing armed robbery and murder and, at the end of the day, the court concludes a 0.9 extent of certainty of his guilt in each of the crimes, then assuming that the reasonable doubt standard is quantified as a 0.95 extent of certainty, the defendant will be acquitted (0.9<0.95). Harel & Porat's claim is that in the described scenario, the probability that the defendant did not commit either of the two offenses is 0.1*0.1=0.01. That is to say, the certainty that the defendant committed at least one of the two crimes (without us knowing which one) is 1-0.01=0.99 (0.99>0.95). Thus, according to Harel & Porat, the defendant should be convicted for one offense and receive, at least for one of the possibilities, the penalty for the lesser offense.

In contrast to Harel & Porat, my analysis will not restrict itself to cases of conjunction or multiple charges, but will instead consider the implementation of probabilistic decisions in the general class of cases, including instances in which a defendant is accused of only one criminal offense.

82.   Lando's model serves as a good starting point for understanding the deterrence advantage associated with probabilistic sentencing. *See* Lando, *supra* note 66 *passim*. However, the picture he presents is only partial, for the Lando model derives the deterrent effect of each unit of punishment solely from certainty of guilt, thereby entailing an additional, implicit conclusion: namely, that punishment should be imposed only at the highest measures of certainty. *Id.* at 277–78. Thus, while the Lando model justifies correlating punishment with probability of guilt, in focusing exclusively on the deterrent effect of certainty of guilt it constricts the epistemic space of conviction and punishment to only the highest levels of certainty. In this sense, and in complete contradiction to its original objective, the Lando model may actually support the binary all-or-nothing regime.

II. THE PROBABILISTIC MODEL FROM A DETERRENCE PERSPECTIVE

A. THE DETERRENCE-BASED CASE FOR PROBABILISTIC SENTENCING

Similar to the tort law context, where systematic under-deterrence serves as a central argument for the implementation of market share liability (and other forms of probabilistic recovery), in the criminal law context as well, deterrence presents a strong case for probabilistic sentencing. Elsewhere,[83] I have developed the economic model supporting the deterrence-based argument for probabilistic sentencing, which is grounded upon three underpinnings: (1) the positive correlation between certainty of guilt and the deterrent effect of punishment (to which Lando pointed);[84] (2) risk-seeking tendencies of potential offenders; and (3) the adverse effect of false acquittals on deterrence. Since deterrence is not the main focus of this current discussion, I will discuss only the general contours of these foundations of the deterrence-based case for probabilistic sentencing, starting with the causal connection between the deterrent effect of punishment and the probability of guilt.

In order for the criminal system to advance deterrence goals, criminal verdicts must reflect the factual truth. Random allocation of convictions and acquittals will not further deterrence, for the deterrent value of the criminal sanction is determined by whether it is imposed on the factually guilty or the innocent. As Lando pointed out, just as punishment of the factually innocent yields a lower deterrent effect than punishment of the factually guilty, punishment of defendants whose probability of guilt is low yields a lower deterrent effect than identical punishment imposed upon defendants whose certainty of guilt is high.[85] Since deterrence effectiveness lessens with diminishing probability of guilt, correlating the magnitude of the sanction with the certitude of guilt will be expected to yield more deterrence per given social expenditure on punishment. A move to a sliding scale punishment regime—based on certainty

---

83. Talia Fisher, *Constitutionalism and the Criminal Law: Rethinking Criminal Trial Bifurcation*, 61 U. TORONTO L.J. 811, 819–30 (2011).

84. *See* Lando, *supra* note 66, at 282.

85. *Id.* Lando discusses the differential deterrent value of imprisonment years, as a function of the defendant's certainty of guilt. *Id.* at 278.

of guilt—would thus increase deterrence utility per sanctioning cost (per given level of social expenditure on punishment).[86]

The positive link between the deterrent effect of punishment and the certainty of the defendant's guilt is imperative for understanding the deterrence advantage associated with probabilistic sentencing. Lando's argument for correlating punishment severity with certainty of guilt is based exclusively upon this causal connection. In my opinion, however, the positive association between the certainty of guilt and the deterrent effect of punishment does not offer a complete justification for such a sentencing regime (even from a deterrence perspective). In fact, contrary to Lando's initial intention, this argument—when standing alone—may actually support the threshold model. For, deriving the deterrent effect of each unit of punishment (incarceration year) solely from certainty of guilt would seem to imply that punishment should be imposed only at the highest measures of certainty, thereby actually reinforcing the binary all-or-nothing regime, where conviction and punishment are limited to a narrow epistemic range of maximal levels of certainty.[87]

---

86. In simpler terms, punishing more severely those with a higher probability of guilt than those with a lower probability of guilt can be justified on the grounds that the greater the certainty of the defendant's guilt, the lesser the concern of "wasting" punishment resources while obtaining a weaker deterrence effect, and vice versa.

87. In other words, the conclusions that arise from Lando's model are double layered. The first layer, which Lando explicitly endorses, relates to the promotion of deterrence goals by deviating from the notion of fixed punishment and correlating severity of punishment with measure of certainty, all things being equal. Lando, *supra* note 66, at 278. This layer of the model is a foundational component in the constitution of a probabilistic punishment regime, but it constitutes only half the picture. The second layer implicit in the Lando model relates to the restriction of the ambit of punishment to the highest levels of certainty on the evidentiary spectrum. This layer can be derived from Lando's model, since only at the highest levels of certainty regarding guilt is there a minimization of the probability of the imprisonment years being "wasted" on innocents, and since this is the only variable that Lando's model considers as impacting deterrence effectiveness. *Id.*

This second prong of the model is problematic, in my opinion, in that it ignores additional variables that affect the deterrent utility of imprisonment years, such as the defendant's attitude to risk and the deterrence cost of wrongful acquittals. Thus, Lando's model assumes risk-neutrality on behalf of potential criminals. However, prospective offenders may exhibit risk-seeking tendencies in light of a diminishing marginal cost to each additional year of incarceration. Another explicit assumption underlying Lando's model is that the different types of evidence evidence establishing the higher probability of guilt and evidence substantiating a lower probability of guilt are equally likely to appear in court following the commission of the crime. *Id.* at 280. However, in light of increasing marginal costs of evidence gathering, there is room to

This conclusion, implicitly arising out of the Lando model, regarding the deterrence advantage of restricting the ambit of punishment to the highest levels of certainty on the evidentiary spectrum, ought to be qualified. Two additional variables which impact the deterrent utility of imprisonment years should also be taken into account: the defendant's attitude to risk and the deterrence cost of wrongful acquittals. When these two variables are taken together with the positive link between the deterrent effect of punishment and probability of guilt, to which Lando pointed, they collectively substantiate the deterrence-based claim for probabilistic sentencing. Starting with attitude to risk, prospective offenders may exhibit risk-seeking tendencies in light of a diminishing marginal cost to each additional year of incarceration.[88] The reasons for the decrease in the marginal cost of additional imprisonment years are varied, and include the postulation that over time inmates grow accustomed to the prison environment, making each year more tolerable than the year preceding it.[89] Under conditions of risk-seeking tendencies and decreasing marginal cost of incarcera-

contest this assumption. Evidence establishing a higher probability of guilt is less likely to appear in court, and this may lead to a high incidence of false acquittals under a regime that places maximal evidentiary requirements as a precondition for conviction. Incorporating risk-seeking tendencies and false acquittals into the equation facilitates the expansion of the epistemic space for conviction and the imposition of criminal punishment at lower levels of certainty (i.e., the lowering of the evidentiary threshold, perhaps even to levels falling below the reasonable doubt standard).

88. *See* Gary S. Becker, *Crime and Punishment: An Economic Approach*, 76 J. POL. ECON. 169, 178 (1968) (claiming that offenders are deterred more by the certainty of conviction than by the severity of the sanction); William Spelman, *The Severity of Intermediate Sanctions*, 32 J. RES. CRIME & DELINQ. 107, 113 (1995) (surveying empirical evidence suggesting that offenders regard a five-year term as only twice to four times as severe as a one-year term and discussing some of the reasons for the decrease in the marginal cost (disutility) of prison years).

89. Additional considerations include the lower degree of certainty attributed to the latter part of the prison term being served, due to the increasing likelihood of pardon or early release on account of illness over time. *See* Alon Harel & Uzi Segal, *Criminal Law and Behavioral Law and Economics: Observations on the Neglected Role of Uncertainty in Deterring Crime*, 1 AM. L. ECON. REV. 276, 297 (1999) (discussing some of the reasons why defendants prefer uncertain sentences over uncertain probability of detection). For further discussion, see generally A. Mitchell Polinsky & Steven Shavell, *On the Disutility and Discounting of Imprisonment and the Theory of Deterrence*, 28 J. LEGAL STUD. 1 (1999) (asking whether deterrence is enhanced more by increasing the length of imprisonment or raising the likelihood of imposing imprisonment and calculating the optimal combination of severity and probability of imprisonment).

tion years, the probability of conviction has a greater marginal deterrence effect than punishment severity[90]—so a low probability of conviction coupled with a proportionately higher sanction would yield less deterrence than a higher probability of conviction coupled with a proportionately lower sanction. In such situations, in which the disutility to the defendant increases at a sub-proportional rate to the increase in length of imprisonment, punishment should not be limited to the highest evidentiary threshold (to maximal punishment at maximal levels of certainty).

The effect of the incidence of false acquittals may also justify, from a deterrence perspective, the expansion of the epistemic space for conviction and punishment to sub maximal evidentiary levels. In light of increasing marginal costs associated with the gathering of evidence,[91] evidence establishing a higher probability of guilt is less likely to appear in court, and this may lead to a higher incidence of false acquittals under a regime that places maximal evidentiary requirements as a precondition for conviction. The force of this effect is a product of the distribution of wrongful convictions and acquittals at different epistemic levels and is likely to vary across different classes of cases. Either way, inserting the false acquittal variable into the equation allows us to identify situations in which optimal deterrence is achieved by lowering the standard of proof to enable conviction and punishment at sub maximal levels of certainty of guilt (alongside maximal punishment at the higher levels of certainty). These situations can be regarded as the criminal counterpart to the abovementioned tort cases, in

90. *See* Michael K. Block & Robert C. Lind, *An Economic Analysis of Crime Punishable by Imprisonment*, 4 J. LEGAL STUD. 479, 483 (1975) (noting that an increase in the certainty of conviction has a greater deterrence capacity than an identical increase in punishment severity).

91. *See*, e.g., Talia Fisher, *The Boundaries of Plea Bargaining: Negotiating the Standard of Proof*, 97 J. CRIM. L. & CRIMINOLOGY 943, 950–51 (2007) ("One can posit a situation where the task of proving the final X percent of the prosecution's case requires a vast investment in resources on its part, such as the monetary cost of obtaining evidence from out-of-state witnesses, the emotional price paid by child witnesses, or the cost of revealing evidence where the prosecution wants to preserve the cover of police agents. The prosecution may regard this evidence as crucial for proving its case 'beyond a reasonable doubt', but find it unnecessary when a move to a lesser standard of proof has been made."); Harel & Porat, *supra* note 53 at 291 ("It is typically much harder—and more costly—to collect the tenth item of evidence than the ninth item, the eighth item, and so on.").

which structural failures in proving causation justify a shift to a probabilistic-remedies regime.[92]

The fundamental explanation for why shifting to a probabilistic punishment regime would promote deterrence can thus be summed up as follows: due to the varying deterrence effect of incarceration, as a function of the variance in probability of guilt, there is cause to bind certainty of guilt to severity of punishment. This is the deterrence-based rationale for varying sanctions for the same behavior based on certainty of guilt. Yet it should not be inferred from this deterrence advantage that conviction and criminal sanctions should be limited to only the maximal levels of certainty of guilt. When risk-seeking tendencies and wrongful acquittals are factored in, they may justify the lowering of the standard of proof necessary for conviction. The combined effect of the abovementioned variables—namely, the utility in terms of deterrence associated with varying the size of punishment with the probability of guilt, as well as the costs in deterrence associated with false acquittals and with the effects of risk-seeking tendencies—leads to the deterrence advantage of a probabilistic sentencing regime. These factors stand at the heart of the deterrence-based case which can be made for correlating severity of punishment with ex post certainty of guilt, while reducing the standard of proof for conviction to sub-maximal levels.

B. ERROR-COST QUALIFICATIONS

In the previous Section I have argued that the uniformity of punishment in the epistemic space above the beyond-a-reasonable-doubt-standard and the lack of punishment at the levels of proof below that standard entail a deterrence cost, which could be avoided under a probabilistic-sentencing regime. But as exemplified by Kaye's argument, formulated against probabilistic remedies in the civil case context,[93] the ex ante deterrence perspective is not the sole consideration, even from an efficiency perspective, and can be qualified by ex post error- cost criteria.[94] In fact, in the criminal context of probabilistic sentencing, the issue of error costs may constitute an even greater constraint on deterrence-based considerations, as compared to the civil scenario of probabilistic recoveries.

---

92. *See supra* Part I.E.
93. *See* Kaye, *supra* note 76.
94. *See id.* at 496–503.

The social costs of the two types of errors that occur in the framework of criminal proceedings—wrongful convictions and wrongful acquittals—are not commensurate.[95] The social costs of wrongful conviction are regarded as significantly greater than those associated with false acquittal. Minimization of the aggregate social costs of error in criminal proceeding thus entails lowering the incidence of false convictions, even by way of increasing the prevalence of false acquittals.[96] The threshold model is compatible with this utilitarian calculus, for it allocates the risk of error between the defense and prosecution in a way that promotes error in favor of defendants—which is considered less costly to society—at the expense of error in favor of the prosecution—which entails a more substantial social cost.[97] A probabilistic-decision regime would function in the reverse manner: the reduction in false acquittal errors would be

---

95. *See, e.g.*, Frederick Schauer & Richard Zeckhauser, *On the Degree of Confidence for Adverse Decisions,* 25 J. LEGAL STUD. 27, 34 (1996).

96. Alex Stein, *The Refoundation of Evidence Law*, 9 CAN. J.L. & JURIS. 279, 324 (1996) (the risks of error must be allocated between the defense and prosecution so as to reflect the disutility ratio between wrongful conviction and wrongful acquittal).

For further discussion of the cost-minimization approach to evidence law, see Richard A. Posner, *An Economic Approach to Legal Procedure and Judicial Administration*, 2 J. LEGAL STUD. 399, 408–17 (1973); Schauer & Zeckhauser, *supra* note 95, at 33–41. Assuming, for example, that the social cost of a false conviction is nine times greater than the cost of false acquittal, the evidentiary threshold for conviction should be set at a 90 percent certainty level. Such a standard of proof would be aimed at reducing the likelihood of erroneous conviction, by compromising on the certainty of the innocence of acquitted defendants proportionate to the 9:1 disutility ratio between wrongful convictions and acquittals. *See also* David M. Appel, Note, *Attorney Disbarment Proceedings and the Standard of Proof*, 24 HOFSTRA L. REV. 275, 277 (1995) (stating that the risk of error in the administration of criminal justice is borne predominantly by the prosecution).

97. Stein, *supra* note 96; *see also* Schauer & Zeckhauser, *supra* note 95. Assuming that the beyond-a-reasonable-doubt standard of proof is set at a 90 percent certainty level (probability of 0.9), the social harm inflicted by erroneously convicting an innocent defendant is considered to be about nine times costlier than the social costs of wrongful acquittal. For further discussion on the desirability of quantifying the reasonable doubt standard, see Henry A. Diamond, Note, *Reasonable Doubt: To Define, or Not to Define*, 90 COLUM. L. REV. 1716 (1990) (arguing that jury instructions defining reasonable doubt should always be given); *see also* Peter Tillers & Jonathan Gottfried, 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 L. PROBABILITY & RISK 135, 140–41 (2006) (arguing that the usual reasons given for the unquantifiability of reasonable doubt are unsatisfactory, with the recent case of *Copeland* serving as a reminder that there are strong considerations in favor of quantifying at least some standards of persuasion).

achieved by increasing the incidence of false conviction, with the likely general outcome being an increase in the aggregate social cost of error in judicial decisions. This increase in error might outweigh the marginal deterrence utility that would result from a move to a probabilistic regime. Therefore, even in situations in which the threshold regime is inefficient from an ex ante deterrence perspective, it may nonetheless lead to more efficient outcomes ex post.

Incorporating the ex post error-costs into the equation may, indeed, constrict the probabilistic regime's scope of operation. Nonetheless, in my opinion, the mere existence of this effect does not mandate a general sweeping preference of the threshold model. One reason is that the social cost of wrongful conviction is not exogenous and constant. Rather, it is a function of the severity of the accompanying punishment: the social cost of a wrongful conviction that results in a sentence of one day in prison is not equivalent to the cost of a wrongful conviction resulting in life imprisonment.[98] Under the probabilistic model, the cost of wrongful conviction at the lower epistemic levels is relatively low, for the punishment imposed is only partial. This weakens the force of the ex post error-costs consideration and must be taken into account when calculating the standard of proof for conviction.[99] Moreover, even assuming that error-cost minimization is possible only under the threshold model, there is still the matter of the accompanying price in deterrence. The deterrence costs of wrongful acquittals cannot be ignored when determining the desirable decision regime, and when they are extensive, there is no justification for a priori preference of the binary model. Shavell has made a similar assertion in the civil setting, claiming that "it is a mistake to take error-cost minimization as the social goal, and a mistake which explains such anomalous implications as the recommended use of the more-probable-than-not threshold even

---

    98. This runs contrary to the existing literature on the criminal standard of proof and the allocation of risk of error in criminal proceedings, which tends to ignore the differential social costs of wrongful convictions and implicitly assumes a fixed social cost to such errors. *See, e.g.,* Posner, *supra* note 96, at 408–15.

    99. It should be noted that under conditions of diminishing marginal cost of imprisonment years, upon which the deterrence argument rests, *see supra* Part II.A, a shorter sentence would indeed mitigate error costs, but it would mitigate the costs in a disproportionate manner to the number of years by which punishment is shortened. Put differently, the mitigation of error costs under the probabilistic model is qualified by the role of diminishing disutility of additional imprisonment years.

where it would result in defendants' always escaping liability for harm done."[100]

III. EXPRESSIVE DIMENSIONS OF THE PROBABILISTIC MODEL

The probabilistic model presents an interesting test case when viewed through the expressive prism. In what follows, I will contend with the ways in which expressive considerations emerge in the context of probabilistic decision making. I will start out by pointing to the hypothetical expressive-based criticisms that can be leveled against probabilistic decision-making in the criminal arena. I will then turn to refute these objections, and show that the probabilistic model may actually fare better than the threshold model in fulfilling the expressive functions of criminal trials.

Expressive theories of law are concerned with the expression of collective attitudes through legal action.[101] The expressive function of criminal law is particularly potent, for concurrent to promoting either deterrence or retribution, the criminal trial serves a communicative function.[102] It constitutes a natural arena for clarification of, and reflection on, the social value scale.[103] Labeling certain behaviors as criminal grants one mor-

100. Shavell, *supra* note 79, at 605.

101.   *See* Elizabeth S. Anderson & Richard H. Pildes, *Expressive Theories of Law: A General Restatement*, 148 U. PA. L. REV. 1503, 1514–18 (2000). As a social practice, law has a significant expressive function, which can be understood in two distinct senses. One is purely symbolic and non consequential. Many people support or object to law not for any consequential reasons (such as the law's deterrent effect), but due to its symbolic content, namely, the declaration it makes about the community's morals and values. *See* Cass R. Sunstein, *On the Expressive Function of Law*, 144 U. PA. L. REV. 2021, 2022–23 (1996). The other aspect of law's expressivity is consequential and relates to its power to shape, change, and reinforce social norms. Law's expressive function is manifested, in this sense, in its ability to influence normative behavior by making statements that create and sustain shared social norms, rather than controlling behavior directly. *See* Paul H. Robinson & John M. Darley, *The Utility of Desert*, 91 NW. U. L. REV. 453, 471 (1997).

102.   Martin Sabelli & Stacey Leyton, *Train Wrecks and Freeway Crashes: An Argument for Fairness and Against Self Representation in the Criminal Justice System*, 91 J. CRIM. L. & CRIMINOLOGY 161, 209 (2000).

103.   For additional discussions of expressive theories of law, see Dan M. Kahan, *The Secret Ambition of Deterrence*, 113 HARV. L. REV. 414 (1999); Dan M. Kahan, *What Do Alternative Sanctions Mean?*, 63 U. CHI. L. REV. 591 (1996); Richard H. McAdams, *A Focal Point Theory of Expressive Law*, 86 VA. L. REV. 1649 (2000); Sunstein, *supra* note 101, at 2021. For criticism of the expressive approach, see Matthew D. Adler, *Expressive Theories of Law: A Skeptical Overview*, 148 U. PA. L. REV. 1363, 1462 (2000).

al approach precedence over contradicting visions of justice.[104] Conviction conveys a message of moral opprobrium of the offender and of the offender's value scale.[105] The severity of criminal punishment, in turn, reflects the force of the moral reprobation. As stated by Dan M. Kahan, "What a community chooses to punish and how severely tells us what (or whom) it values and how much."[106] A probabilistic decision making regime could potentially undermine each of these expressive functions of the criminal trial. In what follows, I will discuss the different ways in which probabilistic decision making may jeopardize the communicative functions of the criminal trial, with emphasis on three tenets: the expressive harm associated with the fragmentation of guilt under probabilistic conviction; the expressive harm associated with the relativity of punishment under probabilistic sentencing; and the potential damage to the conceptualization of judicial decision making as a result of its exposure to probabilistic logic.

A. THE EXPRESSIVE HARM OF CONVICTION RELATIVITY

A probabilistic decision making regime would infuse the notion of guilt with relativity by effectively breaking the category of criminal conviction into varying sub-types, segregated according to their evidentiary strength (conviction on guilt beyond all doubt, conviction on guilt beyond a reasonable doubt, conviction on guilt by clear and convincing evidence, and so forth). A possible objection to such fragmentation is that it would expose the existence of an evidentiary continuum at the base of the criminal verdict and could, therefore, undermine the labeling power of the criminal conviction.

Under the threshold model, the institution of criminal conviction enjoys a monolithic status, due in part to the weighty evidentiary grounds that it requires. The underlying intuition

104. The expressive function of criminal proceedings is also reflected in procedural features. Thus, the criminal procedural structure reflects social justice concepts similarly to the substantive law that applies in the proceedings. *See* Barbara A. Babcock, *Fair Play: Evidence Favorable to an Accused and Effective Assistance of Counsel*, 34 STAN. L. REV. 1133, 1140–42 (1982); Charles H. Koch, Jr., *A Community of Interest in the Due Process Calculus*, 37 HOUS. L. REV. 635, 636 (2000).

105. Erik Lillquist, *Recasting Reasonable Doubt: Decision Theory and the Virtues of Variability*, 36 U.C. DAVIS L. REV. 85, 136 (2002) ("Behavior is criminalized, in part, in an effort to express society's moral condemnation of the behavior, as well as the values that the behavior symbolizes.").

106. Dan M. Kahan*, Social Meaning and the Economic Analysis of Crime*, 27 J. LEGAL STUD. 609, 615 (1998).

is that for the criminal conviction to carry out its denouncing and stigmatizing function, it must be founded on credible information regarding the offender's guilt and must be distinguishable from all other decisions made by the State vis-à-vis its citizenry.[107] Assigning a maximal standard of proof as a prerequisite for conviction protects the criminal conviction from dilution. It sustains the labeling power of conviction by perpetuating the credibility of the information on which it rests and by preserving its unique nature.[108] In light of the high probability of guilt in a conviction, the public is willing to subject the convicted offender to various social sanctions, while internalizing the costs that this entails. Under a probabilistic regime, in contrast, the evidentiary threshold for criminal conviction would be moderated and a dimension of variability would be introduced into it. This could weaken the force of criminal conviction in conveying stigma and moral culpability. The greater extent of uncertainty with regard to the convicted offender's factual guilt may produce negative expressive effects and lead to the dilution of the criminal conviction as a legal institution, resulting in public reluctance to impose social sanctions upon offenders.

B. THE EXPRESSIVE HARM OF PUNISHMENT RELATIVITY

In addition to relativity regarding the notion of guilt, a probabilistic regime would also infuse the function of criminal punishment with relativity, as the sentence imposed following conviction would reflect the epistemic certainty of guilt. There could be room to claim that the leniency of partial punishment, as a result of epistemic doubt, may fail to validate the social value scale. Thus, the realization of the expressive function of criminal punishment is contingent on the ability of the public to forge a link between severity of the punishment and the force of the moral repudiation of the offence and the offender.[109] Since, under the probabilistic model, punishment severity

---

107. Carol S. Steiker*, Punishment and Procedure: Punishment Theory and the Criminal-Civil Procedural Divide*, 85 GEO. L.J. 775, 808 (1997) ("[F]or punishment to be able to perform its blaming function, it must be distinguishable to the wrongdoer and the community at large from all of the other things that governments may do in the course of governing and all of the other burdens citizens must bear as citizens. Thus, special procedures are necessary to make punishment more, rather than less, powerful."); *see also* Lando, *supra* note 66, at 285.

108. *See* Streiker, *supra* note 107, at 809.

109. *See* Richard A. Bierschbach & Alex Stein, *Overenforcement*, 93 GEO. L.J. 1743, 1749–50 (2005).

would be the product not only of the severity of the crime, but also of the probative weight of the incriminating evidence, an ambiguous expressive message may emanate from the courtroom. The reduction of the sanction and imposition of partial punishment, albeit due to evidentiary deficiencies, might be interpreted by the public as stemming from the negligibility of the moral harm, thereby sabotaging the expressive goals of criminal punishment. Put differently, the accepted social meaning of punishment places expressive constraints on the ability to tinker with punishment severity for the realization of exogenous goals, including deterrence goals in the context of probabilistic sentencing.[110]

C. The Expressive Harm of a Probabilistic Conceptualization of Criminal Fact-Finding

Another possible expressive harm which may be attributed to probabilistic decision making relates to the adverse effect associated with exposure of the evidentiary continuum at the basis of the court's findings. Such transparency may impair the expressive role of the criminal trial in preserving the proper conceptualization of judicial decision making. For the judiciary to fulfill its purpose in a liberal society, judicial decision-making must be construed as premised upon factual truth.[111] While factual truth is categorical in nature (did or did not occur; valid or invalid),[112] adjudicative fact-finding is inherently infused with uncertainty.[113] Under the threshold model, when the court reconstructs the events underlying the trial, it converts its intermediate certainty level into a categorical finding

---

110. *See id.*

111. David S. Kaplan & Lisa Dixon, *Coerced Waiver and Coerced Consent*, 74 DENV. U. L. REV. 941, 953–54 (1997) (discussing the public policy goal of truth-seeking).

112. The claim that factual truth is categorical and unambiguous in nature can be contested, in light of the interdependency between fact and law and between fact and value. *See* HOCK LAI HO, A PHILOSOPHY OF EVIDENCE LAW 7 (2008). Philosophical skepticism also poses a challenge to the categorical nature of fact-finding. *See* Laurence BonJour, *Can Empirical Knowledge Have a Foundation?*, 15 AM. PHIL. Q. 1, 1 (1978).

113. Stein, *supra* note 96, at 297–301 (claiming that adjudicative fact-finding is bound to rest upon probabilities rather than certainty for a number of reasons: it does not always relate to the past but can instead consist of predictions of future occurrences; in the criminal trial context in particular, it deals not only with the empirical question of whether something happened, but also why something happened; and the fact that some facts are not easily severable from value judgments).

of conviction or acquittal.[114] This requirement for conversion of adjudicative fact-finding into a categorical outcome can be justified on the background of the desire to preserve the potential of a link to the factual truth. Since the factual truth is categorical in nature, the verdict that allegedly manifests that truth must have a similar epistemic structure.

Under a probabilistic regime, in contrast, such conversion would not take place. Judicial fact-finding decisions would systematically and a priori deviate from what occurred in reality, in all those situations where partial probability of (full) criminal responsibility would be translated into a determination of "partial criminal responsibility."[115] That is to say, the difference between the probabilistic regime and the threshold model can be conceptualized as the tension between a priori systematic error and ex post random error—or as the difference between situations in which the court makes a "half-error" in 100 percent of the cases and those in which it completely errs in 50 percent of the cases.[116] Given this, a possible criticism of the probabilistic punishment regime could be that creating such a built-in discrepancy between the judicial decision and the factual truth would undermine the way in which judicial decision-making is publicly portrayed. It would alter the perception of judicial decision-making from a process linked to factual truth to a probabilistic game premised on the allocation of risk.

The described expressive harms to the public perception of judicial decision making are commonly presented in the literature from the slightly different perspective of legitimacy and public acceptability of the criminal legal system. Nesson[117] and Tribe[118] are the central authors to have linked public trust in

---

114. Bray, *supra* note 4, at 1307.

115. This is analogous to the blurring of the distinction between a determination that "the door is wide open in half of the cases" and the determination that "the door is half-open."

116. John E. Coons, *Compromise as Precise Justice*, 68 CALIF. L. REV. 250, 258 (1980).

117. Nesson, *supra* note 79, at 1358 (claiming that the successful projection of legal rules and the facilitation of public trust in the judicial system depend on a court's ability to cast the verdict as a statement about the actual occurrence rather than as a statement about the evidence); *see also* Abramowicz, *supra* note 75, at 256; Scott E. Sundby, *The Virtues of a Procedural View of Innocence—A Response to Professor Schwartz*, 41 HASTINGS L.J. 161, 162–63 (1989).

118. Laurence H. Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329, 1376–77 (1971) (advising caution in us-

the adjudication system to the binary threshold regime. Under their approach, courts' legitimacy is contingent on the public's perception of the criminal verdict as "statements about what actually happened."[119] Nesson and Tribe both maintain that the justification for the beyond-a-reasonable-doubt evidentiary threshold derives not only from its positive effect on the actual accuracy of judicial decisions, but also from the public trust embodied in maintaining an appearance of accuracy.[120] From this standpoint, the probabilistic punishment model can be criticized in that its admission of the existence of an evidentiary continuum would "transform the substantive message from one of morality . . . to one of crude risk calculation," thereby undermining the legitimacy of the criminal justice system in the eyes of the public.[121]

D. THE EXPRESSIVE CASE FOR PROBABILISTIC CONVICTIONS

Though at first glance expressive considerations seem to support the threshold model in light of the abovementioned expressive drawbacks associated with probabilistic decision making (namely, dilution of the criminal conviction's stigmatizing power, impairment of the communicative function of criminal punishment, and the adverse conceptualization of judicial decision-making), closer scrutiny reveals that a probabilistic decision-making model is actually better-equipped to fulfill the expressive functions of criminal trials. The incorporation of relativity into criminal conviction and the incorporation of variability into criminal sentencing allow for the refinement and fortification of the criminal trial's expressive message. I will turn to substantiate each of these claims, starting with the expressive advantages associated with a fragmented conception of guilt.

The question of criminal guilt invokes the most complex and tangled categories dealt with in law, interweaving the descriptive and the normative. Findings of guilt depict the alleged actions in a morally laden manner, and entail the weighing of a

---

ing statistical evidence and claiming that mathematical information must be converted for the verdict to facilitate public trust in the judicial process).

119. Nesson, *supra* note 79, at 1358.

120. Lillquist, *supra* note 105, at 177 ("[T]heir writings suggest the possibility that, irrespective of any costs and benefits from accurate and inaccurate verdicts, the way in which the current reasonable doubt instruction is implemented can be explained . . . by the benefits that arise from promoting these other values.").

121. Nesson, *supra* note 79, at 1362.

rich assortment of factual, normative, social, and emotional variables.[122] The binary threshold model dictates that the manifold aspects of criminal culpability be ultimately translated into the legal lexicon's strict, one-dimensional terms of conviction or acquittal. It can be argued that such impoverished conceptualization of the criminal trial's outcome results in the loss of valuable information.[123] A probabilistic regime, in contrast, would allow for a more nuanced and sophisticated answer to the question of criminal responsibility. By creating multiple standards of criminal conviction, the probabilistic model would facilitate a more accurate reflection of the evidentiary gray areas that permeate criminal decision making, and would enable finer regulation of the accompanying social sanctions.[124] The fragmentation of the criminal conviction into varying sub-types, reflective of differential probative values, would pave the way for transformation of the question of criminal responsibility from a qualitative question of "yes or no" to the quantitative "how much".[125] Put differently, while the expressive function of the

---

122. Take the classic example of murder. When the defendant is found guilty of murder, the finding does not simply record information about the occurrence. The verdict also passes a value judgment about what happened. *See* HO, *supra* note 112, at 9; *see also* Weigend, *supra* note 56, at 170 (discussing the notion of "the whole truth," which encompasses the "myriad [of] facts including the psychological and biographical factors that might help explain why the offense was committed").

123. For further discussion, see Fisher, *supra* note 91, at 988–89.

124. An infinite fragmentation of the criminal conviction into sub classes and categories may, indeed, impair its force in transmitting information regarding factual guilt. But as noted earlier, under the probabilistic punishment model, the conviction need not be broken down endlessly; four or five conviction categories may suffice. *See supra* Parts I.A–D.

125. For a similar approach with respect to contractual liability, see HANOCH DAGAN, THE LAW AND ETHICS OF RESTITUTION 268–72 (2004) (challenging the conceptual binarism underlying the distinction between breach of contract and performance). It is true that in the criminal culpability context the abovementioned transformation is only partial, in the sense that it would allow for refinement of the judicial verdict and the concept of criminal liability only with regard to the underlying evidentiary dimension. The normative aspect of the question of criminal culpability remains binary and continues to be relegated to either full guilt or innocence. In other words, there is indeed a basic distinction between a verdict dealing with (full) criminal liability at partial probability and a verdict that deals with "partial criminal liability." The probabilistic model allows for verdicts of the first kind, which depend on an evidentiary base, but this would not suffice to establish the essential infrastructure for a verdict of the second kind. However, the very conceptualization of criminal liability on a hierarchical scale, even if only in relation to the evidentiary stratum, constitutes a first step toward opening a pluralistic and graduated discourse in relation to the normative question as well.

criminal conviction is indeed contingent on the ability to infer an informative decree from the conviction regarding the defendant's involvement in the crime and, accordingly, on the existence of some minimal standard of proof as a pre condition for conviction, this does not preclude a plurality of conviction categories based upon sub-maximal levels of proof. There is nothing sacred, natural, or pre-political about the existence of *one* monolithic category of conviction.[126] There is nothing in the expressive role of the criminal conviction that mandates division of the world into two pure and mutually exclusive categories of guilt and innocence.[127]

Moreover, when thinking about the expressive function of the criminal verdict, focus should not be placed solely on the expressive function of the criminal conviction. Rather, the expressive role played by the institution of the criminal acquittal must also be taken into account. Under the threshold model, acquittal covers a vast epistemic space, stretching from the end-point of full certainty regarding innocence to just below the beyond-a-reasonable-doubt threshold as to guilt. In such circumstances, criminal acquittal can serve only a limited expressive function: it cannot clear a defendant's name because it does not necessarily indicate sufficiently high levels of certainty regarding innocence.[128] In fact, under the binary configuration of the threshold model, the defendant cannot avoid both conviction and a stigmatizing acquittal, for the protection from conviction—in the form of a high evidentiary threshold—is itself the source of stigma that acquittal cannot remove.[129] Under a probabilistic regime, on the other hand, the expressive trade-off is different. While the message of acquittal is fine-tuned due to the narrowing of the epistemic space it encompasses, this does not necessarily entail a higher prevalence of beyond-a-reasonable-doubt convictions. The options of lower-level convictions allow for moderated findings of guilt, matched with lenient punishments. In sum, the expressive advantages associated with a fragmented conception of guilt include positive effects on

---

126. Bray, *supra* note 4, at 1324.

127. *See* Mark J. Osiel, *Ever Again: Legal Remembrance of Administrative Massacre*, 144 U. PA. L. REV. 463, 582 (1995) (discussing the binary logic inherent to criminal law "with its insistence on dividing the world into mutually exclusive categories of people: legally, into guilty and innocent; sociologically, into blamers and blamed").

128. Bray, *supra* note 4, at 1325.

129. *Id.*

the expressive role played by the institution of criminal acquittal.[130]

E. THE EXPRESSIVE CASE FOR PROBABILISTIC SENTENCING

There is also room to reconsider the hypothetical objections, portrayed above, regarding the expressive toll of sentencing variability—namely, the argument that probabilistic punishment would fail to validate the social value scale. Under a probabilistic punishment regime, the probative weight of the evidence would be expressed explicitly in the court's actual determination regarding type of conviction (and in the distinctions between conviction on guilt beyond a reasonable doubt, convictions on guilt by clear and convincing evidence, and so forth). The transparency as to the extent of epistemic doubt, incorporated into the verdict, would allow the public to calculate the gravity of condemnation of the given criminal conduct under the assumption of maximal certainty.

More importantly, even under the threshold model, severity of punishment is determined on the basis of a multiplicity of variables. The defendant's age, family status, or chances of rehabilitation all affect sentence gravity. There is no reason to assume that the expressive function of punishment, or its capacity to reflect the full extent of social condemnation of the criminal activity, would be impaired by including certainty of guilt in the long list of considerations already incorporated into punishment severity.

F. THE EXPRESSIVE CASE FOR THE PROBABILISTIC CONCEPTUALIZATION OF CRIMINAL FACT-FINDING

The objection to probabilistic decision-making, rooted in the potential damage to public trust in the criminal justice system, can also be contested. First, the conceptualization of public acceptability as an intrinsic normative end is, in itself, questionable. The system must merit acceptability; the matter of trust cannot be detached from the objective performance of the courts. It is the improvement in the court's performance that is the ultimate end, rather than preservation of public trust per se. In terms of objective performance, a probabilistic decision making regime fares better than the threshold model in several respects: probabilistic decision making promotes deterrence, fa-

---

130. For further discussion of the expressive costs associated with false acquittal, see Lillquist, *supra* note 105.

cilitates a fine-tuned expressive message as to guilt and punishment, and allows for a fairer ex post imposition of punishments among offenders (as will be elaborated below).[131] Second, even if one accepts the aspiration to public acceptability as an intrinsic goal of the criminal trial, it is doubtful that this goal can be attained by denying the risks of error inherent to all judicial decisions.[132] There is no empirical basis to the contention that creating trust in the criminal justice system is contingent on the transmission of a misleading message of certainty as to the guilt of every person convicted in court.[133] In fact, quite the opposite may be true: the dissonance between public awareness of wrongful convictions and the façade of certainty of guilt could fatally undermine the system's public legitimacy. As Abramowicz has asserted, "Even if legal institutions have developed methods of fooling the public because such fooling in individual cases increases the prestige of the courts, the general practice of trying to fool the public may decrease prestige even more."[134] Indeed, making the probabilistic foundation of the judicial decision transparent may foster public trust in the judicial system far more effectively than camouflaging the risks of error.

IV. RETRIBUTIVE DIMENSIONS OF THE PROBABILISTIC MODEL

Unlike consequentialist theories of punishment—which justify criminal punishment in light of its instrumental role in pursuing desirable social ends, such as deterrence or communication of social repudiation—retributivism holds the criminal

---

131. *See infra* Parts IV.A–D.

132. *See* Talia Fisher & Issachar Rosen-Zvi, *The Confessional Penalty*, 30 CARDOZO L. REV 871, 908 (2008).

133. *See* Daniel Shaviro, *Statistical Probability Evidence and the Appearance of Justice*, 103 HARV. L. REV. 530, 544 (1989) ("Perhaps the best way to create a perception of verdict accuracy is to create the reality of verdict accuracy. Perceptions created at the expense of the reality may be unstable, especially over the long term, because not everyone is likely to be fooled. Some people may know that the system as a general matter is pursuing perception at the expense of reality. Other people may simply know or believe that verdicts are inaccurate in particular cases; more such cases will create more such people. The question then becomes one of determining how widely the disillusioning knowledge about verdict error will spread over time. If society learns that the legal system pursues perception at the expense of reality, it may experience greater disillusionment than if it merely discovered some inaccuracy.").

134. Abramowicz, *supra* note 75, at 239 n.104.

punishment of the culpable to be an intrinsic good.[135] At the center of retributivism stands the premise that criminal punishment, like any other social institution, must be valid from a moral standpoint and its ramifications for public welfare cannot constitute a justifying rationale.[136] Under the retributivist approach, the sole justification for criminal punishment is the connection between the sanction and the criminal act committed.[137] The conjectural retributivist line of criticism, which can be formulated against the probabilistic model, is that such a decision-making regime would generate moral harm that cannot be outweighed by its social outcomes, including any deterrence utility. The potential moral harm posed by probabilistic decision making can be described as having two central manifestations. First, by allowing the imposition of criminal punishment in the epistemic space below the beyond-a-reasonable-doubt threshold, the probabilistic regime deviates from retributivist principles, which prohibit inflicting punishment absent moral certainty regarding guilt. Second, by incorporating epistemic uncertainty into the magnitude of the sanction, the probabilistic model departs from the retributivist principle of pro-

---

135. *See* Joel Feinberg, *The Classic Debate*, *in* PHILOSOPHY OF LAW 727, 727–29 (Joel Feinberg & Jules Coleman eds., 6th ed. 2000) (claiming that consequentialists look forward with regard to criminal punishment and focus on the instrumentality of the criminal sanction in advancing desired social outcomes, whereas retributivists take a retrospective view of punishment); *see also* Russell L. Christopher, *Deterring Retributivism: The Injustice of "Just" Punishment*, 96 NW. U. L. REV. 843, 862 (2002) (arguing that retributivism focuses on the appropriate and the just as opposed to the desirable or effective in terms of various teleological goals).

136. Erik Luna, *Punishment Theory, Holism, and the Procedural Conception of Restorative Justice*, 2003 UTAH L. REV. 205, 216 ("[I]nstitutions should produce morally correct decisions regardless of ultimate effects on society . . . . This conception of justice punishes not in service of some greater societal goal but in proportion to the criminal's moral blameworthiness and the harm caused by his offense.").

137. R. A. DUFF, TRIALS AND PUNISHMENTS 4 (1986) (describing retributivist theories as rooted in "the idea that punishment is justified as merited retribution for a past offence"). In Kant's words, "*Punishment by a court* . . . can never be inflicted merely as a means to promote some other good for the criminal himself or for civil society. It must always be inflicted upon him only *because he has committed a crime*." IMMANUEL KANT, THE METAPHYSICS OF MORALS 105 (Mary Gregor ed., trans., 1996). Only a link between the criminal act and the act of punishment can ensure that the defendant's human dignity is preserved and prevent his transformation into an instrument for realizing social goals, "For a human being can never be treated merely as a means to the purposes of another or be put among the objects of rights to things: his innate personality protects him from this, even though he can be condemned to lose his civil personality." *Id.*

portionality, which mandates formulaic sanctions correlating solely with moral guilt. I will hereby address each of these hypothetical retributive objections to probabilistic decision making.

A.   PUNISHMENT IN THE ABSENCE OF MORAL CERTAINTY REGARDING GUILT

A principal element in retributivist thought is that the sole justification for punishment is the existence of guilt,[138] and that punishing in the absence of the highest humanly possible level of certainty as to guilt is morally illegitimate.[139] Retributivists would thus criticize the probabilistic model for eroding the evidence threshold and thereby weakening the moral legitimacy of inflicting criminal punishment. Imposing punishment based on evidence with a probative weight falling below the beyond-a-reasonable-doubt standard (i.e., below the standard of moral certainty) results in the defendant's objectification, causing moral harm that is incommensurable with deterrence utilities.[140]

---

138.   A. M. Quinton, *On Punishment,* 14 ANALYSIS 133, 134–38 (1954) (discussing the ontological connection between punishment and a past offense and arguing that punishment of the innocent is impossible because if the person is not guilty, then what is imposed upon him cannot be deemed punishment). Some retributivists posit that this refers to moral guilt; others contend that it refers to legal guilt. For further discussion of the distinction between "moralistic" and "legalistic" retributivism, see Christopher, *supra* note 135, at 881.

139.   Michael S. Moore, *The Moral Worth of Retribution*, *in* RESPONSIBILITY, CHARACTER, AND THE EMOTIONS: NEW ESSAYS IN MORAL PSYCHOLOGY 179, 179 (Ferdinand Schoeman ed., 1987) ("*Retributivism* is the view that punishment is justified by the moral culpability of those who receive it. A retributivist punishes because, and only because, the offender deserves it.").

The scope of this Article precludes an extensive discussion of the complex matter of defining moral certainty, although I will address further the link between moral certainty and the standard of proof in criminal proceedings. For the time being, I will suffice with the following definition of moral certainty: "[T]he form of certainty that any person is capable of achieving from an understanding of the nature of things, applying reason and thought to the testimony of others, along with personal observation and experience." Steve Sheppard, *The Metamorphoses of Reasonable Doubt: How Changes in the Burden of Proof Have Weakened the Presumption of Innocence*, 78 NOTRE DAME L. REV. 1165, 1177 (2003).

140.   *See* Lillquist, *supra* note 105, at 140 (describing the incommensurability principle). In other words, from the retributivist viewpoint, the deterrence approach turns things upside down in extolling certainty in punishing the guilty when what should in fact stand, absolutely, at the center of every theory of punishment is certainty of not punishing the innocent. *See* Erik Lillquist, *The Puzzling Return of Jury Sentencing: Misgivings About* Apprendi, 82 N.C. L. REV. 621, 695 (2004) [hereinafter Lillquist, *The Puzzling Return*].

B. DEVIATION FROM THE PRINCIPLE OF PROPORTIONALITY

The existence of guilt, albeit a mandatory requirement, is not sufficient to legitimize criminal punishment from a retributivist standpoint. In addition, the severity of the punishment must be reasonable and proportionate to the severity of the crime committed.[141] This principle of proportionality, which prescribes the offender's just desert, was described by Kant as follows:

> [W]hatever undeserved evil you inflict upon another within the people, that you inflict upon yourself. If you insult him, you insult yourself; if you steal from him, you steal from yourself; if you strike him, you strike yourself; if you kill him, you kill yourself. But only the *law of retribution* (*ius talionis*) . . . can specify definitely the quality and the quantity of punishment; all other principles are fluctuating and unsuited for a sentence of pure and strict justice because extraneous considerations are mixed into them.[142]

Hegel explained the principle of proportionality on the fact that both the crime and act of punishment constitute coercion.[143] The latter act of coercion—infliction of punishment—has the power to eliminate the coercive element of the earlier action—the criminal act. This can occur only if there is correspondence between the second and first acts of coercion.[144]

141. A distinction should be made in this context between retributivist approaches that advocate mandatory punishment and those that view the existence of guilt as a justification for punishment but do not mandate its imposition. Kant and Hegel believed that the existence of guilt creates an obligation on the part of society and the sovereign to inflict criminal punishment. *See* G.W.F. HEGEL, ELEMENTS OF THE PHILOSOPHY OF RIGHT § 100 (Allen W. Wood ed., H.B. Nisbet, trans., 1991); KANT, *supra* note 137. According to Kant, "the law of punishment is a categorical imperative, and woe to him who crawls through the windings of eudaemonism in order to discover something that releases the criminal from punishment . . . ." KANT, *supra* note 137. In contrast, there are thinkers who contend that the existence of guilt creates the authority, and not the duty, to punish. *E.g.,* Michael T. Cahill, *Retributivist Justice in the Real World*, 85 WASH. U. L. REV. 815, 826–28, 828 n.36 (2007).

142. KANT, *supra* note 137, at 105–06. For an expansion on this principle, see Richard A. Bierschbach & Alex Stein, *Mediating Rules in Criminal Law*, 93 VA. L. REV. 1197, 1204–05 (2007); Christopher, *supra* note 135, at 860 n.79; Jeffrie G. Murphy, *Does Kant Have a Theory of Punishment?*, 87 COLUM. L. REV. 509, 530–32 (1987). There is a distinction in this context between negative retribution approaches, which set only the uppermost standard of punishment, and positive retribution approaches, which also set a minimal standard of punishment.

143. *See* HEGEL, *supra* note 141, §§ 94–95.

144. For further discussion of Hegel's articulation of the *lex talionis* principle, see Russell L. Christopher, *The Prosecutor's Dilemma: Bargains and Punishments*, 72 FORDHAM L. REV. 93, 127 (2003); *see also* Markus Dirk, *Rediscovering Hegel's Theory of Crime and Punishment*, 92 MICH. L. REV. 1577, 1581–83 (1994) (discussing Hegel's view of crime and punishment).

The proportionality principle mandates the imposition of formulaic sanctions suited to the moral gravity of the underlying crime.[145] The principle gives rise to another possible retributivist objection to the probabilistic regime, namely, its deviation from the principles of just desert due to epistemic uncertainty. Indeed, retributivists would strongly reject the merging of the penal and epistemic dimensions of criminal proceedings and the accompanying breach of proportionality between crime and punishment. In cases where the level of proof against the defendant does not reach the beyond-a-reasonable-doubt threshold, a probabilistic decision-making regime would allow the epistemic gulf to be bridged through penal means, by eroding—in retributivist terms—the criminal sanction to a level below what is prescribed by the proportionality principle.[146] Even in cases in which the defendant's guilt has been proven beyond a reasonable doubt, probabilistic punishment would constitute a deviation from the just desert requirement, due to the penal variability that it entails. Penal variability—as a function of the probative weight of the evidence in any given case—collides head-on with the principle of proportionality, which prescribes the imposition of a sanction that reflects only the gravity of the crime, upon moral certainty of guilt.[147]

C. The Retributive Critique Reconsidered

Admittedly, from an internal-retributive viewpoint, the attempt to justify the probabilistic model is likely to appear unconvincing. In my opinion, however, the retributivist theory cannot serve as a standard for assessing the normative desirability of any given decision regime, including the probabilistic decision-making model. The reason is that the retributivist theory focuses on the individual case and does not address the general systemic level.[148] Retributivists are concurrently com-

---

145. For an extensive discussion of this principle, see Thomas E. Hill, *Kant on Wrongdoing, Desert and Punishment*, 18 LAW & PHIL. 407, 428–37 (1999); *see also* MICHAEL S. MOORE, PLACING BLAME: A GENERAL THEORY OF THE CRIMINAL LAW 202 (1997) (discussing the substantive proportionality of punishment in context of completed versus attempted crimes).

146. This type of harm can be understood as a sort of mirror-image of the harm deriving from punishment in the absence of moral certainty, discussed above. Under the retributivist approach, just as moral certainty of guilt cannot be waived as a precondition for criminal punishment, so is non-proportional punishment due to epistemic doubts intolerable. *See* Lillquist, *The Puzzling Return*, *supra* note 140, at 623–24.

147. *See* Luna, *supra* note 136, at 223.

148. *See* HEGEL, *supra* note 141, § 99.

mitted to two fundamental principles: punishing the guilty and not punishing the innocent.[149] Deviation from either one of these outcomes is considered a departure from the principles of just desert.[150]

But due to the focus on the individual case in retributivist thought, these two mandates are construed in absolute terms, without addressing the question of which takes priority. Indeed, retributivism takes no stance with regard to the trade-off between wrongful conviction and wrongful acquittal over the general cross-section of cases. It is, in this sense, a utopian theory, disregarding the difficult day-to-day reality of unavoidable errors in legal fact-finding, where decreasing the incidence of wrongful convictions comes at the cost of increasing the overall rate of wrongful acquittals.[151] In light of this lack of a systemic outlook and the absolutist conception of the principles of non-punishment of the innocent and punishment of the guilty—with no inquiry into the comparability between the two—many contend that the retributivist theories cannot, in fact, justify any type of decision regime, the present threshold model included. Reiman and van den Haag, for example, assert that the retributivist theories cannot justify the beyond-a-reasonable-doubt threshold, as it is inconsistent with the retributivist commitment to punishing the guilty: "Should we try to convict fewer innocents and risk letting more of the guilty escape, or try to convict more of the guilty, and, unavoidably, more of the innocent? Retributivism (although not necessarily retributivists) is mute on how high standards of proof ought to be . . . ."[152] Similarly, Michael Moore asserts that the retributivist approach no less justifies the relatively lower preponderance-of-evidence standard in the criminal sphere than it does the heightened beyond-a-reasonable-doubt standard. In his opinion, absent any a priori commitment to the appropriate systemic ratio between wrongful convictions and wrongful acquittals, "[t]he retributivist might adopt a principle of symmetry here—the guilty going unpunished is exactly the same magnitude of evil as the innocent being punished—and design his institu-

---

149. Christopher, *supra* note 135, at 848.

150. *Id.*

151. *See* Luna, *supra* note 136, at 219 (noting that retributivism in practice results in "factually innocent individuals, as well as legally justified or excused defendants, . . . be[ing] subject to undeserved punishment").

152. Jeffrey Reiman & Ernest van den Haag, *On the Common Saying That It Is Better That Ten Guilty Persons Escape Than That One Innocent Suffer: Pro and Con*, 7 SOC. PHIL. & POL'Y, Spring 1990, at 226, 242–43.

tions accordingly."[153] At the very most, then, retributivist theories embody pure ethics, while they are unable to lay an implementable normative groundwork upon which to evaluate different decision regimes. In neglecting systematic thinking, retributive theories cannot serve as a source of valid criticism against probabilistic decision making.[154]

---

153. MOORE, *supra* note 145, at 157 n.11.

154. Moreover, as a practical matter, epistemic mistakes in the administration of criminal justice are unavoidable: "[B]ecause punishment is administered by the state rather than by god . . . it is inevitable that the practice of punishment will suffer from (at least) each of the following three deficiencies: It will be tremendously expensive, subject to grave error, and susceptible to enormous abuse." Douglas N. Husak, *Why Punish the Deserving?*, 26 NOÛS 447, 450 (1992). It is therefore impossible to ensure that both mandates—punishing the guilty and not punishing the innocent—are simultaneously upheld.

Retributivists are thus susceptible to the same criticism they direct at consequentialist approaches. For so long as they do not out-and-out reject criminal punishment, retributivists can be accused of subjecting the innocent to punishment to further the goal of punishing the guilty and thus of turning them into a means for furthering external goals: "The retributivist remains 'willing to trade the welfare of the innocents who are punished by mistake for the greater good of the punishment of the guilty' and thus, it would seem committed to sacrificing—'using'—the mistakenly convicted for the benefit of society in general." David Dolinko, *Three Mistakes of Retributivism*, 39 UCLA L. REV. 1623, 1632–33 (1992) [hereinafter Dolinko, *Three Mistakes of Retributivism*]. The retributivist does this "by appealing to the inevitability of mistaken convictions," for this reason "retributivism itself can be accused of using convicted offenders, and thus stripped of its cloak of Kantian respectability." *Id.* at 1632.

The common retributivist response to these claims is that an a priori systematic error is quite different from an ex-post random error. Retributivist theorists contend that adopting a criminal penal system, with all of its accompanying dangers, creates the risk of unjust desert for an unidentified public, but that it is qualitatively different from exposing specific defendants to a deliberate risk. *See, e.g.*, Larry Alexander, *Retributivism and the Inadvertent Punishment of the Innocent*, 2 LAW & PHIL. 233, 235–36 (1983). These theorists base the moral legitimacy of the act of criminal punishment also on the doctrine of double effect propounded by Aquinas, under which moral norms are absolutely binding only in cases of intentional outcomes. THOMAS AQUINAS, ON LAW, MORALITY, AND POLITICS 227–28 (William P. Baumgarth & Richard J. Regan eds., 1988). An action that breeds concurrent positive and negative results will be considered illegitimate from a moral perspective if the negative result is an intentional outcome—whether as an end or as a means to other ends. *See id.* The action will be allowed, in contrast, if the negative by-product—even though anticipated in advance—is not intended as a final end or an intermediate means. *See* David Dolinko, *Retributivism, Consequentialism, and the Intrinsic Goodness of Punishment*, 16 LAW & PHIL. 507, 512 (1997); Michael S. Moore, *Justifying Retributivism*, 27 ISR. L. REV. 15, 20 (1993).

The retributivist rationale for the institution of criminal punishment is, therefore, that the difficult outcome of wrongful convictions, despite being

D. The Mixed Retributivist-Utilitarian Perspective

The following discussion will address a slightly different type of critique which can be formulated against the probabilistic model—one that incorporates retributivist elements, while offering the systemic perspective neglected in purely retributivist theories. According to the potential line of criticism, the specific trade-off between false convictions and deterrence, expected in the framework of a probabilistic regime, is morally flawed. This type of criticism can be described as a mixed retributivist-utilitarian critique in that it rests on retributive foundations yet recognizes the possibility, in principle, of creating a moral balance between wrongful conviction and other social values.

Many attempts have been made to reconcile the consequentialist and retributivist theories of punishment, with Hart's punishment theory at their center.[155] Hart's justification for criminal punishment is two-layered. The first layer, termed the "general justifying aim," relates to the justification of the

known and anticipated in advance, is not intentional—not as a final end nor as a means for punishing the guilty. In other words, under the retributivist paradigm, the moral legitimacy of criminal proceedings and the act of punishment is determined according to the efforts taken to prevent wrongful conviction and not the outcome of those efforts. In this sense, the wrongful conviction of a given defendant under the existing criminal proceedings model, whereby he receives the maximum punishment of ten years in prison, is not comparable to the parallel scenario likely to arise under a probabilistic punishment model: the conviction of a number of innocents by a lower standard of evidence and the infliction of one year of imprisonment on each. *See* Albert W. Alschuler, *The Changing Plea Bargaining Debate*, 69 Calif. L. Rev. 652, 714 (1981); Christopher, *supra* note 135, at 901.

There is a clear difference between the two scenarios even if we assume that their outcomes are identical from a substantive perspective. This difference is rooted in the fact that the first wrongful conviction is not systematic to the criminal proceeding or one of its intentional outcomes. In contrast, the wrongful convictions in the second scenario are a built-in component of the proceeding and unjust from the perspective of the decision regime itself. Here the question arises as to whether the probabilistic punishment model also meets the standards set in the double effect doctrine, and there may be conflicting views in this regard. For an analogous discussion in the context of reducing the criminal standard of evidence, see Christopher, *supra* note 135, at 913–15. For additional criticism of the application of the double effect doctrine, see Dolinko, *Three Mistakes of Retributivism, supra* at 1633–36.

155. H.L.A. Hart, Punishment and Responsibility: Essays in the Philosophy of Law 8–13 (reprt. 1988); *see also* Hamish Stewart, *Legality and Morality in H.L.A. Hart's Theory of Criminal Law,* 52 SMU L. Rev. 201, 203 (1999) (discussing Hart's theory of liability and criminal punishment, and describing it as an "alternative to viewing criminal punishment as purely consequentialist or purely utilitarian").

general practice of punishment and rests on consequentialist grounds.[156] It addresses the useful outcomes of criminal punishment in terms of deterrence. The second layer—the "principles of distribution"—relates to the question of who should be punished and how punishment should be meted out in society.[157] The answer to this second question, according to Hart, is grounded on retributivist precepts: only those who have committed a crime should be punished.[158] Hart offered a similar mixed retributive-utilitarian approach with regard to the severity of the criminal sanction. While still insisting that the punishment should be consistent with principles of just desert, he also acknowledged the possibility of deterrence goals constraining the level of punishment.[159]

Despite Hart's incorporation of deterrence elements into the theory and justification of criminal punishment, the probabilistic punishment model may still not align with it.[160] From the Hartian perspective, the difficulty with a probabilistic trade-off that renders more convictions but less severe penalties on average is not rooted in the reduction of the sanction's severity per se; for, as Hart himself acknowledges, deterrence goals can constitute an upper limit to punishment.[161] Rather, what is problematic is the possibility of imposing criminal punishment on defendants whose guilt has not been proven beyond a reasonable doubt. The potential for criminal punishment absent moral guilt clashes with the distributional component of Hart's justification and, thereby, forms the crux of the conjectural resistance to a probabilistic decision-making model.

In my opinion, however, considerations of punishment distribution actually support the probabilistic model, rather than undermine it. Under the threshold model, minimal differences in the probative weight of the incriminating evidence result in dramatic discrepancies in the outcomes of the judicial process. Whereas when the probability of guilt exceeds the reasonable doubt threshold even slightly, full-blown punishment is im-

---

156.   HART, *supra* note 155, at 8–11.

157.   *Id.* at 11–13.

158.   Stewart, *supra* note 155. For a further discussion of Hart's distinction between the "general justifying aim" of punishment and its distribution, see Christopher, *supra* note 135, at 868–69.

159.   *See* HART, *supra* note 155, at 88.

160.   This is a central point of criticism made by retributivists against Hart's approach. *See, e.g.*, Igor Primoratz, *Mixed Rationales*, *in* THE PHILOSOPHY OF LAW: AN ENCYCLOPEDIA 559, 560 (Christopher Berry Gray ed., 1999).

161.   *See* HART, *supra* note 155, at 88.

posed; when the probability of criminal culpability falls slightly below the evidentiary threshold, the defendant is acquitted and completely exempt from criminal punishment. It can be argued that such dramatic divergences in outcome due to only minimal epistemic disparities violate the principles of justice and fairness. In other words, the binary dichotomy of the threshold model may be contested in light of its effect on the distribution of punishment across the class of defendants. The distributive outcome of this all-or-nothing regime is that many of the factually guilty are released from all responsibility and go totally unpunished, while a minority of defendants incurs a high level of criminal punishment, in excess of what they would have received were the punishment resources distributed across a wider group of defendants. Although all of these potential defendants are exposed ex ante to the same punishment expectancy, ignoring the ex post distributive outcomes of the threshold model may nonetheless be problematic. The distributive trade-off under the probabilistic model could be more appropriate in this context: turning the dichotomy into a continuum and punishing a greater number of the factually guilty while reducing sanction severity would reduce the penal disparities between cases of similar epistemic infrastructure. Probabilistic sentencing lessens the ex post outcome distortions among cases exhibiting minimal epistemic divergences, leading to a more just and equal distribution of punishment across the general class of defendants.[162]

---

162. Another possible caveat of the hypothetical retributive and Hartian critiques of probabilistic decision making is that when considering the moral trade-off, it is important to weigh not only the rights of the accused but also those of the victims of crime. Both the classic retributivist approach and the distribution component of Hart's theory can be criticized for disregarding the actual and potential victim of the crime. *See* Christopher, *Deterring Retributivism*, *supra* note 135, at 874 (claiming that "though deterrence theories may be unjust in justifying the intentional punishment of innocents, retributivism is unjust in exposing the innocent general public (innocent future crime victims) to a greater risk of victimization").

Similarly to how the deterrence approach uses the defendant to realize external social goals, classical retributivism makes instrumental use of the present and future victims of crime, in its protection of defendants as subjects. *Id.* at 951 ("On the one hand, incorporating the interests of the victim into a determination of the deserved punishment of the victimizer seems incompatible with retributivism. On the other hand, using one set of persons (crime victims) as mere means in order to treat another set of persons (offenders) as ends in themselves, if not incompatible with retributivism, renders retributivism subject to one of its principal criticisms of consequentialism." (citation omitted)). Prominent contemporary retributivist scholars, such as George Fletcher, have even proposed expanding the conception of retribution so as to also include

V. THE POLITICAL LEGITIMACY OF THE PROBABILISTIC MODEL

Irrespective of the deterrence, expressive, and distributive considerations supporting the probabilistic model, this decision-making regime does pose a challenge to the liberal state model. As a distinct part of the law wielded by state organs, criminal law is the strongest expression of the state-law link. Criminal punishment is a potent manifestation of state coercion and poses the most substantial challenge to the justification of state action: "If punishment can be justified, so can other, lesser, forms of coercive state action. If it cannot, what's the point of legitimizing, say, taxation?"[163] Questions of political legitimacy arise in each of the criminal law arenas, from the very definition of the criminal offense (substantive criminal law), to its implementation in concrete cases (procedural criminal law), to its enforcement through the imposition of criminal punishment (enforcement).[164]

As a general rule, the liberal state is committed to ideological neutrality toward the different conceptions of good and therefore must refrain from moral denunciation, which is perceived as an infringement on individual autonomy.[165] Yet the State's act of defining the criminal offense is aimed entirely at giving certain conceptions of good precedence over competing anti-ethical notions, and at imposing these moral stances on the public at large. Criminal law thus constitutes a significant exception to the liberal principle of non intervention.[166] The

---

victims. GEORGE FLETCHER, WITH JUSTICE FOR SOME: VICTIMS' RIGHTS IN CRIMINAL TRIALS 202–05 (1995).

163.   Markus Dirk Dubber, A Political Theory of Criminal Law: Autonomy and the Legitimacy of State Punishment 1 (Mar. 15, 2004) (unpublished manuscript), *available at* http://ssrn.com/abstract=529522.

164.   Markus Dirk Dubber, *The Criminal Trial and the Legitimation of Punishment*, *in* THE TRIAL ON TRIAL: TRUTH AND DUE PROCESS 85, 85 (Antony Duff et al. eds., 2004).

165.   Stephen A. Gardbaum, *Why the Liberal State Can Promote Morals After All*, 104 HARV. L. REV. 1350, 1351 (1991).

166.   *See* Sharon Dolovich, *Legitimate Punishment in Liberal Democracy*, 7 BUFF. CRIM. L. REV. 307, 310 (2004) ("Any theory of state punishment in a liberal democracy must grapple with the problem of political legitimacy."); *see also* R.A. DUFF, PUNISHMENT, COMMUNICATION AND COMMUNITY 35 (2001) ("A normative theory of punishment must include a conception of crime as that which is to be punished. Such a conception of crime presupposes a conception of the criminal law—of its proper aims and content, of its claims on the citizen. Such a conception of the criminal law presupposes a conception of the state— of its proper role and functions, of its relation to its citizens. Such a conception

same holds true for the procedural and enforcement arenas. By punishing a person, the State not only strips him of property and liberty or otherwise inflicts on him pain and humiliation, it also brands him as morally culpable. In so doing, the State acts in a way that exceeds its ordinary capacity and authority in a liberal democracy: "Punishment . . . is prima facie illegitimate; in punishing its constituents, the state harms the very people it is supposed to protect, by interfering with the very rights it claims to guarantee, in the name of guaranteeing them."[167]

The extraordinary functions that the state performs in the various arenas of criminal law raise a host of political legitimacy issues. Against this background of exceptional state interference in individual autonomy, and in light of the concern that this unique capacity will be exploited for wrongful concentration of coercive state power, the traditional liberal view advocates restricting substantive criminal law to the *mala per se*.[168] On the procedural front as well, the heightened procedural safeguards extended to defendants and the requirement that conviction be based on a high likelihood of guilt are considered imperative for authorizing such extraordinary state action and for preventing the establishment of guilt by association. They are intended to circumscribe the State's right to exceed its ordinary role and ascribe moral culpability to its citizens, while at the same time ensuring limited use of this coercive state power.[169] The threshold model, and its pivotal requirement of proof beyond a reasonable doubt, are, thus, a matter of political morality and serve as one of the main limitations on state power in modern systems.[170] The probabilistic model, in contrast, enhances the coercive capacity of the state by expanding the

of the state must also include a conception of society and of the relation between state and society.").

167. Dubber, *supra* note 163, at 1; *see also* Stephen P. Garvey, *Lifting the Veil on Punishment*, 7 BUFF. CRIM. L. REV. 443, 443 (2004) ("When the state punishes a person, it treats him as it ordinarily should not. It takes away his property, throws him in prison, or otherwise interferes with his liberty.").

168. For further discussion of this claim, see Bierschbach & Stein, *supra* note 142, at 1254–55.

169. *See* George P. Fletcher, *Two Kinds of Legal Rules: A Comparative Study of Burden-of-Persuasion Practices in Criminal Cases*, 77 YALE L.J. 880, 888–90 (1968) (explaining that the heightened burden of persuasion in criminal trials is attributable to the need to justify the use of criminal sanctions as a means of moral condemnation).

170. *See* ALEX STEIN, FOUNDATIONS OF EVIDENCE LAW 13 (2005); Claire Finkelstein, *Positivism and the Notion of an Offense,* 88 CALIF. L. REV. 335, 382 (2000) (claiming that the presumption of innocence is a presumption in favor of individual liberty).

range of situations in which individuals are exposed to moral denunciation to include even instances in which guilt is not proven beyond a reasonable doubt. This extension of the space of coercive state action, it can be asserted, both violates the liberal principle of nonintervention and lacks political legitimacy.

While the challenge posed by probabilistic decision making to the liberal state model is, indeed, a strong case against its implementation, there is room to qualify it. First, this objection refers to only one component of the probabilistic model—that relating to the category of cases in the epistemic range beneath the beyond-a-reasonable-doubt threshold. It is the possibility of conviction and punishment based on lower certainty of guilt that poses the threat of intensifying state coercion. The liberal model does not inherently conflict with the second component of the probabilistic model, which refers to the epistemic space above the reasonable doubt threshold, and which prescribes variability of punishment as a function of the weight of remaining doubt. Put differently, the political legitimacy perspective can, at most, restrict the applicability of the probabilistic model to the epistemic space above the beyond-a-reasonable-doubt threshold, but it does not preclude the possibility of infusing the criminal arena with probabilistic logic per se.

Second, even with respect to the epistemic range stretching beneath the beyond-a-reasonable-doubt threshold, a case can be made for the probabilistic model. As explained above, under the probabilistic model the expansion of the epistemic space of conviction to sub maximal evidentiary levels is rooted in the effort to remedy a higher than optimal level of false acquittals due to the existence of systematic evidentiary obstacles.[171] Such situations of systematic under-deterrence constitute the criminal counterpart to probabilistic remedies in tort law, where structural failures in proving causation justified a shift to a probabilistic remedies regime.[172] There may be room to claim that insisting on the insurmountable beyond-a-reasonable-doubt threshold under such circumstances—as prescribed by the threshold regime—effectively precludes the execution of the substantive social choice to prohibit the underlying criminal activity. Thus, the very inclusion of a particular activity in the list of criminal offenses expresses the social choice to deter individuals from engaging in that activity. Yet heightened evidentiary requirements in the face of systematic proof barriers

---

171.   *See supra* notes 84–87 and accompanying text.
172.   *See supra* note 76 and accompanying text.

effectively prevent the conviction of individuals who engage in the underlying forbidden activity. This leads to a failure to realize the social choice with respect to that activity and to a de facto deviation from the predetermined processes for resolving political disputes in society. In such cases, it is the very departure from the maximal evidentiary requirements that would enable the effective execution of the substantive social choice. The probabilistic model therefore validates the accepted social processes for assembling the collective value scale from the conflicting visions of the good life.[173] This line of thought better aligns with the Rawlsian understanding of autonomy as contingent upon society's ability to predetermine the processes by which to choose between conflicting social values.[174]

Lastly, the political legitimacy criticism can be qualified in yet another respect, which relates to the victims of crime and to the state-victim link (existing parallel to the state-defendant axis). There may be room to claim that the liberal state's duty toward victims of crime, and the need to balance victims' rights with those of defendants, can serve as possible justifications for expanded use of the state's coercive power, as would occur under a probabilistic regime. This interpretation of the liberal model is naturally debatable and raises a host of possible objections,[175] a deliberation that is beyond the scope of this Article. Suffice it to say for present purposes that upholding victims' rights, if we assume them to be consistent with the underpinnings of the liberal tradition, can constitute yet another rationale for the expanded space of state coercion under the probabilistic decision-making model.

CONCLUSION

The threshold model—with its on-off configuration of guilt and all-or-nothing sentencing—has been established to the point of being considered an inevitable reality. This Article attempted to expand the horizons of legal imagination by bringing to light the alternative of probabilistic decision-making in the criminal arena. As demonstrated, the criminal justice sys-

---

173. A similar measure (but in the opposite direction) was taken by Bierschbach & Stein, *supra* note 142, at 1254.

174. JOHN RAWLS, POLITICAL LIBERALISM 98 (expanded ed. 2005).

175. For further discussion of the claim that victims' rights conflict with the premises of the liberal model, see Ahmed A. White, *Rule of Law and the Limits of Sovereignty: The Private Prison in Jurisprudential Perspective*, 38 AM. CRIM. L. REV. 111 (2001).

tem is overrun by legal practices that reflect probabilistic decision-making logic. Central doctrines in the criminal law world—such as the residual doubt doctrine, the recidivist premium, and even the jury trial penalty—have paved the way for incorporating certainty of guilt into the severity of punishment meted out. Moreover, as this Article shows, the probabilistic model is supported by deterrence and expressive considerations. It is also compatible with fairness and justice considerations, as it leads to a fairer ex post distribution of criminal punishment. The possible drawbacks of the probabilistic regime, whether based on ex post error-cost considerations or on political legitimacy grounds, fail to collapse the model. At the very most, these claims serve to circumscribe the scope of application of a probabilistic decision-making regime in the criminal sphere. Notwithstanding this, however, the question of the regime's parameters is secondary in a reality in which the threshold model is construed as the only possible option. The ideal of binary decision making at trial was questioned already by Hume, when he made the following claim:

> [T]ho' abstract reasoning, and the general maxims of philosophy and law establish this position, that property, and right, and obligation admit not of degrees, yet in our common and negligent way of thinking, we find great difficulty to entertain that opinion, and do even secretly embrace the contrary principle. . . . An action must either be perform'd or not. The necessity there is of choosing one side in these dilemmas, and the impossibility there often is of finding any just medium, oblige us, when we reflect on the matter, to acknowledge, that all . . . obligations are entire. But on the other hand, when we consider the origins of . . . obligation and find that they depend on public utility, and sometimes on the propensities of the imagination, which are seldom entire on any side; we are naturally inclin'd to imagine, that these moral relations admit of an insensible gradation . . . . Half rights and obligations, which seem so natural in common life, are perfect absurdities in their tribunal; for which reason they are often oblig'd to take half arguments for whole ones, in order to terminate the affair one way or other.[176]

   In this Article, I have sought to present a decision making model that does not mandate that we "take half arguments for whole ones" and that recognizes the possibility of gradation in the criminal verdict and sentencing. It is my hope that this will serve as a preliminary opening for continued dialogue on the subject of the place of probabilistic decision making in the criminal arena and for a reevaluation of the threshold paradigm.

---

   176. DAVID HUME, A TREATISE OF HUMAN NATURE 530–31 (1888), *quoted in* Joseph Jaconelli, *Solomonic Justice and the Common Law*, 12 OXFORD J. LEGAL STUD. 480, 481 (1992).