RETRIBUTIVE JUSTICE IN THE REAL WORLD

MICHAEL T. CAHILL[*]

TABLE OF CONTENTS

INTRODUCTION

In 2003, Attorney General John Ashcroft announced a set of Department of Justice (DOJ) policies designed to curtail federal prosecutors' discretion to plea bargain.[1] The new DOJ rules sought to ensure uniform criminal punishments and prevent criminal offenders from cutting deals to avoid deserved punishment.[2] Yet the would-be ban on bargaining seems to have had little, if any, effect.[3] Such a result is hardly surprising, as the policy's spirit of pursuing justice at all costs, however noble, is unworkable if the pursuers cannot actually pay all costs. Limitations of budget, resources, or evidence inevitably demand some compromises, such as a tradeoff between seeking the maximum punishment for each offender and ensuring *some* punishment for all

---

1. *See* Memorandum from John Ashcroft, U.S. Att'y Gen., to All Federal Prosecutors, *Department Policy Concerning Charging Criminal Offenses, Disposition of Charges, and Sentencing* (Sept. 22, 2003), *reprinted in* 16 FED. SENT'G REP. 129 (2003) [hereinafter Ashcroft Memo]. This memorandum attracted a great deal of attention but followed earlier memoranda by preceding attorneys general who also sought to reduce discretionary bargaining. *See, e.g.*, Reno Bluesheet on Charging and Plea Decisions, from Janet Reno, U.S. Att'y Gen., to Holders of U.S. Attorneys' Manual, Title 9, *Principles of Federal Prosecution* (Oct. 12, 1993), *reprinted in* 6 FED. SENT'G REP. 352 (1994); Instructions from Richard Thornburgh, U.S. Att'y Gen., to U.S. Attorneys, *Plea Policy for Federal Prosecutors* (Mar. 13, 1989), *reprinted in* 1 FED. SENT'G REP. 421 (1989). The Ashcroft memo charged prosecutors with a duty to "charge and pursue the most serious, readily provable offense or offenses that are supported by the facts of the case," while avoiding charges "filed simply to exert leverage to induce a plea." Ashcroft Memo, *supra*, at 130. Prosecutors were also not to "request or accede to" a downward sentencing departure except under limited circumstances. *Id.* at 132.

2. *See, e.g.*, Gwen Filosa, *Ashcroft: New Police Powers Make U.S. Safer*, TIMES-PICAYUNE (New Orleans, La.), Sept. 26, 2003, at A7 ("'Like federal judges, federal prosecutors have an obligation to be fair, uniform and to be tough,' Ashcroft said. 'The amount of time in prison should depend on the severity of the crime, not on which judge happens to be sentencing him. This is about equal justice under the law.'"); Joe Gyan, Jr., *Ashcroft Seeks End to "Revolving Door" for Violent Criminals*, ADVOC. (Baton Rouge, La.), Sept. 26, 2003, at 9 ("Ashcroft also defended his get-tough policy, saying the intent is to bring 'uniformity and equality in justice across the board. . . . We need a policy that says people who commit serious crimes will be held accountable for those crimes . . . .'"); *cf.* George Fisher, *A Practice as Old as Justice Itself*, N.Y. TIMES, Sept. 28, 2003, § 4, at 11 ("There are good reasons to dislike plea bargains. Justice is dealt behind closed doors rather than in the open air of a courtroom. Prosecutors can hijack sentencing authority from judges, and criminals can get off with less than they deserve.").

3. *See, e.g.*, Albert W. Alschuler, *Disparity: The Normative and Empirical Failure of the Federal Guidelines*, 58 STAN. L. REV. 85, 113 (2005) ("Like similar directives from Attorneys General before him, Ashcroft's apparent restriction of plea bargaining made headlines. On the surface, it appeared to be a significant reform. The memorandum's exceptions, however, largely swallowed its rule."); *cf.* Mary Patrice Brown & Stevan E. Bunnell, *Negotiating Justice: Prosecutorial Perspectives on Federal Plea Bargaining in the District of Columbia*, 43 AM. CRIM. L. REV. 1063, 1078 (2006) ("Although the Ashcroft Memo did not have a major impact in D.C. on the content of plea agreements, it did tend to streamline plea negotiations . . . ."); Daniel Richman, *Institutional Coordination and Sentencing Reform*, 84 TEX. L. REV. 2055, 2068 (2006) ("It is far from clear whether the Ashcroft Memorandum had any effect on line prosecutors in U.S. Attorneys' offices.").

offenders. Because Ashcroft's memorandum could not articulate clear rules for how or when prosecutors could take these practical constraints into account, the result was more business as usual: unguided and unreviewable exercises of prosecutorial discretion in individual cases.[4]

Both the adoption and the failure of the DOJ antibargaining policy underscore a significant but often-ignored difference between the two major theories of criminal law, utilitarianism and retributivism.[5] The policy reflected the aspirations of a retributive-justice agenda, seeking fair, uniform, and deserved punishment for all offenders. Yet the policy also reflected the limitations of such an agenda, highlighting its restricted scope relative to that of the rival utilitarian theory. In a meaningful way, utilitarianism provides a complete theory of criminal justice, while retributivism apparently does not.

Utilitarianism, which bases punishment on the forward-looking goal of preventing future crime, is not only a *justificatory* theory explaining why

---

4. As Marc Miller pointed out at the time, the Ashcroft memo had more bark than bite, as its policies failed to "confront the major issues that control charging and screening decisions, including the wise use of prosecutorial resources (or in a more negative light, the need for triage) . . . ." Marc L. Miller, *Domination & Dissatisfaction: Prosecutors as Sentencers*, 56 STAN. L. REV. 1211, 1257 (2004). Miller concluded: "Even if every line lawyer takes the 2003 policies to heart (a practical impossibility), or applies them aggressively but only fully to address plea bargains (more plausibly), the tough-sounding 2003 policies include exceptions that any wise prosecutor (and there are many wise prosecutors) could drive a truck through." *Id.*

5. *See, e.g.*, JOSHUA DRESSLER, UNDERSTANDING CRIMINAL LAW § 2.03 (4th ed. 2006) (describing utilitarianism and retributivism); Russell L. Christopher, *The Prosecutor's Dilemma: Bargains and Punishments*, 72 FORDHAM L. REV. 93, 98–99 n.19 (2003) (providing support for the claim that "[t]he two main justifications of punishment are retributivism and consequentialism").

These are the two "major" or most commonly cited punishment theories, but there are others as well: "expressive" theories, "character" (or "virtue ethics," or "aretaic," or "Aristotelian") theories, and "restorative justice" theories, among others. *See, e.g.*, R.A. DUFF, TRIALS AND PUNISHMENTS 267 (1986) (describing "an ideal account of criminal punishment as a communicative enterprise in which we engage with the criminal"); Peter Arenella, *Character, Choice and Moral Agency: The Relevance of Character to Our Moral Culpability Judgments*, *in* CRIME, CULPABILITY, AND REMEDY 59 (Ellen Frankel Paul et al. eds., 1990) (focusing on the relevance of character to criminal law); John Braithwaite, *A Future Where Punishment Is Marginalized: Realistic or Utopian?*, 46 UCLA L. REV. 1727 (1999) (discussing a "restorative justice" model of criminal law); Kyron Huigens, *Virtue and Inculpation*, 108 HARV. L. REV. 1423 (1995) (describing an aretaic or virtue-based conception of criminal responsibility); Dan M. Kahan, *What Do Alternative Sanctions Mean?*, 63 U. CHI. L. REV. 591 (1996) (offering an "expressivist" view focusing on punishment as societal expression of condemnation). The existence of these other theories does not affect my analysis of retributivism, although that analysis may well apply to some of these other theories as well.

Further, it is significant to note that terms such as utilitarianism and retributivism do not (without further elaboration) describe full-fledged *theories* of punishment but rather mere aspects of a theory. *See* Kyron Huigens, *On Commonplace Punishment Theory*, 2005 U. CHI. LEGAL F. 437, 439–44. For example, deterrence and retribution may be *aims* of (or possible *justifications* for) punishment, but specifying the end(s) is merely one aspect of a theory. *See id.* In a sense, this article's claim is that a commitment to utilitarian ends can guide or at least imply the choice of means, and accordingly can form the basis of a fully developed theory more easily than a commitment to retributivist ends.

criminal punishment should exist, but also a *prescriptive* theory explaining how punishment institutions should work. The utilitarian agenda encompasses both the purposes and the practices of the criminal justice system, seeking in all cases and at all stages of the process to minimize or prevent social harms (in the most cost-effective way).[6] Because its project aims to specify both the content of criminal law and the proper means for enforcing that law, utilitarianism is not only a theory of *punishment,* but a complete theory of *criminal justice*.[7]

By contrast, retributivism, which adopts a backward-looking perspective focusing on the moral duty to punish past wrongdoing, is a justificatory theory, but seemingly not a prescriptive one.[8] It offers retribution as a justifying ideal but does not explain how legal institutions are supposed to make retribution real.[9] To the extent retributivism offers guidance about its own operation in practice, it speaks only to the content of criminal law rules, and not to their implementation.[10] Retributive

---

6. *See, e.g.*, Richard A. Bierschbach & Alex Stein, *Mediating Rules in Criminal Law*, 93 VA. L. REV. 1197, 1210 (2007) ("In contrast to retributivism, [utilitarian] deterrence is system-focused. Deterrence gauges the appropriateness of criminal liability and punishment by reference to the social costs of wrongful conduct and the social costs of measures necessary to prevent it. Harmful conduct should be criminalized and punished whenever doing so is the most cost-effective way to disincentivize that conduct.").

7. *See* LOUIS KAPLOW & STEVEN SHAVELL, FAIRNESS VERSUS WELFARE 310 n.35 (2002).

8. Indeed, even as a justificatory theory, retributivism may serve only a narrow function. According to one recent account, for example, retributivism does not provide an affirmative general basis for punishing but merely counters a *particular* objection to the imposition of punishment—namely, that it inflicts suffering. *See* Mitchell N. Berman, *Punishment and Justification*, 118 ETHICS (forthcoming Jan. 2008) (manuscript at 23, on file with author) ("Retributivists take themselves to be offering a tailored justification for punishment—tailored to the demand basis that punishment inflicts suffering. Unless they become persuaded that they must offer something more, the proposition that [retributivism] cannot establish that it is permissible all things considered for the state to inflict the deserved punishment looks more like an observation [retributivists] are permitted to accept than an objection to which they need respond."); *see also id*. at 50 ("A sympathetic or charitable understanding of contemporary retributivism requires one to recognize that its proponents are rarely purporting to provide an all-things-considered justification of punishment. More often, they are offering a tailored justification against the fact that punishment inflicts suffering . . . .").

9. *Cf.* Kenneth L. Avio, *Economic, Retributive and Contractarian Conceptions of Punishment*, 12 LAW & PHIL. 249, 263 (1993) ("[R]etributivist explanations . . . fail to indicate how many resources should be devoted to criminal justice, and how to allocate these resources between apprehension and punishment."); Mark D. White, *Retributivism in a World of Scarcity*, *in* THEORETICAL FOUNDATIONS OF LAW AND ECONOMICS (Mark D. White ed., Cambridge Univ. Press, forthcoming 2009) (manuscript at 2, on file with author) ("[R]etributivism does not take costs or resource constraints into account when it requires universal just punishments, so there is no room for optimization or trade-offs within its boundaries.").

10. *See* KAPLOW & SHAVELL, *supra* note 7, at 308–11 (criticizing retributive theory for its asserted failure to address practical enforcement issues of error and resource constraints); *cf.* Bierschbach & Stein, *supra* note 6, at 1206 (2007) ("For the retributivist, substantive rules and sanctioning rules are—or at least should be—generally more important than evidentiary rules.").

principles may identify what the law should criminalize,[11] and might even say something about the proper idealized level of punishment for those crimes relative to each other.[12] As to matters of application, however, retributivists tend to focus only on the resolution of individual (often hypothetical) cases where an offender's behavior is known or stipulated.[13]

---

11. Some critics of retributivism question whether it can do even this much; some supporters claim it does not purport that it can do even this much. *See* KAPLOW & SHAVELL, *supra* note 7, at 303 (pointing out that retributive "theory does not include a definition of what constitutes wrongful behavior deserving punishment"); *id.* at 304 n.24 ("That the retributive theory does not state a definition of wrongfulness deserving of punishment, and is therefore incomplete, is frequently noted." (citing JOHN BRAITHWAITE & PHILIP PETTIT, NOT JUST DESERTS: A REPUBLICAN THEORY OF CRIMINAL JUSTICE 36 (1990) (claiming retributive theory does "not offer guidance on what is to be criminalized"))); *see also* Dan Markel, *Against Mercy*, 88 MINN. L. REV. 1421, 1434 (2004) ("[T]here is no specifically retributivist contribution to the determination of what kind of behavior ought to be the subject of criminal legislation. The subject of criminal legislation is one that philosophers, economists, and judges dispute . . . but it is not one in which retributivists have a stake as retributivists . . . ."); *cf.* Mitchell N. Berman, *On the Moral Structure of White Collar Crime*, 5 OHIO. ST. J. CRIM. L. 301, 301–02, 326–27 (2007) (describing reasons for enacting legal rules that do not precisely track the contours of moral norms, and stating that legal theorists "will see farther if they more fully appreciate the extent to which the problem of legal drafting departs from the problem of identifying the underlying moral wrong that the law might be intended in the first instance to capture").

Even if the retributivist project is limited to justifying punishment rather than articulating what is supposed to be punished, retributivism does offer a basic criterion—moral wrongfulness—by which it is possible to assess whether criminalizing any particular conduct is prima facie acceptable. Notably, to the extent that (as the above critiques claim) even a strongly retributive substantive criminal law must inevitably take account of factors other than moral wrongfulness, this suggests that the consequentialist-retributivist approach I endorse in this article might be useful or even necessary at the stage of formulating, as well as implementing, criminal law. *See infra* Part II.C (discussing the impact of real-world concerns on the legislature, even if the legislature pursues a retributive agenda); Part III.C (defending consequentialist retributivism and noting its possible implications for the substance of criminal rules).

12. The extent to which retributive theory can do this is debated. *Compare, e.g.*, ANDREW VON HIRSCH, PAST OR FUTURE CRIMES: DESERVEDNESS AND DANGEROUSNESS IN THE SENTENCING OF CRIMINALS 63–76 (1985) (claiming potential to identify the relative severity of different offenses with reasonable precision), *with* David Dolinko, *Three Mistakes of Retributivism*, 39 UCLA L. REV. 1623, 1636–42 (1992) (critiquing work seeking to determine the proportionate level of punishment that different offenses deserve).

Theoretical matters aside, some empirical work asserts that lay intuitions about deserved punishment show fairly wide consensus regarding the relative seriousness of different offenses. *See, e.g.*, PAUL H. ROBINSON & JOHN M. DARLEY, JUSTICE, LIABILITY, AND BLAME: COMMUNITY VIEWS AND THE CRIMINAL LAW (1995); Paul H. Robinson & Robert Kurzban, *Concordance and Conflict in Intuitions of Justice*, 91 MINN. L. REV. 1829 (2007).

13. *Cf.* Douglas Husak, *What Moral Philosophers Might Learn from Criminal Theorists*, 36 RUTGERS L.J. 191, 192 (2004) ("Moral philosophers and criminal theorists tend to employ a similar methodology in deciding whether a given principle or doctrine is fair, just, or reasonable. . . . They frequently begin by describing examples, and ask how the behavior of the persons in these cases should be assessed. Answers or intuitions are thought to provide good reason to prefer one principle rather than another; they function as data that competing doctrines must struggle to accommodate."); Erik Luna, *Punishment Theory, Holism, and the Procedural Conception of Restorative Justice*, 2003 UTAH L. REV. 205, 206 (discussing the problems with "use of dubious arguments by punishment theorists and critics, employing surreal hypotheticals that have little resonance in perceived reality"); Paul H. Robinson, *The Role of Moral Philosophers in the Competition Between Deontological and*

Their theories offer no clear guidance as to more general issues of implementation in a system bound by resource constraints, imperfect information, and other limitations. They do not tell underfunded police and overworked prosecutors how to prioritize, or when, if ever, to compromise. If all offenders cannot be caught, how should the police set priorities in their enforcement agenda? Can, or should, prosecutors enter into plea bargains or give one offender less (or no) punishment for the sake of convicting another? Retributivism offers no obvious answers to these questions, or else seems to give unrealistic answers. Hence the failure of the Ashcroft policy, which shared retributive theory's focus on desert but also its myopia regarding practical constraints.

This Article seeks to determine whether any retributive theory can do any better than Ashcroft's memorandum in setting a feasible law enforcement agenda. Can a retributivist theory of punishment be translated, or fleshed out, into a complete theory of criminal justice—what this Article will call a theory of *real-world retribution*? (This Article will refer to such a theory as a "retributive" theory or a theory of retribution rather than a "retributivist" theory, as some might contend that retributivism per se properly should be viewed as a theory of punishment only. I make no claim that retributivists *must* expand their theory into some version of real-world retribution; the claim is merely that failure to do so limits the reach of retributivism as a theory.)

Such a theory would have to address two related aspects of the practical application of its substantive moral commitments.[14] First, the

---

Empirical Desert, 48 WM. & MARY L. REV. 1831, 1839 (2007) ("A standard analytic form, if not *the* standard form, among moral philosophers today is to use hypotheticals and philosophers' own intuitions about the proper resolution of the hypothetical as a basis for building moral principles.").

14. The application of a retributive-justice scheme might give rise not only to the *practical* concerns this Article addresses but also to conflicts between retribution and other *principled* commitments. For example, a commitment to certain aspects of procedural justice or fairness may sometimes frustrate the system's ability to impose punishment on those who deserve it. I, and others, have discussed elsewhere the relation between desert-based punishment and these other rules or principles. *See* PAUL H. ROBINSON & MICHAEL T. CAHILL, LAW WITHOUT JUSTICE: WHY CRIMINAL LAW DOESN'T GIVE PEOPLE WHAT THEY DESERVE 137–85 (2006) (discussing procedural rules that may conflict with desert); *see also* MICHAEL MOORE, PLACING BLAME: A GENERAL THEORY OF THE CRIMINAL LAW 739–95 (1997) (discussing limits liberal political theory might place on the reach of retributive criminal law); Bierschbach & Stein, *supra* note 6, at 1206–09 (discussing relationship between retributivism and evidentiary rules).

Ultimately, resolution of such conflicts between principled commitments depends on some decision about which principle merits priority in the abstract. This Article does not focus on such conflicts at the level of choosing among principles but on another set of issues: Presupposing *some* level of commitment to retributive justice, could even an adherent of retribution offer practical guidance about how to enforce its commands? And if so, what would that guidance look like? These issues would arise even if retributive justice were the *only* goal of the criminal system, for even then,

theory must offer guidance not only to the legislature charged with enacting criminal law, but also to the government officials—police and prosecutors—charged with enforcing that law.[15] A complete theory of justice should recognize, consider the relations between, and offer guidance to all of the players and institutions that comprise the legal system.[16]

Second, a real-world retribution theory must confront the scarcity of resources. Any justice system operates on a limited budget. Accordingly, any such system faces constraints in its ability to catch, prosecute, and punish all criminals.[17] Any theory that expects to make retribution real should have the capacity to explain how the system should set priorities within and between these tasks.[18]

If retributivism says nothing about how to make retribution real, it is incomplete as a theory of justice or of criminal law. A defense of retribution as a moral principle is not the same as a model of retributive justice, for it does not tell us how—or, perhaps, even *whether*—actual legal institutions should go about the task of imposing punishment that is morally deserved in the abstract.[19] Such a theory provides a moral philosophy but not a legal system.

---

resource and informational constraints would inevitably deny the system the ability to identify, prosecute, and punish all offenders.

15. The theory must also instruct the judges and juries who interpret the law, but because those actors exercise their authority case by case, any real-world retribution theory would presumably tell them to "do justice" in each case—i.e., to ensure, within the confines of their discretion, that each defendant is punished in accordance with his or her desert. Retributivism's lacuna is its failure to guide the enforcement authorities who have the power (and responsibility) to determine which cases are pursued to the fullest extent, which are subject to compromise, and which are dropped entirely.

16. As Michael Moore has put it, retributive justice is "not a role-specific" basis of punishment, but rather one that applies to judges as well as to legislators. MOORE, *supra* note 14, at 71. Presumably, it should apply to police and prosecutors as well. This Article seeks to explore the guidance that a retributive model (or different possible retributive models) would or could offer these other players. *Cf. id.* at 150 (referring to necessary transition "from moral theory to the political theory needed to justify the setting up of institutions" designed to achieve principled goals).

17. For an example of how law enforcement officials acknowledge resource constraints and the need to prioritize which aspects of their mission to pursue, see generally FEDERAL BUREAU OF INVESTIGATION, STRATEGIC PLAN 2004–2009, *available at* http://www.fbi.gov/publications/strategicplan/strategicplanfull.pdf.

18. For another recent effort to elaborate and evaluate possible ways to apply retributive principles in a world of limited resources, see generally White, *supra* note 9.

19. *See* Douglas Husak, *Holistic Retributivism*, 88 CAL. L. REV. 991, 996 (2000) ("[T]he demonstration that some good is produced by giving persons what they deserve does not suffice to justify the institution of punishment—even for retributivists. . . . Retributivists must show not only that giving culpable wrongdoers what they deserve is intrinsically valuable, but also that it is sufficiently valuable to offset . . . the drawbacks of punishment—negative values that inevitably are produced when an institution of punishment is created."); *cf.* Guyora Binder, *Punishment Theory: Moral or Political?*, 5 BUFF. CRIM. L. REV. 321, 321–22 (2002) ("Because punishment is part of a system of

Such incompleteness is not only problematic as a matter of theoretical nicety but presents serious moral and practical concerns for anyone committed to the principle of punishment according to desert. In the absence of any well-developed explanation of how best to pursue retributive goals in a world of limited resources, law enforcers are left with two options that any committed retributivist should find highly problematic: either rely on post hoc (and ad hoc) case-by-case enforcement discretion, or else simply give up on desert-based punishment and adopt an explicitly utilitarian agenda.[20] Accordingly, the project of setting out a consistent and principled scheme for making retribution real would appear to have significant value.

Unfortunately, deriving a set of real-world enforcement priorities from retributive theory is no easy task. Indeed, retributivism's orientation as a justificatory theory seems to undermine, or at least complicate, its own capacity to offer prescriptive guidance. Retributive theorists often justify punishment by reference to the moral duty (not just the authority) to punish those who deserve it.[21] That understanding of punishment as an unwavering obligation indicates that the retributivist perspective does not lend itself to the practical compromises often necessary in the real world, where we cannot punish all wrongdoers to the full extent of their moral desert.

This Article explores possible ways to flesh out a retributive theory of punishment that might, at least in principle, offer guidance to real-world decisionmakers about how to enforce criminal law, as well as how to write it. The present project does not aim to provide a solution for every specific situation these decisionmakers might confront. Rather, the task here is to ask whether a real-world retribution theory is even *possible*, to determine what general approaches such a theory might prescribe, and to evaluate those approaches in terms of their general practical ramifications. At the very least, it is a call to retributive theorists to pay more attention to real-world implementation issues. One would think that such theorists would be eager to improve the potential to realize retributive principles by offering guidance to all relevant players in the criminal justice system,

---

institutional authority, it is not amenable to a simple moral analysis."); *id.* at 333 ("[I]t makes more sense to think of the justification of punishment as a problem of political theory than as a problem of ethics.").

20. *See, e.g.*, White, *supra* note 9, at 21 ("While retributivism is strictly a theory of punishment, few retributivists would be comfortable with wide latitude and discretion on the part of prosecutors, and most may endorse the extension of justice-based constraints to the actions of prosecutors as well.").

21. *See* sources cited *infra* Part I.A.

rather than leaving critical matters of implementation unexplored and unexplained, thereby essentially surrendering a great deal of ground at the outset.

In some respects, then, this Article is strongly critical of retributivism, since the Article's premise is that retributivist theory, as typically formulated and advanced, is unsatisfying or incomplete. At the same time, the Article is fundamentally sympathetic to the retributivist agenda, as it assumes that the principle of desert-based punishment[22] merits at least some presumptive respect.[23] I point to three sources for evidence of the significance of retributive justice. First, though perhaps least important, is the vast scholarly literature advancing and defending retributive theory. As between retributive and utilitarian perspectives, retributivism is often recognized as occupying the dominant position, at least within the scholarly literature, over the last several decades.[24] Second, philosophers are not the only ones who value retributive justice; empirical studies indicate that most people consider desert-based punishment to be the fundamental goal of criminal law and intuitively respond to actual cases in desert-based terms.[25] Finally, at least some degree of commitment to a

---

22. This Article refers interchangeably to "retributive" punishment and "desert-based" punishment. It is reasonably common for commentators to equate the two terms in such fashion, though some take issue with this. *See* Dan Markel, *State, Be Not Proud: A Retributivist Defense of the Commutation of Death Row and the Abolition of the Death Penalty*, 40 HARV. C.R.-C.L. L. REV. 407, 426–27 (noting but criticizing the typical understanding of retribution as grounded in moral desert).

23. *See* Lawrence H. Davis, *They Deserve to Suffer*, 32 ANALYSIS 136, 139 (1972) (concluding it "very likely" that imposing deserved punishment has value "on the grounds (a) that there is no convincing argument against it, and (b) that inclination to believe it seems very widespread among the people whose moral intuitions constitute the main data we have for settling questions of value").

One who attaches no value to the pursuit of retributive justice would obviously see nothing worthwhile in an effort to present a scheme for pursuing it effectively. For example, committed utilitarians enjoy deriding retributivism's lack of practicality, but they make no attempt to explain whether or how the theory could be made more practical since they have no interest in making it so. *See generally, e.g.*, KAPLOW & SHAVELL, *supra* note 7, at 308–11, 326–29, 339–41 (critiquing the practical shortcomings of theoretical retributivist literature for the sake of advancing the authors' welfarist project).

24. *See* MOORE, *supra* note 14, at 83 & n.1 (noting, and citing authorities supporting, retributivism's "resurgence among the legal theorists of the past thirty years"); Christopher, *supra* note 5, at 97 n.13 (collecting authorities referring to retributivism as the dominant contemporary justification of punishment).

25. Studies indicate that the typical person's reaction to criminal cases is driven more by desert-based considerations than by utilitarian considerations. *See, e.g.*, Kevin M. Carlsmith et al., *Why Do We Punish? Deterrence and Just Deserts as Motives for Punishment*, 83 J. PERSONALITY & SOC. PSYCHOL. 284 (2002) (empirical study suggesting that laypersons consider neither difficulty of detection nor publicity—classic deterrence factors—in setting punishment); John M. Darley et al., *Incapacitation and Just Deserts as Motives for Punishment*, 24 LAW & HUM. BEHAV. 659 (2000) (empirical study suggesting that, in setting punishment, laypersons emphasize culpability over the correlation to future criminality, especially when protective mechanisms other than the criminal justice system are available); Cass R. Sunstein et al., *Do People Want Optimal Deterrence?*, 29 J. LEGAL

notion of desert or moral blame is almost universally evident within the history and traditions of actual criminal systems.[26]

Part I of this Article looks to retributive theory to see what it says or suggests about how to enforce the law, either in terms of imposing constraints (telling officials what they may not do) or offering affirmative guidance (telling officials what they should do). Different versions of retributive thinking imply different options or mandates for law enforcement. Canvassing and generalizing from the literature, Part I describes three possible implementation approaches that a legitimately retribution-oriented theory might adopt. Part I also considers the position that retributivism simply has nothing to say about enforcement, and notes two possible approaches grounded in that position.

Part II discusses in more detail what a real-world retribution scheme might look like under each of the three versions of a retribution-based criminal justice theory sketched in Part I. The discussion explores what practices each theory might endorse or prohibit for police and prosecutors seeking to do justice as best they can. In doing so, Part II raises practical considerations and undesirable potential consequences that might complicate, or militate against, the implementation of these general options.

Part III evaluates the options for real-world retribution, both conceptually and practically, in light of the obstacles and drawbacks that

---

STUD. 237 (2000) (discussing two reported experiments suggesting that people do not spontaneously think in terms of optimal deterrence, and that people would have objections to policies based on the goal of optimal deterrence); Mark Warr & Mark Stafford, *Public Goals of Punishment and Support for the Death Penalty*, 21 J. RES. CRIME & DELINQ. 95, 99–101 (1984) (describing citizens' identification of retribution as the goal that should govern the criminal justice system).

For good or ill, such retributive judgments of blameworthiness may be natural, and perhaps even inevitable, based on human biology. *Cf.* J.L. Mackie, *Morality and the Retributive Emotions*, 1 CRIM. JUST. ETHICS 3 (1982) (pointing to biological bases for moral intuitions about retribution); Lawrence M. Solan, *Cognitive Foundations of the Impulse to Blame*, 68 BROOK. L. REV. 1003 (2003) (discussing features of human psychology contributing to our natural tendency to impose moral judgments of blame).

Some studies also suggest that people value achieving a desert-based result more highly than they value procedural fairness. *See, e.g.*, Linda J. Skitka & David A. Houston, *When Due Process Is of No Consequence: Moral Mandates and Presumed Defendant Guilt or Innocence*, 14 SOC. JUST. RES. 305 (2001); *see also* Linda J. Skitka, *Do the Means Always Justify the Ends, or Do the Ends Sometimes Justify the Means? A Value Protection Model of Justice Reasoning*, 28 PERSONALITY & SOC. PSYCHOL. BULL. 588 (2002); Linda J. Skitka & Elizabeth Mullen, *The Dark Side of Moral Conviction*, 2 ANALYSES OF SOC. ISSUES & PUB. POL'Y 35 (2002); Linda J. Skitka & Elizabeth Mullen, *Understanding Judgments of Fairness in a Real-World Political Context: A Test of the Value Protection Model of Justice Reasoning*, 28 PERSONALITY & SOC. PSYCHOL. BULL. 1419 (2002).

26. *See, e.g.*, Paul H. Robinson, *The Criminal-Civil Distinction and the Utility of Desert*, 76 B.U. L. REV. 201 (1996) (noting that all, or nearly all, legal systems appear to recognize the distinction between civil and criminal law and see the latter as uniquely concerned with moral condemnation).

surface in Part II. The Article concludes that only one option seems to offer a genuine and (relatively) workable vision of how to implement retributive justice in the real world. That approach, "consequentialist retributivism" (CR), would view retributive punishment as a goal to maximize rather than a categorical *ex ante* commitment, as other approaches would have it. Interestingly, though the CR approach has considerable intuitive appeal in this context, it has thus far apparently received *no* explicit, sustained defense in the scholarly literature. One significant contribution this Article seeks to make is to provide such a defense of the CR approach, which might be the only version of retributive justice possible, or at least worth pursuing, in the real world.

I. SURVEYING THE OPTIONS

Retributive theorists tend to say little about matters of implementation, but one can make some inferences about the application of retributive principles to law enforcement authorities based on what the theorists say about retribution's general demands. These discussions imply at least five possible views on how to apply retributive justice in real-world cases: three variations involve how such an application might or should be done, and two variations involve how it cannot or should not be done. This Part first discusses the three views that allow for some form of real-world retribution scheme, in order of decreasing strictness (i.e., starting with the view under which retributive demands place the most severe restrictions on the options available to officials). The options are as follows: (1) retributive justice imposes a moral duty to punish all (identified) offenders to the full extent their moral desert demands, with no exceptions;[27] (2) the same offender-specific duty exists as in (1), but some exceptions are allowed;[28] and (3) the duty to punish does not demand punishment in every particular case but instead imposes a general obligation to maximize the overall amount of (deserved) punishment imposed.[29]

It is also possible that, in the end, retributivism is only a justificatory theory and not a prescriptive one (nor capable of being developed into or forming the basis of a prescriptive one). If so, at least two interpretations of that conclusion are possible, offering two additional visions of the relation between retributivism and the real world. The first interpretation, which itself might generate two distinct prescriptions for law enforcement,

---

27.   *See infra* Part I.A.
28.   *See infra* Part I.B.
29.   *See infra* Part I.C.

is that (4) retributivism makes no demands regarding implementation, so officials should either (4a) pursue an explicitly utilitarian agenda, or (4b) do whatever they think is best (perhaps applying one of the three putatively retribution-oriented models above, or perhaps simply using their discretion ad hoc to resolve individual cases or respond to enforcement situations as they arise). The second interpretation is that (5) retributivism offers no guidance about implementation, so it is altogether useless or inappropriate as an influence on any aspect of criminal law, including substantive law; in other words, retributivism and the real world are incompatible, so the real world wins.[30]

This Part sets out, in turn, each of the five possible positions described above and surveys the available literature for examples of thinkers who either explicitly adopt that view, tacitly hint at it, or describe it as a plausible version (or extension) of retributivism while declining to support it.

A. The Absolutist Model: Avoid All Failures of Desert

The strictest version of retributivism sees its moral rules as clear and absolute mandates applying to each person in all situations. Such moral rules would include not only the imperatives captured by criminal law itself but also, and significantly for this discussion, the moral duty to punish all those (and only those) who deserve punishment. Because this version is rooted in *a priori* moral commitments, it is often described as a *deontological* understanding; because its commands apply to each individual actor and action, rather than seeking to maximize "the good" overall or in the long run, it is sometimes described as an *agent-relative* approach.[31]

Retributivist literature is rife with references to the principle of desert-based punishment as a moral duty and to the corresponding claim that the retributive principle does not merely authorize punishment but affirmatively calls for its imposition on those who deserve it.[32] The view

---

30. This is not to say that all retributive thinkers, if forced to decide between the two, would necessarily choose the real world over the theory.

31. *See* MOORE, *supra* note 14, at 156 ("The 'deontological' or 'agent-relative' retributivist regards the act of punishing the guilty as categorically demanded on each occasion, considered separately.").

32. A similar claim, or perhaps the same one in different dress, asserts that imposition of deserved punishment is not merely a typical policy objective but a dictate of (retributive) *justice*, making its pursuit obligatory in a way that pursuit of other goals or goods is not. *See* Berman, *supra* note 8, at 28 ("Those who doubt that the good of a wrongdoer suffering (on account of and in proportion to his blameworthy wrongdoing) provides a reason of sufficient weight, or of the right sort,

dates back at least as far as Kant[33] and Hegel,[34] and it continues to appear in contemporary writing as well.[35] Indeed, some version of this claim

---

to render permissible the [imposition of punishment] might think otherwise once justice is introduced as a mediating norm or value. Retributivists can now argue that a wrongdoer's suffering is demanded by justice, and therefore that what provides reason of sufficient weight to justify [punishment] is (at least in part) the state's obligation to see that justice is done.").

33. *See* IMMANUEL KANT, THE METAPHYSICS OF MORALS 105 (Mary Gregor ed. & trans., Cambridge Univ. Press 1996) (1797) (describing retributive punishment as a "categorical imperative" that "must always be inflicted upon [a criminal] only *because he has committed a crime*"); *id.* at 109–10 (asserting that "failure to punish . . . is the greatest wrong" a state can inflict against its citizens).

34. *See* G.W.F. HEGEL, ELEMENTS OF THE PHILOSOPHY OF RIGHT 127 (Allen W. Wood ed., H.B. Nisbet trans., Cambridge Univ. Press 1991) (1821) ("Both the nature of crime and the criminal's own will require that the infringement for which he is responsible should be cancelled."); *see also* IGOR PRIMORATZ, JUSTIFYING LEGAL PUNISHMENT 74 (1997) ("A full-fledged retributive theory sees in the offense committed not only the ground of the right to punish, but the source of the duty to punish as well. In this respect, Hegel is a true retributivist: his doctrine . . . is primarily an attempt to establish the second of the two claims.").

35. *See, e.g.*, Stephen P. Garvey, *Is It Wrong to Commute Death Row? Retribution, Atonement, and Mercy*, 82 N.C. L. REV. 1319, 1324 (2004) ("Retributive justice obligates the state to punish an offender because and to the extent, but only to the extent, he deserves to be punished. . . . [T]he state cannot shirk its obligation to do justice."); *see also infra* note 42 (citing Michael Moore's references to duty to punish).

The extent to which contemporary retributivist thinkers continue to adhere to the classic absolutist position is unclear. Some recent general ascriptions of the absolutist position to retributivist thinking come from critics who seek to point out its difficulties. *See, e.g.*, KAPLOW & SHAVELL, *supra* note 7, at 295 n.9 (referring to, and providing citations to support, the "standard retributive view under which wrongdoers should be punished"); *id.* at 297 n.12 (providing additional references to twentieth-century retributivist literature); Christopher, *supra* note 5, at 125–26 (discussing and supporting claim that "[o]ne of retributivism's most important principles is the duty to punish culpable wrongdoers"). Such critics may be connecting retributivism with absolutism for the sake of lampooning retributivism more easily.

Other references to the absolutist duty to punish appear within general summaries of the retributivist position, which are not always put forth by scholars who are themselves retributivists. *See, e.g.*, PRIMORATZ, *supra* note 34, at 148 ("The offense committed is the sole ground of the state's right and duty to punish . . . . 'Justice' and 'just deserts' are not meant merely negatively, as constraints, but also positively, as demands for punishment of the guilty and the full measure of proportion between the punishment and the offense."); Stanley I. Benn, *Punishment*, *in* 7 ENCYCLOPEDIA OF PHILOSOPHY 29, 30 (Paul Edwards ed., 1967) ("The most thoroughgoing retributivists . . . maintain that the punishment of crime is right in itself, that it is fitting that the guilty should suffer, and that justice, or the moral order, requires the institution of punishment."); Kent Greenawalt, *Punishment*, *in* 4 ENCYCLOPEDIA OF CRIME AND JUSTICE 1336, 1338 (Sanford H. Kadish ed., 1983) (describing retributive justification as holding that "[s]ociety not only has a right to punish a person who deserves punishment, but it has a duty to do so").

Accordingly, the absolutist view may be more a relic of intellectual history than a centerpiece of modern retributivist thought. Even if that is true, discussion of this version of the retributivist position remains useful for two reasons. First, in this context, situating absolutism within a broader array of genuinely retributive models makes clear that the absolutist position describes just one version of retributivism rather than the whole of it, thereby defusing some of the antiretributivists' general attack even while agreeing with its validity as to this particular vision of retributive justice.

Second, although explicit endorsement of the absolutist position within retributivist commentary seems increasingly uncommon, many retributivist writers do not explicitly endorse any other view either. Although some contemporary retributivist writers have directly, and laudably, confronted the fundamental nature of the retributivist commitment—and this Article tries to incorporate and advance

seems necessary to a retributivist justification of punishment, as the weaker claim that wrongdoing only *permits* but does not demand punishment—which also surfaces in the literature at times[36]—would not suffice to justify punishment, for many practices may be permissible without being desirable (much less required).[37]

In theory, the duty-to-punish view creates a clear affirmative goal for the application of the retributive principle: punish everybody who deserves it, to the full extent of their desert. In practice, however, limitations of resources, evidence, and knowledge make this goal impossible to achieve. Even if we wanted to punish every wrongdoer or offender, we could not find them all or muster sufficient proof of their crimes.

---

their work—others take that commitment as given and spend their time analyzing the nature and demands of moral desert in specific situations. If nothing else, this Article might encourage such retributive thinkers to reflect on the nature of their commitment to the retributive principle and the extent to which different understandings of that commitment might bear on the formulation or application of retributive rules in particular contexts.

36. *See, e.g.*, KAPLOW & SHAVELL, *supra* note 7, at 295 n.9 (noting and citing sources for the alternative view "that holds only that punishment of wrongdoers is *permissible*"—though questioning, in some cases, whether ascription of this view to the relevant commentators is accurate).

37. *See id.* (noting that a theory of retribution as allowing but not requiring punishment "must be complemented by another theory that indicates when punishment that is permissible should in fact be meted out; ordinarily, it seems to be understood that some form of consequentialist theory would be employed to do so"); *cf.* J.L. Mackie, *Retributivism: A Test Case for Ethical Objectivity*, *in* PHILOSOPHY OF LAW 677, 678 (Joel Feinberg & Hyman Gross eds., 4th ed. 1991) [hereinafter PHILOSOPHY OF LAW] (noting that "negative" retributive claims about avoiding punishment of the innocent, as opposed to accounts of retribution as an affirmative basis for punishing, "take us no distance at all towards a positive reason for punishing the guilty just because they are guilty").

Mitchell Berman offers a distinct account of retributivism whose ultimate effect is similar—i.e., defending the acceptability of punishment without offering an affirmative basis for imposing it. Berman contends that retributivism provides a (partial) justification of punishment—not by defining a moral duty, but rather by offering a "cancelling condition" that negates a particular objection to punishment (namely, that it inflicts suffering). *See* Berman, *supra* note 8, at 10–11 (describing the notion of a cancelling condition, which Berman borrows from the work of Joseph Raz); *id.* at 42–43 (summarizing position that retributivism offers partial justification for punishment and does so by cancellation). Accordingly, retributivism, standing alone, does not necessarily provide a basis for imposing punishment in the face of other objections (such as its cost)—though, somewhat in keeping with the CR perspective, Berman points out that the good of achieving desert can be one consideration within a broader consequentialist analysis of punishment's costs and benefits. *See id.* at 51 ("[R]etributivism lacks the resources to provide . . . even a tailored justification for peripheral cases of punishment—those in which the imposition of punishment rests on a mistake regarding the defendant's responsibility for the offense. So if peripheral cases are to be justified against the fact that punishment inflicts suffering, and if punishment simpliciter is to be justified all-things-considered, then consequentialist considerations must be relied upon."); *id.* at 41–42 (noting that "good consequences can include (what I have supposed to be) the good of bringing about a state of affairs in which deserved suffering obtains").

Theorists rarely confront this problem head-on or describe how best to pursue the retributive obligation in practice,[38] so any effort to flesh out the theory in this regard involves some extrapolation. In its application, an absolutist deontological version of retributivism does not provide a positive objective so much as a negative injunction or constraint: where (we know[39]) an identifiable person deserves punishment, we are obliged to punish that person.[40] At the same time, we are obliged *not* to punish a person where (we know) she deserves no punishment, or where (we know) the punishment exceeds what she does deserve.[41]

The general orientation of this version is a concern with punishing individual offenders (as opposed to offenses). Because it sees the duty to punish as categorical, this vision does not offer a clear means for prioritizing punishment of more serious offenses over less serious ones— the *amount* of punishment would vary according to desert, but the *duty* to punish would presumably apply with equal force in all cases. For example, although a murderer would ultimately receive more punishment than a shoplifter based on their relative blameworthiness, the categorical duty to impose punishment in all cases of blameworthy wrongdoing would seemingly exist as to both in equal measure.

---

38. *See, e.g.*, KAPLOW & SHAVELL, *supra* note 7, at 308◘10 (pointing to real-world problems of adjudicative error and resource limitations and criticizing retributive theory for failing to address these issues).

39. Here, and in the following textual sentence, I place the phrase "we know" in parentheses because here, as generally with retributivist rules, there is some question as to what kind of mental state or level of awareness would suffice to make one's breach of a duty morally culpable, therefore itself wrongful. Punishing an innocent person on purpose would clearly be forbidden; doing so by accident might be improper or unfortunate but would not be morally blameworthy. This issue is relevant to the more concrete question of what safeguards an actual criminal justice system would have to erect, or what kinds of behavior an individual official may or must perform or avoid, to satisfy the demands of the deontological view. *See infra* Part II.A.1 (discussing debate as to whether absolutist deontology even permits the formation of punishment institutions, given that some errors are bound to occur).

40. *See, e.g.*, George P. Fletcher, *The Place of Victims in the Theory of Retribution*, 3 BUFF. CRIM. L. REV. 51, 61 (1999) ("Kant's intuition represents a fundamental plank in the tradition of retributive justice. The reason we must punish is to avoid liability for *impunidad*—for allowing criminals to go [un]punished."); *see also id.* at 62 (noting that the obligation to punish all offenders— to avoid systemic complicity in the evil of "*impunidad*"—arises only after institutions of punishment have been developed, and relates to their implementation).

41. Even "weak" or "negative" retributive theories, which do not claim an affirmative duty to punish the deserving, recognize an obligation *not* to punish the innocent, and not to punish beyond what is deserved. *See supra* notes 36–37 (referring to such theories and the sources discussing them); Mackie, *supra* note 37, at 678 ("Negative retributivism (for which 'minimalism' may be another name) holds that those who are not guilty must not . . . be punished, that the absence of a crime morally requires the non-infliction of a penalty. There is also a quantitative variant of negative retributivism, that even if someone is guilty of a crime it is wrong to punish him more severely than is proportional to the crime.").

B. *The Threshold Model: Avoid Serious Failures of Desert*

A second version of retributivism shares the first version's position that imposing desert-based punishment is a moral duty, but it considers that duty less categorical or universal than the absolutist approach. Under this view, sometimes called *threshold deontology*, moral duties are generally binding *but* may be violated if adherence would lead to some particularly bad result.[42] As with the absolutist approach, theorists sometimes provide tacit rather than overt support for this general view,[43] though the threshold view does have the explicit support of at least one major retributivist thinker: Michael Moore.[44]

---

42. *See* MOORE, *supra* note 14, at 1568 n.13 ("A 'threshold' deontologist refuses to violate a categorical norm of morality until not doing so produces sufficient bad consequences as to pass some threshold—then, he will override such categorical norms."). I confess that I am not sure what "categorical" is supposed to mean in the preceding quotation, as the whole idea of introducing a threshold is that the relevant moral norms are *not* "categorical," at least as I understand that term. In any event, it is certainly apparent from Moore's statements elsewhere that, as a threshold deontologist, he still takes the duty to punish very seriously. *See id.* at 91 ("[Moral desert] gives society more than merely a *right* to punish culpable offenders. . . . For a retributivist, the moral responsibility of an offender also gives society the *duty* to punish."); *id.* at 154 ("The desert of offenders certainly gives [state] officials *permission* to punish offenders . . . . But retributivism goes further. As a theory of a kind of justice, it *obligates* us to seek retribution through the punishment of the guilty."); *id.* at 156 ("[T]he retributivist regards the punishment of the guilty to be categorically imperative whenever the opportunity to give such punishment presents itself.").

43. For example, Igor Primoratz offers the following discussion of the potential for other considerations (specifically, mercy) to override the duty to punish:

> Sometimes, when the offense committed is not very grave, and the suffering . . . that would be brought about by meting the full measure of deserved punishment would be very, very great, the call for mercy will override the duty to punish, and the penalty will be considerably reduced. There will also be cases in which the facts calling for mercy will be so weighty . . . that the final decision will be a full pardon.
>
> [This] view of the duty to punish is actually the dominant one in retributivism. . . . The duty is not seen as an absolute one, allowing for no exception and no mitigation, whatever the circumstances, but rather as a duty of paramount, but not absolute importance, which sometimes gives way to mercy and pardon.

PRIMORATZ, *supra* note 34, at 110 (citing, as support for the claim that this view is "dominant," H.J. McCloskey, *A Non-Utilitarian Approach to Punishment*, *in* CONTEMPORARY UTILITARIANISM 239, 255–57 (Michael D. Bayles ed., 1968); Alwynne Smart, *Mercy*, *in* THE PHILOSOPHY OF PUNISHMENT 212, 221–27 (H.B. Acton ed., 1969)).

44. *See* MOORE, *supra* note 14, at 158 (describing himself as a threshold deontologist); *id.* at 719–24 (sketching and giving support for threshold version of deontology). *But cf. infra* notes 146–49 (discussing Moore's recent support for a consequentialist perspective as to the general design of criminal-justice institutions).

Other theorists, dealing with matters other than criminal law specifically, have also implied their support for the threshold-deontologist view. *See, e.g.*, CHARLES FRIED, RIGHT AND WRONG 10 (1978); THOMAS NAGEL, MORTAL QUESTIONS 53–74 (1979); ROBERT NOZICK, ANARCHY, STATE, AND UTOPIA 30 n.* (1974).

The very task of setting, or admitting the possibility of, a threshold suggests that this model finds the duty to impose desert-based punishment to vary in its binding effect. Typically, retributive concerns will outweigh other (i.e., consequentialist) considerations, but as the relevant bad consequences of following the desert principle become serious enough, they may trump retribution. Moore's discussion of this dynamic looks primarily at the situation where retributive goals mandate *refraining* from punishment, but the good consequences that would follow from punishment are *large* enough to outweigh the desert principle. This is essentially the "ticking-time-bomb" scenario where torture, which retributivism would otherwise prohibit, seems attractive for consequentialist reasons.[45]

By the same logic, it may also be true that in some cases the retributive concerns *favor* punishment but are *small* enough that significant consequential costs will trump them, militating decisively against punishment. Such variations in the level of obligation to punish would presumably track the moral gravity of the offense in question: we may forgo punishment for less serious crimes, but must punish more serious ones.[46] Moore himself sometimes hints at this point,[47] and on at least one occasion explicitly states that a balancing of retribution and other concerns might disfavor retributive goals not only because other costs are large, but

---

45. *See, e.g.*, MOORE, *supra* note 14, at 719 ("It just is not true that one should allow a nuclear war rather than killing or torturing an innocent person. It is not even true that one should allow the destruction of a sizable city by a terrorist nuclear device rather than kill or torture an innocent person."). The original article in which Moore provides this discussion of threshold deontology was written to address the specific question of the permissibility of torture. *See* Michael S. Moore, *Torture and the Balance of Evils*, 23 ISR. L. REV. 280 (1989).

46. Primoratz suggests such a scheme in discussing the scope of retribution as a duty to punish:

Admittedly, retribution as a positive principle seems to have greater force in cases of more serious crimes than in those of petty offenses. In cases of the latter sort we may be more inclined to let the offender "get away with it" if no effects of deterrence are to be achieved by punishment. If so, we should be able to account for the difference in our moral judgment in terms of a general theory of the gravity of offenses which would justify the different treatment: a theory which would allow, and even call for, the application of the principle of just retribution regarding the most heinous crimes, and others somewhat less abominable, and then others less abominable still—and then suspend the principle at a certain point on the scale of crimes and replace it by considerations of an entirely different nature: those of deterrence.

PRIMORATZ, *supra* note 34, at 149.

47. *Cf.* MOORE, *supra* note 14, at 187 (referring to "the (minor) retributive good of punishing minor moral wrongs"); *id.* at 724 (noting that "where the threshold lies depends in part on the degree of wrong done"—though in this context, "wrong" appears to refer to the violation of the retributivist principle itself, rather than a wrongful act meriting retributive punishment; whether Moore views the two as equivalent is not entirely clear).

because the moral wrong in question is "trivial in the degree of [its] wrongfulness."[48]

Further, if the relative seriousness of different crimes bears on the strength of the duty to punish in this way, presumably that significance should hold (at least to some degree) above and below the threshold, rather than merely at the threshold point.[49] That is, if the duty to punish tracks the severity of the wrong, even as to cases that exceed the threshold level and warrant punishment, this perspective should endorse prioritizing more serious crimes or wrongs over less serious ones.

To summarize, the threshold deontological view of retribution suggests a second possible model of real-world retribution. For the sake of convenience, this Article will call it the "threshold model." An important caveat is in order here, however. In part because no one has previously elaborated the threshold deontological view along quite these lines, it is not clear whether that view would necessarily dictate or even endorse all aspects of the "threshold model" presented here.[50] In other words, though I take the "threshold model" to at least be *consistent* with threshold deontology, my project here is not to provide a scheme that maps precisely onto the threshold theory or is implicit in the views of any or all of that theory's adherents.[51] Rather, the current aim is merely to offer the model as a plausible alternative candidate for advancing some version of retributive justice in the real world.

This second model differs from the absolutist view in two ways. First, though it resembles absolutist deontology in recognizing a duty to punish, this model allows that in some (presumably rare) cases, other considerations could trump that duty and permit a violation of the desert

---

48. *Id.* at 663 ("For types of crimes that are relatively trivial in the degree of their wrongfulness, such as jaywalking, one might well think that public resources should be better spent on more important issues.").

49. *Cf. id.* at 723 ("[F]or a threshold deontologist consequences always 'count'. . . . Even before [the] threshold was reached, consequences counted but were of insufficient weight to determine the rightness of actions.").

50. Of course, the same might also be true for the absolutist and consequentialist models this Article sketches, but the risk of mischaracterization seems lower for those two models—for the absolutist model, because its moral commitments are so clear, and for the consequentialist model, because it apparently has no devotees to offend.

51. To note one adherent of the theory in particular, I have pointed to statements or positions offered by Michael Moore, a threshold deontologist, as support for the relationship between threshold deontology and the threshold model I present. *See supra* notes 47–49. It is possible, though, that Moore does not hold those positions *by virtue of* his being a threshold deontologist, but rather holds them independently of, or in addition to, his commitment to threshold deontology. It is also entirely possible that, notwithstanding both his general commitment to threshold deontology and his other statements, Moore would completely disavow any support for the "threshold model" this Article describes.

principle. Thus, where the absolutist model would categorically forbid intentional infliction of injury or death on an innocent person under any circumstances, the threshold model might allow torturing or even killing if doing so would, for example, save an entire city from destruction.[52] Second, this model would appear to *scale* the magnitude of that duty according to the magnitude of the moral wrongdoing involved in an offense.[53] Unlike the absolutist view that seeks to punish the greatest possible number of offenders, then, the threshold view might more narrowly target (and demand) punishment of the greatest number of *serious* offenders, with less serious ones receiving lower priority.

C. Consequentialist Retributivism: Maximize Desert

A third option abandons the agent-relative view of moral duties and adopts an *agent-neutral* understanding. On this account, the nature of the moral call for desert-based punishment does not demand punishment of each individual offender but rather sets the goal of maximizing the *total* amount of desert-based punishment,[54] even if this means sacrificing deserved punishment in some cases for the sake of pursuing it in others. Because this approach interprets desert-based punishment as a "good" to maximize, rather than an *ex ante* and inviolate command, it has been described as *consequentialist* (as opposed to deontological) retributivism,[55] which this Article will sometimes abbreviate as "CR."

Significantly, the CR perspective does not merely seek to find room for retribution as one component of a standard consequentialist scheme, as some defenses of desert-based punishment do, for those accounts defend the practice as *instrumentally* valuable insofar as it generates other good consequences (such as making people happy or preventing crime).[56] The

---

52. *See supra* note 45.

53. The relative moral gravity of different offenses might turn on various factors, but two prominent factors would surely be the level of culpability required (so that intentional wrongdoing would be more serious than reckless wrongdoing) and the extent of the harm caused or risked (so that causing death would be more serious than causing injury). Other factors or subfactors, such as whether harm is actually caused or "merely" risked, might be more controversial in terms of their bearing on the seriousness of the offense.

54. Obviously, maximizing *deserved* punishment is distinct from maximizing punishment altogether. Any retributivist system would seek to avoid undeserved punishment at least as strenuously as it would seek to impose deserved punishment, if not more so. *See infra* note 87 (discussing the relative significance, under a retributivist view, of punishing the guilty versus not punishing the innocent). Thus, any undeserved punishment would not only fail to increase the total deserved punishment but would reduce the total.

55. *See* MOORE, *supra* note 14, at 156–59.

56. *See, e.g.*, A. Mitchell Polinsky & Steven Shavell, *The Fairness of Sanctions: Some*

CR account holds that imposing deserved punishment is *itself* an intrinsically good and valuable outcome, regardless of whether any other positive consequences attend it.

The CR approach has been noted in the literature but never explicitly embraced. For example, Michael Moore has alluded to the possibility of consequentialist retributivism, though without subscribing to that approach himself.[57] Dan Markel may be the only self-described retributivist whose elaboration of a retributive commitment adopts a CR perspective. Markel views retribution as "worthwhile in itself,"[58] but also recognizes that punishment institutions cannot be "indifferen[t] to the consequences and costs of punishment."[59] As a result, he aligns himself with the fundamental CR conception of retribution as a goal to be pursued subject to real-world constraints, rather than as a binding moral duty:

> [T]he practice of retribution is only one attractive social practice among many. . . . [R]esources spent on the project of retributive justice are resources not spent on feeding the hungry, housing the homeless, and healing the sick. Thus, to say that retributive justice justifies punishment under ordinary circumstances does not mean

---

Implications for Optimal Enforcement Strategy, 2 AM. L. & ECON. REV. 223 (2000) (incorporating "taste" or preference for retribution into economic analysis of punishment); Paul H. Robinson & John M. Darley, *The Utility of Desert*, 91 NW. U. L. REV. 453 (1997) (defending desert-based punishment as the best way to achieve deterrence).

   Jean Hampton has also suggested one version of such a theory, defending punishment (including retributive punishment) as a means for advancing goals external to the punishment itself: specifically, moral education. *See, e.g.*, Jean Hampton, *Correcting Harms Versus Righting Wrongs: The Goal of Retribution*, 39 UCLA L. REV. 1659, 1659 n.2 (1992) (citing Hampton's own previous work offering a "moral education" theory of punishment, and asserting continuing adherence to the position that "moral education and deterrence should be goals of any well-designed system of punishment"); *id.* at 1701 n.65 ("[T]he theory of retribution I have elaborated understands punishment in a way that may often be consistent with and supportive of the goal of moral education."). Yet Hampton is certainly no pure consequentialist. She adopts a "pluralist approach," *id.*, that also involves a commitment to retribution as an intrinsic good. *See, e.g.*, *id.* at 1659, 1701 (defending retribution as moral imperative); Jean Hampton, *The Retributive Idea*, *in* JEFFRIE G. MURPHY & JEAN HAMPTON, FORGIVENESS AND MERCY 111, 129–33 (1988) (distinguishing "retributive motive" from other, instrumental goals of punishment, and cautioning against mistaking "retributive punishment for moral education").

   57. *See* MOORE, *supra* note 14, at 155–59. *But cf. infra* notes 146–49 (discussing Moore's recent support for a consequentialist perspective as to the general design of criminal justice institutions).

   58. Dan Markel, *Are Shaming Punishments Beautifully Retributive? Retributivism and the Implications for the Alternative Sanctions Debate*, 54 VAND. L. REV. 2157, 2193 (2001); *see also* Markel, *supra* note 22, at 435 (describing retribution as an "intrinsic good"); Markel, *supra* note 11, at 1450 ("[Retribution] is a practice that, generally speaking, can be justified apart from the contingent benefits it might generate.").

   59. Markel, *supra* note 58, at 2193; *see also id.* (asserting retribution is a valid basis for punishment if it "does something good and . . . the good it achieves is not outweighed by some other course of action"); *id.* at 2194 (asserting a version of retributivism in which "states are justified and (*defeasibly*) obligated to respond to legal wrongdoing") (emphasis added).

that punishment ought to be imposed under all circumstances such that the ceaseless pursuit of justice consumes our every and last unit of social resources.[60]

Markel has even applied his understanding of retribution to real-world enforcement and prosecution issues such as imposition of the death penalty[61] and other sanctions[62] and the role of mercy in adjudicating cases.[63] At the same time, Markel does not overtly adopt or defend a CR account covering all aspects of criminal justice, such as law enforcement priorities or certain aspects of prosecutorial discretion. For example, he clearly finds that retribution offers a justification for creating institutions of punishment, but he also hints that at least some aspects of the system's implementation decisions about what or how to punish might best be left to democratic or otherwise political processes.[64]

Clearly the CR view eases the demands on both police and prosecutors. Police may pursue a more flexible strategy that targets whatever mix of offenders combine to generate the largest total amount of desert-based punishment, rather than necessarily striving to apprehend the largest possible raw number of criminals (as under the first view) or having to focus their attention on the most serious subset of offenses (as under the second view). Prosecutors would gain even more freedom, as they would be able to forgo punishment (or accept reduced punishment) for some offenders, if doing so would enhance the total measure of desert-based punishment. This approach might also allow prosecutors to seek departures from desert that impose *excessive* punishment for the sake of achieving more deserved punishment (or avoiding undeserved punishment) in other cases.

---

60. Markel, *supra* note 22, at 436–37; *see also* Markel, *supra* note 58, at 2212–13. Mitchell Berman similarly suggests the need to situate retributive principles within a broadly consequentialist scheme if one is to justify retribution relative to its costs or to other social goods. *See supra* note 37.

61. *See* Markel, *supra* note 22, at 440–77.

62. *See* Markel, *supra* note 58, at 2215–40.

63. *See* Markel, *supra* note 11, at 1453–73.

64. *See* Markel, *supra* note 58, at 2205 ("[L]egislators are the ones interpreting the criminal act and they may properly, i.e., democratically, denominate what counts as meriting more severe penalties and what should not, based on a variety of factors . . . ."); Markel, *supra* note 11, at 1433 ("The intensity of [retributive] punishment and the trigger for the punishment are products of democratic deliberation.").

D. Silence: A Limited Scope for Retributivism

The possibility remains that retributivism simply offers no firm rules about how to implement substantive criminal law rules. Under this view, those charged with enforcing the law would be free to pursue other, non-retributive objectives. In this scenario, law enforcement officials are essentially unguided by retributivism in that no part of their project is driven or necessitated by retributivist theory or goals.

For example, one might adopt a version of deontological retributivism that finds many or all of the enforcement matters this Article addresses to simply fall outside the scope of any deontic obligation. Many descriptions of such obligations limit their scope to instances where a situation or result is the product of the actor's affirmative intent or choice, rather than merely being foreseen or allowed to occur.[65] Under these characterizations of deontic duty, any such duty might not bind institutional actors whose choices are in some way forced by resource limitations or other inevitable constraints.[66] Accordingly, whatever investigatory or prosecutorial tradeoffs are necessitated by real-world exigencies would not be subject to the command (or the guidance) of any deontological imperative.

Such a result might seem disappointing for retributivism, as it would mean that retributive principles have little or nothing to say about many aspects of real-world criminal justice. Yet it also suggests that a punishment system can at least remain consistent with retributivism even if that system is unguided by retributivist principles at the enforcement level. Some thinkers suggest that a criminal justice system can still properly be considered "retributivist" so long as retributive justice remains the underlying *justification* for punishment, even if it does not supply the decision rule for every individual case:

> Some of the 'mixed theorists' [i.e., theorists who claim both
> retributive and utilitarian considerations justify punishment] have

---

65. *See infra* notes 90–95 and accompanying text (discussing similar position in context of charge that *any* criminal justice system violates deontological retributivism by allowing punishment of the innocent).

66. *Cf.* Michael S. Moore, *Four Reflections on Law and Morality*, 48 WM. & MARY L. REV. 1523, 1552 n.79 (2007) [hereinafter Moore, *Reflections*] (suggesting that "choice sets" of institutional designers who must allocate scarce resources, such as legislatures, fall "outside the scope" of deontological obligations, as they lack requisite intent or affirmative action—but indicating that individual "governmental actors like judges, policemen, attorneys, etc." remain subject to deontological obligations); Michael S. Moore, *Patrolling the Borders of Consequentialist Justifications: The Scope of Agent-Relative Restrictions*, 27 LAW & PHIL. 35 (2008) [hereinafter Moore, *Patrolling the Borders*].

wished to distinguish the general justifying aim of the criminal law from the justifications of particular decisions within an ongoing criminal law system. While such a distinction certainly has its uses, it is of no use in wedding retributivist to, say, utilitarian concerns in justifying punishment. Even if one were to concede the existence of 'practice rules'—rules which require moral justification but whose application to particular cases is to be made without resort to that moral justification—such concession goes no distance towards introducing any alien concerns into either a retributive or a utilitarian punishment scheme. That is, if one is a utilitarian in his justification of the practice of punishment, one will remain purely a utilitarian even if one bars utility calculations in applying criminal law rules; and *mutatis mutandis* for the retributivist.[67]

Interestingly, the most notable of the "mixed theorists" Moore is describing advocated essentially the inverse approach to the one suggested here, claiming that the general justificatory purpose of criminal law is utilitarian but that specific cases of punishment should satisfy retributive concerns.[68] This view may have much to recommend it, but it does not provide a version of real-world retribution, as it clearly considers utilitarian concerns paramount in justifying punishment. Further, this view may raise the same issues as a "pure" retributive model in terms of its practical comprehensiveness, as it seems to offer guidance to only two of three branches of government: it sets an agenda for legislators who enact criminal law and suggests how adjudicators should resolve individual cases, but it does not help set the enforcement agenda of executive agents such as police and prosecutors.[69]

Two possible versions of the distinct approach advanced here, where retributive considerations might dictate the substantive rules of criminal law but do not govern enforcement issues, suggest themselves most obviously.[70] The first alternative enforcement agenda would be to pursue

---

67.   *See* MOORE, *supra* note 14, at 154–55 (internal footnotes omitted).

68.   *See* H.L.A. HART, PUNISHMENT AND RESPONSIBILITY: ESSAYS IN THE PHILOSOPHY OF LAW 3**B**13 (1968) (identifying utilitarian considerations as the "general justifying aim" for criminal law, but also finding retributive considerations relevant to the "distribution" of punishment, i.e., decisions about which individuals to punish); John Rawls, *Two Concepts of Rules*, 64 PHIL. REV. 3 (1955) (asserting that utilitarian concerns justify the existence of institutions of punishment, whereas retributive concerns should guide the application of punishment in particular cases).

69.   Rawls, for example, distinguishes between the utilitarian "legislator" and the retributive "judge," *see id*. at 6–7, but does not describe whether a police department or prosecutor's office stands in a position more akin to one of these than the other. *Id.*

70.   Larry Alexander suggests a different allocation between deontological and consequentialist concerns:

utilitarian crime control. Under this option, the criminal justice system would basically have two different components, each with different operational goals. The legislature and other substantive lawmakers, such as sentencing commissions and appellate courts, would generate rules governing criminal liability that specify who *deserves* punishment in a retributive sense. The executive agents charged with applying those rules—police and prosecutors—would pursue an enforcement strategy seeking to minimize crime.

The second possibility is that, instead of adopting a utilitarian agenda, law enforcement authorities should not adopt an *ex ante* agenda at all. Instead of expecting authorities to pursue a single goal specified in advance, this option would merely set out the criminal law's general substantive rules and then trust officials to act as sensible agents in enforcing that law on an ad hoc basis. The officials might favor retributive concerns in some situations and utilitarian ones in others, depending on the facts of the specific case or on broader considerations such as the current crime rate, with enforcement focusing more on utilitarian crime control measures as crime becomes more prevalent.[71] Under this view, the critique that retributivism offers too little real-world guidance loses some force—one could as easily accuse the utilitarian project of providing too

---

In my opinion, the best conception of deontology would deem its core principle to be that one may never use another as a resource without his consent. In other words, a person's body, labor, and talents do not exist for others' benefit except to the extent that he freely chooses to benefit others.

Outside the realm of appropriating others' bodies, labor, and talents, deontology is inapplicable. Thus, in settling conflicts over uses of natural resources and incompatible activities, consequentialist considerations will be just as relevant to the deontologist as they are to the consequentialist. . . .

Therefore, the deontologist can, like the consequentialist, consider consequences to be morally relevant *within the domain beyond the deontological core.*

Larry Alexander, *Deontology at the Threshold*, 37 SAN DIEGO L. REV. 893, 911 (2000).

Alexander's scheme, however, would not eliminate the need to offer some deontological account of how to structure the criminal justice system, for two reasons. First, the practice of punishment itself is a "physical appropriation" under Alexander's scheme, and therefore could only be justified or implemented by reference to deontological principles, not consequentialist ones. Second, even if punishment itself is not an appropriation, many crimes would constitute such appropriations. Accordingly, criminal justice institutions following Alexander's approach, which would both impose and redress "physical appropriations" and therefore be subject to deontological constraints, would confront the same inevitable practical tradeoffs that this article challenges any other retributivist theory to address.

71. It may be interesting to observe that the resurgence of retributivism, both in the scholarly literature and in the policy discourse about issues such as sentencing, coincides with falling crime levels. If the crime rate were to increase significantly, considerations of prevention might well become more prominent as a practical matter.

much guidance in claiming that we should always follow a single goal in implementing criminal law.

Perhaps there exists a casual assumption or expectation that the criminal justice system is designed to work, or should work, in one of these ways: either (a) the law's job is to say what is wrong, but police officers' job is to keep us safe; or (b) the law states clear rules, but officials have (and should have) discretion to choose whether, how, and why to implement the rules in individual cases. Crimes are defined so that we can punish someone only based on his or her past wrongdoing;[72] but within the universe of wrongdoers, we try to punish where doing so will "lock up" a dangerous person, "send a message" to others, or otherwise minimize crime.[73] Or, at least, as the second version of unguided enforcement would have it, police and prosecutors may be swayed mainly by such considerations in some cases but focus more on an offender's blameworthiness, the victim's outrage, or other influences in other cases— without any clear "threshold" dictating when moral blame counts more than crime control or vice versa.

These two options might be consistent with retributivism, and might even have some descriptive accuracy in capturing how the system works, or how people expect or want it to work, but a committed retributivist

---

72. Interestingly, this aspect of criminal law may reveal criminal law's general retributive orientation, for a purely utilitarian set of substantive criminal rules would presumably have no such limitation. A utilitarian criminal code might contain descriptions of conduct to avoid, for the sake of giving notice to the public and thereby achieving behavior control. But, then again, a similar document without binding legal force, made available to the public, could accomplish this goal as well. The main *legal* function of a criminal code is to define when a person is subject to prosecution and punishment by the state, and, under a purely utilitarian system, amenability to prosecution and punishment would not depend on a person's having already engaged in prohibited conduct but only on whether the person is dangerous or his punishment would deter others. In other words, the code's rules for *officials* (who ultimately enforce the code as law) would contain no prohibition against punishing one who had not yet committed a "crime." This feature of a utilitarian criminal justice system drives the most fundamental criticism of utilitarianism—in theory, at least, it imposes no bar against punishing the innocent.

Indeed, it is hard to imagine what the rules of a purely utilitarian criminal code would look like, for its features would differ greatly from any contemporary criminal code. Its adjudicatory rules presumably would not focus on matters of culpability or individual responsibility, except as those issues bear on an "offender's" propensity toward future criminality, or on the efficacy of punishment as a deterrent measure. Pure utilitarianism's implications for substantive criminal law may be as unexplored as retributivism's implications for the enforcement of that law.

The above discussion, however, might apply only to an act-utilitarian punishment scheme—as do the standard, well-rehearsed objections to utilitarianism, such as the complaint that it would not preclude punishment of the innocent. *See* Binder, *supra* note 19, at 322–27; Guyora Binder & Nicholas J. Smith, *Framed: Utilitarianism and Punishment of the Innocent*, 32 RUTGERS L.J. 115 (2000). Perhaps a rule-utilitarian criminal code or system would permit punishment of guilty wrongdoers only.

73. *See* ROBINSON & CAHILL, *supra* note 14, at 117–36 (discussing each of these rationalizations for punishment and how each might conflict with the goal of desert-based punishment).

presumably would not view them as desirable—and neither, perhaps, would anyone else reflecting on the matter.[74] Whatever appeal they have lies precisely in the fact that, and the degree to which, they are able to pursue goals entirely distinct from the retributivist's project. Even if either or both might be sound as a policy matter—a contestable point, to be sure, for what looks like beneficial flexibility to one person may look like baseless inconsistency to another[75]—neither amounts to a version of real-world retribution, for both depend on the concession that retributive considerations offer no clear or consistent guidance about real-world implementation matters. And since both options anticipate an enforcement scheme with either no clear objective or else a forward-looking, strictly utilitarian objective of minimizing future crime, it does not seem worthwhile to compare or contrast them with the three other models, all of which share (though in varying forms, and perhaps to varying degrees) a separate common objective: punishing blameworthy offenders based on their desert. Accordingly, although it seems useful to include these potential enforcement schemes in the menu of options set out in this Part, the balance of this Article does not seek to describe or evaluate them further.

E. The Nuclear Option: Abandon Retributivism Altogether

A final possible reaction to retributivism's lack of a prescriptive agenda is to view it as fatal for retributivism altogether, even as a guiding justificatory principle. In one of the few previous articles focusing directly on the difficulties of putting retributive theory into practice, Russell Christopher draws that very conclusion. Christopher's analysis centers on one problem for retributivism in the real world—what he calls "the prosecutor's dilemma."[76] The dilemma arises most clearly in a stylized example Christopher provides: a prosecutor is approached by one offender who offers to provide testimony against two other, equally serious

---

74. *See* KAPLOW & SHAVELL, *supra* note 7, at 316 n.47 ("[The] mixed view seems difficult to defend. . . . [T]he view appears to be illogical at the most basic level. If one principle—be it welfarism or retributivism—is supposed to be ethically attractive, why is that principle not always our ethical guide? How can our ethical objective change just because we happen to be discussing the application of punishment by judges instead of the determination of laws (including those that set punishments) by legislatures?").

75. *Compare* Bierschbach & Stein, *supra* note 6 (suggesting the possible advantages of having procedural rules mediate between retributive and utilitarian goals), *with* KAPLOW & SHAVELL, *supra* note 7, at 317 n.47 (suggesting that having different institutions follow different goals may lead to "a justice system that is inferior with regard to both systems of evaluation").

76. *See* Christopher, *supra* note 5.

offenders, in return for immunity from punishment; without such testimony, by hypothesis, punishment of the other two offenders is impossible. In such a scenario, the prosecutor must necessarily forgo punishment of one offender to pursue punishment of the other. Because retributivism imposes a moral duty to punish, as well as a duty to treat equal cases alike, Christopher concludes that any resolution of the dilemma is a violation of retributivism. He views this result as an indictment of retributivism as a theory rather than of immunity or plea-bargaining as practices.[77]

As this Article's prior discussion makes clear, Christopher is right to seize on the failure of retributive theory thus far to offer anything more than hints about its own proper application, and he deserves credit as one of the few commentators to draw attention to the theory's practical implications. However, his conclusion—that retributivism's inability to satisfactorily resolve the prosecutor's dilemma reduces the entire theory to utter absurdity—is too strong. It is doubtless true that practical obstacles will predictably create situations where the goal of desert-based punishment cannot be achieved, must be traded off against other goals, or must even be traded off against itself. But this is not uniquely true of retributive punishment; the same could be said of our limited ability to pursue any idealized principle in the real world. That we might sometimes limit freedom for the sake of promoting some other goal (say, safety), or limit it for its own sake (curbing one person's freedom to enhance another's), does not indicate that freedom is a meaningless, empty, or incoherent concept or value. Even diehard retributivists have noted that imposing just punishment is merely one principled objective among numerous objectives we might embrace, and as such is bound to compete with other goals[78] or practical cost considerations.[79] Indeed, such

---

77. *See id.* at 105 ("[T]he incompatibility of retributivism and bargain justice now renders illegitimate not bargain justice, but rather retributivism.").

78. *See* MOORE, *supra* note 14, at 172 ("The retributivist like anyone else can admit that there are other intrinsic goods, such as the goods protected by the rights to life, liberty, and bodily integrity. The retributivist can also admit that sometimes some of these rights will trump the achieving of retributive justice . . . ."); *id.* at 186 ("[H]ere again we need to remind ourselves that retributivists are not monomaniacal about the achieving of retributive justice. Of course those who culpably do some (slight or great) moral wrong deserve some (slight or great) suffering, yet there are other intrinsic goods besides giving culpable wrongdoers their due and sometimes these other goods override the achievement of retributive justice. One such collection of goods we have called the principle of legality . . . ."); *see also supra* note 58 and accompanying text.

79. *See* MOORE, *supra* note 14, at 151 ("It is always relevant to ask, 'at what cost do we attain some form of justice?' . . . Trivial immoralities, and immoralities done in private, by consenting individuals, that are strongly motivated, are not worth the enforcement costs to criminalize and to punish."); *cf.* Markel, *supra* note 22, at 429 ("If we insisted that the state actually achieve complete

constraints might not only affect the system's ability to enforce the law, but might also influence substantive decisions about what moral wrongs to criminalize in the first place.[80]

Without ignoring or minimizing retributivism's limitations, abandoning it altogether seems a hasty response. For reasons noted earlier,[81] this Article assumes that retributivism is a prima facie legitimate perspective and proposes to investigate whether it can address the purported plea-bargaining dilemma and other enforcement issues, or whether its project necessarily skirts these issues (and if so, what that means). As discussed earlier in this Part, retributive theory itself suggests at least three possible means of implementing its principles. The balance of this Article seeks to further elaborate what each of those methods would entail as a practical matter and to evaluate the methods based on both their internal logic and their external ramifications.

II. APPLYING THE OPTIONS: MAKING RETRIBUTION REAL

This Part applies each of the three possible versions of real-world retribution sketched above to specific issues, discussing how each might guide actual enforcement decisions. Section A addresses basic issues confronting police departments, while section B deals with prosecutorial practices.[82]

---

enforcement and punishment, we would then be placed in the untenable position of spending all of our collective resources on criminal justice.").

80. *See* MOORE, *supra* note 14, at 68 (noting different kinds of possible "limits to the use of criminal law that should lead us to eschew prohibiting certain conduct even if it is morally wrongful"); *id.* at 661–65, 739–95 (describing and defending various considerations, such as personal liberty, that limit the proper scope of retributivist criminal prohibitions).

81. *See supra* notes 23–26 and accompanying text.

82. The discussion in this Part will not focus on issues related to criminal sentencing, which might also seem to merit consideration as a significant aspect of the application of criminal rules in individual cases. The reason for sentencing's absence from this discussion is that *all* of the retributive models would likely address sentencing questions in a roughly similar way, in keeping with some form of the "limiting retributivism" approach various commentators have described. *See, e.g.*, Richard S. Frase, *Excessive Prison Sentences, Punishment Goals, and the Eighth Amendment: "Proportionality" Relative to What?*, 89 MINN. L. REV. 571, 590–92, 623–25 (2005); Richard S. Frase, *Limiting Retributivism*, *in* THE FUTURE OF IMPRISONMENT 83 (Michael Tonry ed., 2004); Youngjae Lee, *The Constitutional Right Against Excessive Punishment*, 91 VA. L. REV. 677, 704 (2005); Norval Morris, *Desert as a Limiting Principle*, *in* PRINCIPLED SENTENCING 201 (Andrew von Hirsch & Andrew Ashworth eds., 1992); *cf.* ROBINSON & CAHILL, *supra* note 14, at 212–17 (proposing scheme where desert dictates the amount of punishment but utilitarian concerns influence the method of punishment).

The premise of the limiting-retributivism approach to sentencing is that retributive concerns (no matter how seriously one is committed to them) cannot and do not dictate an exact punishment for any given offender but only prescribe a *relative* assessment of the seriousness of the offense, which can then generate a range of acceptable sentencing options. Once retributive concerns have set this range, other considerations—such as incapacitation, rehabilitation, or other utilitarian goals—may influence

A. Police Priorities

A particular example might facilitate comparison of the differences among the three schemes for real-world retribution in terms of the priorities they would set for police departments. I offer the (admittedly stylized and simplified) scenario of a society confronted with the following universe of crimes and criminals each year:

Offense	No. of Offenders	Cost to Apprehend	Error Rate	Punishment	Cost Per Unit of *Deserved* Punishment
Shoplifting	440 (220 + 220)	$100 / $450	0% / 20%	1 month	$100 / $750
Car Theft	280	$450	20%	30 months	$25
Aggravated Battery	200	$225	5%	50 months	$5
Murder	4	$1,080	10%	225 months	$6

The substantive content of the criminal prohibitions is not especially important; other offenses could substitute for those listed with no effect on the scenario. For purposes of simplicity, the hypothetical generally

---

the selection of a specific sentence within the range. Indeed, so long as the punishment is within the prescribed range, nonretributive concerns could guide the *method* of punishment as well as the specific *amount* of punishment without offending retributive principles. *See* ROBINSON & CAHILL, *supra* note 14, at 212‒17. Of course, to the extent that the initial retributive distinctions are highly nuanced, as some research indicates, the available punishment range will narrow, reducing the influence of nonretributive considerations. *See* Paul H. Robinson, *The A.L.I.'s Proposed Distributive Principle of "Limiting Retributivism": Does It Mean in Practice Anything Other than Pure Desert?*, 7 BUFF. CRIM. L. REV. 3, 10‒12 (2003).

Because any sentence within the range prescribed by a limiting-retributivism scheme would comport with the offender's desert, application of such a scheme in any individual case would satisfy any of the models this Article sets out, including the absolutist model. On the other hand, sentencing rules authorizing any amount of undeserved punishment, or less punishment than deserved, would violate the retributivist principle, so all three models would strongly disfavor or entirely prohibit such rules.

assumes that the "cost to apprehend" (meaning the total costs of finding and prosecuting each offender—note that the low dollar figures posited here for convenience's sake could be multiplied across the board by ten, or one hundred, with no impact on the analysis), the "error rate" (that is, the frequency with which an arrestee is not the actual offender),[83] and the corresponding punishment are constant for all commissions of each crime. The hypothetical also introduces one exception to this general uniformity, for the sake of hinting at how more nuanced schemes might play out. For shoplifting, the hypothetical assumes half the shoplifters are caught in the act, thus having low apprehension costs and an effective error rate of zero, whereas the other half get away, becoming just as hard to catch as the car thieves. Increasing the variability of the factors might make the hypothetical more realistic, but would not alter or undermine anything significant in the discussion below.[84]

---

83. This hypothetical assumes that the police department could even ascertain or estimate what its own error rate would be for various crimes, which might be an unwarranted speculation. An actual police department might lack this information and would have to rely on other considerations. Indeed, it appears that little is currently known about error rates. Darryl Brown has discussed some of the relevant, and conflicting, factors that might influence the rate of error across different crimes:

> It is unclear as a practical matter whether one of these categories—wrongful small punishments or wrongful large ones—is realistically more likely. We have certainly learned in recent years that wrongful convictions happen even in cases involving the most serious crime—capital murder. But it may be that the factual nature of serious cases makes them more prone to erroneous outcomes than simpler, low-level cases. On the other hand, more resources are probably committed to investigating and prosecuting serious cases, which might mean a higher error rate in the investigation and charging of low-level crimes. Similarly, political pressure for conviction should be greater in serious cases, but public (and defense-side) scrutiny of the prosecution is probably greater as well.

Darryl K. Brown, *Rationing Criminal Defense Entitlements: An Argument from Institutional Design*, 104 COLUM. L. REV. 801, 819 n.69 (2004).

Defined in terms of how many arrestees are either not prosecuted or not convicted (the relevant factors in a discussion of police arrest decisions), the rates may be rather high. *See, e.g.*, McGregor Smyth, *Holistic Is Not a Bad Word: A Criminal Defense Attorney's Guide to Using Invisible Punishments as an Advocacy Strategy*, 36 U. TOL. L. REV. 479, 481 n.15 (2005) ("In 2002, only 62.1% of all arrests [in the state of New York] resulted in a conviction for any offense. . . . In New York City, only 57.4% of arrests resulted in convictions." (citing statistics available in New York State Division of Criminal Justice Services, Dispositions of Adult Arrests 1999–2006, http://criminaljustice.state.ny.us/crimnet/ojsa/dispos/index.htm)). Defined in terms of how many *prosecuted* defendants are actually innocent, the limited evidence suggests a much lower error rate. *See, e.g.*, Darryl K. Brown, *The Decline of Defense Counsel and the Rise of Accuracy in Criminal Adjudication*, 93 CAL. L. REV. 1585, 1595 n.34 (2005) ("One rough and conservative estimate based mostly on surveys of judges and attorneys posits an accuracy rate of 99.5%." (citing C. RONALD HUFF ET AL., CONVICTED *BUT* INNOCENT: WRONGFUL CONVICTION AND PUBLIC POLICY 53–62 (1996))).

In any case, rate of error, defined either of these ways, would surely be highly relevant for any retribution-oriented police department or prosecutors' office.

84. Further, some of the figures might experience increasing or decreasing marginal rates based on the number of offenders prosecuted in each category, along the lines of the breakdown I have presented for shoplifting but in more complex fashion. A conscious decision by the police and

Across crimes, however, the punishment levels, apprehension costs, error rates, and number of offenders vary, as would clearly be true in any actual situation. The punishment levels listed roughly track the averages for the offenses in question.[85] The figures chosen for the number of offenders, the apprehension costs, and the error rate within each category are essentially, though not entirely, arbitrary,[86] chosen to make the distinctions between the different models more transparent. It will be obvious from the following discussion how manipulating those figures would affect the relative priorities within each retributive law enforcement scheme.

The final column uses the apprehension cost, error rate, and punishment columns to calculate the cost to the system for each unit of

prosecutors themselves to pursue fewer cases might also reduce the error rate. *See* James S. Liebman & Lawrence C. Marshall, *Less Is Better: Justice Stevens and the Narrowed Death Penalty*, 74 FORDHAM L. REV. 1607, 1661 (2006) ("[E]vidence reveals that reducing the numerousness of capital prosecutions and verdicts tends to decrease the rate of error as well.").

These variations would complicate the application but would not alter the enforcement agenda. They would simply require enforcers to set priorities within offense categories that would replicate some features of their priorities across offense categories.

85. *See* U.S. DEP'T OF JUSTICE, BUREAU OF JUSTICE STATISTICS, FELONY SENTENCES IN STATE COURTS, 2002, at 5 tbl.4 (2004), *available at* http://www.ojp.usdoj.gov/bjs/pub/pdf/fssc02.pdf [hereinafter FELONY SENTENCES 2002]. The report shows the following average prison sentences imposed for felonies in state courts in 2002: murder, 225 months; aggravated assault, 54 months; motor vehicle theft, 30 months. *Id.* The estimated amount of actual time to be *served* for those sentences, however, is a mere 49–66% of the sentence imposed. *Id.*

86. As to number of offenders, available FBI data providing nationwide totals for various crimes from 1987 to 2006 reveal an overall ratio for that period of approximately 70 motor vehicle thefts for every 50 aggravated assaults and for every single "murder and nonnegligent manslaughter." *See* FEDERAL BUREAU OF INVESTIGATION, CRIME IN THE UNITED STATES 2006, at tbl.1, *available at* http://www.fbi.gov/ucr/cius2006/data/table_01.html [hereinafter CRIME IN THE U.S. 2006]. The annual rates within that time period vary somewhat but not greatly: each year shows a ratio of 64–77 motor vehicle thefts for each murder, and 43–59 aggravated assaults for each murder. *See id.* The table above maintains the 70:50:1 ratio for those three crimes. Another FBI table, from the same study, showing property crime totals for the last five years, shows a ratio of about 55 cases of shoplifting for every 70 motor vehicle thefts. *See* CRIME IN THE U.S. 2006, *supra*, at tbl.7, *available at* http://www.fbi.gov/ucr/cius2006/data/table_07.html. The table reflects this by maintaining a ratio of 70:55 between the car thefts and the "easy" cases of shoplifting, on the assumption that many other cases of shoplifting (the hard ones to catch) are not reported or otherwise, as the FBI reports put it, "known to law enforcement." Indeed, many may not even be detected.

As to apprehension costs, I am aware of no available hard data that might give any indication of how much, on average, it costs any American criminal justice system to find and prosecute an individual offender in any given offense category. The costs provided in the table were selected partly to generate reasonably simple numbers in the cost-per-unit column, and partly to reflect reasonable assumptions about the difficulty of enforcing different crimes. It would seem, for example, that murder, where the victim obviously cannot report the crime and the offender has a stronger incentive to avoid detection, would have much higher enforcement costs than the other crimes. For aggravated battery, on the other hand, the victim might know the assailant, making costs somewhat lower (though not all victims may come forward to prosecute).

As to error rates, see *supra* note 83.

deserved punishment—which, of course, is what a retributive system seeks. Taking the example of aggravated battery, every 20 arrests will cost a total of $5,400. One of those arrests (5% of 20), however, will be an innocent person, so the total amount of deserved punishment will only be that for 19 offenders (19 times 50, or 950). Yet, if we assume that the erroneously arrested person is actually prosecuted and punished, the amount must be further reduced to account for the *undeserved* punishment inflicted on that person. Making the admittedly contestable assumption that one unit of undeserved punishment is equivalent to (the negation of) one unit of deserved punishment,[87] it is necessary to subtract an additional 50 units from the punishment total, reducing it to 900. Dividing the costs of $5,400 by the total deserved punishment of 1,800 yields a per-unit cost of $5.

Assume that this society's police department has a fixed annual budget. The amount of the budget is unimportant as long as it is less than $296,320, that being the amount it would cost the department to apprehend all offenders.[88] In other words, assume the budget is small enough to require the department to make choices about which crimes to prioritize for enforcement purposes and which offenders must remain at large.[89] The

---

87. It is unclear whether retributivism is neutral as between "false positives" (undeserving people who are punished) and "false negatives" (deserving people who fail to be punished), or whether retributivism should be seen as valuing the avoidance of punishing the innocent more highly than the imposition of punishment on the guilty. *See, e.g.*, KAPLOW & SHAVELL, *supra* note 7, at 342 n.103 (expressing uncertainty as to retributivism's view of how to balance positive value of deserved punishment against negative value of punishing the innocent); Bierschbach & Stein, *supra* note 6, at 1208 (defending claim that retributivism is "agnostic" between these two kinds of error and "expresses no strong preferences for whether [evidentiary] rules skew more toward the side of false positives or more toward false negatives"); Russell L. Christopher, *Deterring Retributivism: The Injustice of "Just" Punishment*, 96 NW. U. L. REV. 843, 912–13 (2002) (discussing the debate within retributivism as to whether punishing the innocent is worse than failing to punish the guilty).

If this assumption were modified to ignore the "retribution losses" from punishing the innocent (e.g., by assuming wrongly arrested persons were later released without prosecution), or to take them more seriously (e.g., by treating a unit of undeserved punishment as offsetting two or three units, rather than one unit, of deserved punishment), the absolute per-unit costs of deserved punishment for the various offenses would change and the ratios between their costs would change (though not radically), but their relative ranking in terms of per-unit cost would not. Thus, the analysis of priorities under CR—the retributive scheme most affected by per-unit costs—might change somewhat if such an alteration were made, but the basic elaboration of the CR approach should make clear how it would work under any of these assumptions.

88. More accurately, this is the amount that it would cost for the department to make arrests for all *offenses*. With a nonzero error rate, some offenders would remain at large—but the police department, not knowing when it had arrested the wrong person, would not know to continue pursuing any of those offenders.

89. One might question whether a police department can exercise the kind of control over its enforcement practices that this discussion might imply. Often the police are merely reactive, responding to complaints of crime and trying to find the offenders, rather than making conscious

question here is how each of the versions of real-world retribution
discussed in Part I would guide those choices.

Notably, a purely utilitarian criminal justice system would, at least in
theory, have a clear agenda for setting police priorities: it would promote
the mix of enforcement that prevents the most crime (or, more accurately,
avoids the most harm) relative to cost. Such an agenda would require
consideration of various factors, including many beyond those appearing
in the chart above—such as the effect of punishing a given crime one year
on the amount of that crime the following year, the relative "harm" each
crime is thought to impose on its victim, or the costs of punishment in
addition to the costs of apprehension—and the balancing of those factors
would surely prove complex. Nonetheless, the enforcement *agenda* would
be clear.

What agenda would a retributive system set for itself? Each of the three
versions of real-world retribution described in Part I will generate a

---

decisions in advance about which crimes they will pursue. Certainly that is true, and it gives rise to its
own question in terms of the practical application of retributivism. At least some versions of
retributivism suggest a duty to punish known offenders. *See supra* Part I.A. On this view, would there
also be some (perhaps weaker) obligation to investigate known *offenses*—as opposed to, say, devoting
resources to a sting or undercover operation that may or may not uncover any wrongdoing? Though
that view may be possible, I am not aware of anything in the existing literature that states or strongly
implies such a view. To the contrary, the principled, nonconsequentialist approach of retributivism
should deny any place for the kinds of concerns that would lead the police to prioritize reported crimes
over unreported ones—for example, satisfying the victim(s) or maintaining the reputation of the police
department itself, both of which are beneficial consequences unrelated to the offender's desert. *Cf.*
Michael Moore, *Victims and Retribution: A Reply to Professor Fletcher*, 3 BUFF. CRIM. L. REV. 65, 67
(1999) ("I think victims should and must be ignored if you are claiming to be doing retributive
theory."). *But cf.* Fletcher, *supra* note 40, at 55 (arguing that Moore's retributivist "view of deserved
punishment would be richer if it included the suffering of victims in his account of why punishment is
deserved and therefore just").

As to the crux of the criticism here—that crimes dictate the police's agenda, rather than the
police's agenda dictating what crimes are investigated—though it is surely accurate for much police
work, I do not think it affects the discussion here, for two reasons. First, to the extent police merely
react to external forces, police work would look exactly the same under a purely utilitarian model, a
purely retributivist model, or any other model; to the extent such forces do drive enforcement
practices, there is nothing much for this Article to say about that aspect of enforcement other than to
note its existence.

Second, that aspect surely does not describe the *whole* of police work. Many aspects of police
work involve something other than an immediate effort to respond to a reported crime: for example,
most undercover work is not of this sort, nor would police reopen "cold cases" if police lacked the
ability to set their own priorities. Further, even for many reported crimes, the level of police
commitment to finding the offender surely varies (and should vary, under either a utilitarian model or
any of the real-world retribution models discussed here). Police are highly likely to devote less
attention to tracking down a nonviolent property offender than a murderer—and note that such a
decision cannot *always* be justified solely on consequentialist grounds, as the facts might indicate that
the car thief is likely to reoffend while the killer is not.

different allocation of police resources. This section considers them in turn.

1. The Absolutist Model

The absolutist model views desert-based punishment as a moral duty. Accordingly, the system has an obligation to punish the deserving and refrain from punishing the undeserving. As noted above, the usual elaboration of this view does not vary the magnitude of the duty based on the magnitude of the relevant punishment: the duty to punish a minor wrongdoer is as strong as the duty to punish a more serious one, though the amount of punishment each would ultimately receive would be different. Accordingly, the law enforcement agenda under this view would focus specifically on two of the factors in the chart above: the error rate (which dictates how many people receive undeserved punishment) and the cost of apprehension (which dictates how many people receive deserved punishment).

As to the error rate, one initial issue to surmount is whether an absolutist deontological system of retributive punishment would tolerate *any* error rate greater than zero. The argument here is that because an absolutist conception holds that punishing the innocent is always categorically forbidden, and because any real-world system of punishment (even a purely retributive one) will inevitably have errors resulting in innocents being punished, the absolutist is bound by her own principles to refrain from creating a system of punishment.[90] Accordingly, the critique proceeds, (absolutist) retributivists are in a double bind: in satisfying the moral duty to punish the guilty, they are bound to violate the moral duty to avoid punishing the innocent.[91]

---

90. *See, e.g.*, KAPLOW & SHAVELL, *supra* note 7, at 340–41 (claiming that retributivists insisting on the categorical impermissibility of punishing the innocent must reject any real-world punishment system); *id.* at 340 n.99 (citing commentators who have offered similar arguments); *cf.* Christopher, *supra* note 87, at 869–88 (arguing that any retributivist system will inevitably punish some innocents, and that while this may not be fatal to retributivism per se, it takes away retributivism's asserted moral superiority to consequentialism); Daniel J.H. Greenwood, *Restorative Justice and the Jewish Question*, 2003 UTAH L. REV. 553 (noting that the Jewish tradition recognized the moral correctness of punishing wrongdoing but denied fallible humans the right to exact such deserved punishment).

91. Guyora Binder and Nicholas Smith describe the tension between the (absolutist) retributive duties:

> Bentham is trying to be clear about two points that retributivists typically obscure: that punishment of the innocent is an unavoidable concomitant of punishing the guilty, but that punishment of the guilty is an avoidable policy choice. Thus, when retributivists claim to adhere rigidly to the principle of no undeserved punishment, they deceive themselves. If they really regarded avoiding punishment of the innocent as an absolute duty they could in fact

Some retributivists have responded to this critique,[92] though their responses sometimes seem to soften the absolutist position and adopt some other version of retributivism.[93] Further, the basis of the standard retributivist defense to this critique—that if the system did not knowingly or intentionally punish the innocent, it would violate no moral duty and would remain acceptable[94]—has been met with some skepticism.[95] Assuming for present purposes that the inevitability of a nonzero error rate would not prove fatal to an absolutist retributivist criminal justice system, the error rate would remain highly relevant under such a system, and the system's absolutist commitments might compel it to dedicate vast resources to minimizing the rate of error.

Aside from error rate, the absolutist model's enforcement scheme would probably concentrate its focus on the per-offender cost of apprehension. The severity of the crime in question should not enter the calculus, for in the absolutist account, the duty to punish wrongdoing is categorical and therefore applies to all offenses, great and small. Facing constraints on its ability to punish *all* deserving offenders, the model's second-best alternative would be to satisfy the duty to punish in as many cases as possible—that is, to maximize the number of deserving offenders receiving punishment.

---

adhere to it by avoiding punishment. They do not appreciate that they are responsible for choosing to punish the innocent because they do not see punishment of the guilty as discretionary.

Retributivists in fact treat punishment of the guilty as an absolute duty which compels sacrifice of the innocent.

Binder & Smith, *supra* note 72, at 143.

92. *See* MOORE, *supra* note 14, at 158 ("The probable punishment of the innocent by any real-world punishment scheme is not much of a worry even for deontological versions of retributivism."); Larry Alexander, *Retributivism and the Inadvertent Punishment of the Innocent*, 2 LAW & PHIL. 233 (1983) (concluding that viable versions of retributivism do not categorically forbid risking punishment of innocents).

93. *See, e.g.*, MOORE, *supra* note 14, at 156–57 (noting, before defending the deontological view against critique based on the inevitable punishment of innocents, that such a critique is a "non-problem" for a *consequentialist* retributivist).

94. *See, e.g.*, *id*. at 158 ("Agent-relative moral norms bind us absolutely only with respect to evils we either intend or (on some versions) knowingly visit on specified individuals. One can thus arrange . . . punishment in ways that predictably hurt some who do not deserve to be hurt, without for a moment ceasing to be an agent-relative theorist about morality."); Alexander, *supra* note 92, at 245 (citing RONALD DWORKIN, *Principle, Policy, Procedure*, *in* CRIME, PROOF & PUNISHMENT, 195, 201–10 (1981)).

95. *See, e.g.*, KAPLOW & SHAVELL, *supra* note 7, at 344 n.106 (describing the distinction between intended and inadvertent punishment of the innocent as "substantially inaccurate in the present setting" as well as conceptually unconvincing); *cf.* Christopher, *supra* note 87, at 869–70, 887–88 (asserting that a retributivist system would allow for *intentional* punishment of the innocent).

For the hypothetical scenario sketched here, then, the absolutist model would advise first pursuing all 220 of the easily-caught shoplifters, for whom both apprehension costs and error rates are lowest. After that, the model would support pursuing the batterers, for whom both apprehension costs and error rate rank second. This allocation of funding would both maximize the number of deserving offenders receiving punishment and minimize the number of undeserving innocents receiving punishment.

2. *The Threshold Model*

The threshold model differs from the absolutist model in its willingness to view the duty to punish as less than categorical—or at least, even if categorical in some sense,[96] not uniform in its binding effect across all cases. While the absolutist model sees the duty to punish as unequivocal, the threshold model allows for two considerations to affect the duty's power to oblige punishment in an individual case: first, the magnitude of the duty itself might vary based on the moral gravity of the offense; second, the magnitude of other relevant factors (such as the good or bad consequences of punishing) might sometimes, though rarely, grow large enough to trump retributive justice.[97]

Because, under the threshold view, the force of the duty to punish an offense rises with the offense's seriousness, this model would not seek to punish the most offenders (as the absolutist model does), but rather the most *serious* offenders. Its first priority would be to apprehend perpetrators of the most serious offense, as to whom the duty to punish is most acute; once all of those offenders had been punished, the model would focus on the second-most serious crime, and so on. Pursuing this agenda would require only an *ordinal* ranking of the seriousness of offenses and not any measure of the absolute magnitude of their seriousness or the "gap" between one and the next. A more serious crime, whether slightly or enormously graver than another crime, calls out more forcefully for punishment, and therefore would merit enforcement priority.

The threshold model would also allow this set of priorities to give way to other concerns, such as cost, if these concerns became overwhelming.[98] Thus, in the more plausible scenario where not all offenses have the same

---

96. *See supra* note 42 (quoting Michael Moore, a threshold deontologist, describing duty to punish as "categorical" but nonetheless allowing that it may sometimes give way to other considerations).

97. *See supra* Part I.B (discussing threshold retributivism).

98. *See generally supra* Part I.B.

enforcement cost, this model would acknowledge that, at some point (the threshold point itself), pursuit of even the most serious offender would not be appropriate if tracking down that offender would necessitate enormous costs—either direct financial costs, or other costs such as the reduction in human liberty from intrusive searches for evidence, or the opportunity cost in terms of a lost chance to punish many other deserving offenders.

In this scenario, then, a police department employing the threshold model would first pursue all the murderers, then pursue as many of the batterers as the budget would allow.

3. Consequentialist Retributivism

The consequentialist-retributivist, or CR, model differs from both other models in that it views imposing deserved punishment not as a duty but as a goal. Each of the other two models understands retributive justice as an obligation driving a particular enforcement agenda: in one case, pursuing the most offenders, and in the other, pursuing the gravest crimes. The CR model is outcome-oriented, seeking to maximize the total amount of deserved punishment the system imposes; that total is driven by a combination of the number of offenders caught, the amount of (deserved) punishment per offender, and the cost of apprehending each offender. Accordingly, a CR system would take into account both resource constraints and offense seriousness, balancing these two concerns by focusing on the per-unit cost of deserved punishment.[99]

A model relying on cross-offense comparisons of the per-unit cost of punishment, however, raises a significant practical concern. To engage in such comparisons, such a model must assume the ability to rank crimes along a *cardinal* scale, not merely an ordinal one. In other words, every unit of punishment must have fixed and constant value, or else a known and identified rate of decreasing (or increasing) marginal value: a 300-unit crime must equal exactly three 100-unit crimes, or at least *some* identified quantity of 100-unit crimes.

As typically understood, however, retributive punishment schemes can provide only ordinal, not cardinal, rankings of the seriousness of different crimes.[100] The sentencing or other punishment ranges actually assigned by

---

99. *See generally supra* Part I.C.
100. *See, e.g.*, Darryl K. Brown, *Cost-Benefit Analysis in Criminal Law*, 92 CAL. L. REV. 323, 365 (2004) ("There is little agreement among retributivists (or others) on how to match wrongdoing to specific levels or forms of punishment rather than a relative ranking—that is, a cardinal rather than merely an ordinal ranking."). *But cf.* ANDREW VON HIRSCH, CENSURE AND SANCTIONS 36B46 (1993) (discussing possible ways to construct a cardinal ranking of offenses).

law are at best imperfect, and possibly misleading, proxies for a genuine moral assessment of the relative severity of different crimes.[101] Accordingly, it may not be proper to assume that a single "20-year crime" is equal in moral weight to four "5-year crimes."

This limitation of retributive theory creates a serious difficulty for the potential to implement CR, which demands a means for determining whether or not a given array of punishments achieves more desert than another. It is important to note, however, that this exact difficulty confronts any other consequentialist model as well. Where CR must determine the relative moral wrongfulness of different offenses, utilitarianism must determine the relative harmfulness of different offenses in terms of setbacks to human welfare. Without such judgments, *neither* theory could manage to set priorities between different offenses. In short, the problem of devising a metric for balancing the costs of punishment against its benefits is profound, but also unavoidable for any comprehensive scheme that allows for such balancing rather than following some binding *ex ante* imperative.

Further, as a practical matter, the legally prescribed punishment levels do specify the extent of the system's real-world *opportunity* to impose punishment, and the role of the enforcement agents within the system is such that they are expected to pursue that opportunity to the extent possible. Thus it might be plausible as a second-best approximation to use actual punishments as a measure of the seriousness of crimes, or at least to conclude that a set of penalties generating more total punishment achieves the desert goal better than one that generates less.

A crucial (and highly contestable) assumption here is that no prescribed punishment definitely exceeds the amount dictated by the offender's moral desert. So long as all authorized punishment is less than or equal to what the offense merits, more punishment means more deserved punishment—assuming the punishment is imposed on the right person, which is a distinct issue. The assumption, in other words, is that the legislature has itself sought to obey the retributive principle in defining crimes and setting their penalties.[102] This may be a dubious assumption in the real world, but it seems fair to allow a retributive *model* of criminal justice to stipulate

---

101. *See, e.g.*, Paul H. Robinson, *Dissent from the United States Sentencing Commission's Proposed Guidelines*, 77 J. CRIM. L. & CRIMINOLOGY 1112, 1113–16 (1986) (criticizing federal sentencing guidelines for failing to consistently distinguish crimes based on relative seriousness, or any other principled considerations, and providing examples).

102. The possibility remains that a retributivist legislature would deliberately enact an "imperfect" criminal code (in a retributivist sense), precisely for the sake of *advancing* retributive goals in the real world. This possibility is addressed *infra* Part II.C.

that the law itself should be retributive in its design. While such a model must account for inevitable resource constraints, it need not design policy prescriptions around flaws or shortcomings in the substantive law, for the obvious policy prescription in such a situation is simply to amend the substantive law.

For the hypothetical scenario set out above, the CR model, setting its priorities according to the per-unit cost of deserved punishment, would focus first on the aggravated battery offenses (and second on the murders), which provide the highest bang for the law enforcement buck and therefore enable the highest overall imposition of deserved punishment.

B. Prosecutorial Discretion

Prosecutors often use their discretion to decline to punish clearly deserving offenders, or to give such offenders less punishment than is legislatively authorized. Indeed, this is the norm rather than the exception. Well over ninety percent of cases are resolved with guilty pleas,[103] almost all of which involve plea bargains, trading off a lesser amount of punishment in return for a certain conviction.[104] The federal sentencing guidelines make the "plea discount" explicit by reducing an offender's guideline sentence if he "clearly demonstrates acceptance of responsibility,"[105] which generally requires a plea of guilty.[106]

---

103. *See, e.g.*, FELONY SENTENCES 2002, *supra* note 85, at 8 tbl.9 (showing that in 2002, ninety-five percent of all state felony convictions were obtained by guilty plea); U.S. SENT'G COMM'N, 2006 SOURCEBOOK OF FEDERAL SENTENCING STATISTICS, at fig.C, *available at* http://www.ussc.gov/ANNRPT/2006/ SBTOC06.htm (figures for 2002 to 2006 showing pleas were involved in 94.5% to 97.1% of cases annually during that period).

104. *See* G. NICHOLAS HERMAN, PLEA BARGAINING 1 (1997) (noting that an overwhelming number of the approximately ninety percent of U.S. criminal convictions resulting from guilty pleas are achieved through plea bargains); *see also* Daniel C. Richman, *Bargaining About Future Jeopardy*, 49 VAND. L. REV. 1181, 1237 (1996) ("[P]lea bargaining is the dominant mode of adjudication . . . ."). *See generally* GEORGE FISHER, PLEA BARGAINING'S TRIUMPH: A HISTORY OF PLEA BARGAINING IN AMERICA (2003).

105. *See* U.S. SENT'G COMM'N, 2006 FEDERAL SENTENCING GUIDELINES MANUAL, ch.7, § 3E1.1, *available at* http://www.ussc.gov/2006guid/3e1_1.html.

106. *See* Robert N. Strassfeld, *Robert McNamara and the Art and Law of Confession: "A Simple Desultory Philippic (Or How I Was Robert McNamara'd into Submission),"* 47 DUKE L.J. 491, 513–14 (1997) ("The notes indicate certain circumstances that will preclude or strongly militate against an adjustment. The Guidelines presuppose that in most instances the defendant who seeks an adjustment will have pled guilty, and note 2 states that a defendant who denies factual elements of guilt, and only admits guilt and expresses remorse after conviction, is ineligible for an acceptance of responsibility adjustment."); *cf. id.* at 514 n.115 (noting that the Guideline commentary recognizes "that there may be rare instances in which the defendant can manifest acceptance of responsibility despite a conviction at trial where the defendant had gone to trial to assert issues unrelated to his factual guilt, such as the constitutionality of the statute in question or the applicability of a statute to his conduct").

Prosecutors also commonly use the testimony of some offenders as a means to facilitate prosecution of others, granting the testifying criminal reduced punishment (by way of sentencing departures for "substantial assistance"[107] and the like) or no punishment (by way of a "cooperation agreement"[108] or a grant of witness immunity[109]) as an inducement.

Do these common practices violate the retributive principle? The balance of this section evaluates their validity under each of the three models.

1. The Absolutist Model

Under the absolutist model, law enforcement at the prosecutorial stage would focus on making certain that all, and only, identified wrongdoers were punished in an amount appropriate to their desert. Accordingly, the absolutist model would categorically ban plea bargaining, witness immunity, downward departures for substantial assistance, or any other failure to impose deserved punishment.[110]

---

107. *See* U.S. SENT'G COMM'N, *supra* note 105, § 5K1.1 (allowing downward sentencing departure for a defendant who has "provided substantial assistance in the investigation or prosecution of another person who has committed an offense").

Between 2003 and 2006, the rate of "substantial assistance" departures varied from 14.4% to 15.9% of *all* federal criminal convictions. *See* U.S. SENT'G COMM'N, *supra* note 103, at fig.G.

108. Cooperation agreements involve a prosecutor's informal promise not to prosecute a potential defendant, as opposed to a formal grant of immunity. *See* John G. Douglass, *Confronting the Reluctant Accomplice*, 101 COLUM. L. REV. 1797, 1811 (2001). Prosecutors prefer such an arrangement for several practical reasons: "The prosecutor purchases not only the testimony, but also the opportunity to prepare the accomplice to testify. And an accomplice who undergoes hours of pretrial preparation is more predictable, and hence more valuable, than one who is forced to testify under a simple grant of immunity." *Id.* at 1826–27. In addition, unlike statutory immunity, a promise not to prosecute may be conditioned on the witness providing full cooperation and truthful testimony.

109. Even if one considers only the federal jurisdiction, prosecutors make hundreds of immunity requests each year (in the federal system, prosecutors must receive authorization from the Attorney General before granting immunity). Federal immunity requests reached their peak in 1986, with prosecutors making 2,550 requests (involving 5,013 witnesses); since then, the number of requests has decreased, with 743 requests (involving 1,280 witnesses) occurring in 2006, the most recent year for which information is available. *See* U.S. DEP'T OF JUSTICE, BUREAU OF JUSTICE STATISTICS, SOURCEBOOK OF CRIMINAL JUSTICE STATISTICS 2006, at tbl.5.1, *available at* http://www.albany.edu/sourcebook/pdf/+512006.pdf.

The main reason for the decline in formal immunity grants is not that the practice of trading testimony for nonprosecution has waned but rather that cooperation agreements and other informal promises have replaced formal immunity grants. Further, the figures above relate only to the federal system; immunity grants in one form or another are provided in each of the fifty states as well. *See* 8 JOHN HENRY WIGMORE, EVIDENCE IN TRIALS AT COMMON LAW § 2281 n.11 (John T. McNaughten ed., rev. ed. 1961 & Supp. 1998) (listing state and federal immunity statutes).

110. If the legislature sometimes enacted laws that departed from the dictates of desert—which even a purely retributivist legislature might see reason to do, *see infra* Part II.C—a prosecutor might also confront cases in which a person deserved some punishment, but the punishment that would be

2. *The Threshold Model*

As under the absolutist model, the threshold model adopts a view of punishment as a moral duty according to which it would be incumbent on prosecutors to pursue the agent-relative duty to punish each identified offender to the full extent of his or her desert. Thus, the threshold model would also generally ban plea bargaining, witness immunity, and the like.

The only difference here is that the threshold view would not *categorically* ban these prosecutorial practices, but might allow them in cases where some especially bad outcome would result from not using them—as Moore has put it, they would be unavailable except when "needed to punish some very deserving criminals."[111] Moore does not further specify, however, just how "deserving" those criminals must be before the practices would be allowed.

3. *Consequentialist Retributivism*

The CR model would allow and even endorse plea bargaining and other such prosecutorial practices if they facilitated increases in the total amount of desert-based punishment the system could achieve. Unlike the absolutist and threshold models, the CR approach enables this result because it views the moral value of imposing deserved punishment in agent-neutral, not agent-relative, terms.[112] That is, the CR view sees the goal of desert-based punishment in *systemic* terms as something to maximize overall, so that a sacrifice of desert is allowed in one case if it enables a more-than-offsetting gain in another case. The other two models, by contrast, see retribution as a duty binding institutional actors in each individual case, prohibiting such tradeoffs. As Michael Moore explains:

> [T]he consequentialist-retributivist will intentionally refuse to punish guilty persons whenever more guilty persons (or greater

---

imposed was *more* than the person deserved (say, a draconian mandatory minimum). In such a situation, no matter what the prosecutor did, there would be a failure to impose all (and only) deserved punishment. What, then, is a retributivist prosecutor to do?

The absolutist model's likely answer is that the prosecutor should pursue conviction, on the ground that failing to do so would involve a voluntary decision on the prosecutor's part to violate a moral duty, whereas the overly severe punishment was not the prosecutor's own choice but was imposed by the legislature. Notice that if the legislature's expectation in enacting the law was that prosecutors would use the authorized sentence to bargain down to a lesser charge, or use their discretion to decline to prosecute except in serious cases, the absolutist model would lead to prosecutorial decisions frustrating the actual legislative agenda.

111. MOORE, *supra* note 14, at 158.
112. *See supra* Part I.C.

guilt) will be punished thereby. For the consequentialist-retributivist, no matter how intrinsically good it is that the guilty receive their deserts, more of that good is to be preferred to less of it.[113]

Thus, CR is the only model of real-world retribution that would condone various existing, and indeed routine, prosecutorial practices.

C. Feedback Effects on Legislative Action

A final aspect of real-world retribution worth noting is the possibility that the legislature itself, even if pursuing a retributive agenda, would be explicitly practical rather than principled in doing so. Such a legislature might enact rules that violate an ideal set of retributive rules, but do so for the sake of making retributively "right" outcomes more likely in actual cases. For example, the legislature might refuse to enact certain defenses that might properly bear on a person's desert in the abstract, but are seen as susceptible to manipulation and abuse and therefore likely to frustrate punishment of the deserving, rather than promote exoneration of the innocent.[114] The legislature might also devise substantive or evidentiary rules designed to promote accurate decisions in the majority of cases, though it is acknowledged that these rules would generate results that run counter to desert in some cases.[115]

Finally, at least under a CR model of criminal justice, the legislature might manipulate offense definitions or punishment levels based on its awareness of how police and prosecutors exercise their discretion.[116] For example, since CR allows prosecutorial plea bargaining, a CR legislature that anticipates bargaining and seeks to facilitate desert-maximizing bargains might deliberately authorize punishments exceeding an offense's deserved punishment, expecting that prosecutors will use those higher punishments as a bargaining chip to negotiate pleas that approximate the offender's actual desert. A retribution-oriented legislature might also

---

113. MOORE, *supra* note 14, at 157.

114. *See* ROBINSON & CAHILL, *supra* note 14, at 27**B**51.

115. *See id.* at 52**B**71; *see also* Bierschbach & Stein, *supra* note 6 (discussing interaction of substantive and evidentiary rules in promoting retributive justice).

116. *See* William J. Stuntz, *The Pathological Politics of Criminal Law*, 100 MICH. L. REV. 505, 550–52 (2001) (discussing political and practical incentives for legislatures—whether driven by retributivist considerations or not—to enact overbroad criminal laws and overly high criminal penalties based on expectations about how enforcement will work); *cf.* Paul H. Robinson & Michael T. Cahill, *The Accelerating Degradation of American Criminal Codes*, 56 HASTINGS L.J. 633, 635–44 (2005) (providing examples of the trend toward overcriminalization and overly high penalties).

knowingly define offenses to be overbroad (in the sense of capturing some morally blameless conduct), relying on retribution-oriented prosecutors to exercise their discretion not to prosecute in cases where a legally guilty person is not also morally blameworthy.

The possible spillover effects of real-world considerations into the substance, rather than merely the enforcement, of a retribution-seeking criminal law are worth noting here, but are somewhat tangent to this Article's project. The goal here is to elaborate whether a retributive theory has any potential to provide a comprehensive scheme of criminal justice, rather than to anticipate or resolve every complication involved in implementing such a scheme. Still, it is worth noting that a real-world retributive model might have to recognize and account for the possibility that practical considerations might affect not only the priorities of enforcement agents but also the content of the substantive law.

III. WEIGHING THE OPTIONS

A. The Absolutist Model: A Practical Fiasco

The absolutist view seems patently unworkable, and probably affirmatively undesirable, as it leads to real-world outcomes nobody would want. First, under at least some versions, this approach offers no clear reason to distinguish between more serious crimes and less serious ones. The amount of punishment must be proportionate to the offense, but the moral *duty* to punish is not scaled to reflect the offense's gravity: as an absolute *ex ante* imperative, it categorically mandates punishment of petty thieves and murderers both.[117]

Second, and perhaps more importantly, this view demands that we punish every known offender to the full extent of that person's desert.[118] The police would not be able to let anyone off with a warning, even if doing so would give them time to find and arrest other, more serious offenders.[119] Prosecutors would have to refuse to grant immunity to an offender who offers to identify a more serious offender (or ten other more serious offenders), and they would also be unable to balance the severity of punishment against its certainty by offering a plea bargain to prevent a

---

117. *See supra* notes 38–41 and accompanying text (discussing the demands of the absolutist view).
118. *See supra* notes 38–41 and accompanying text.
119. *See supra* Part II.A.1.

potentially inaccurate acquittal (or possible jury nullification).[120] Beyond these particular practical difficulties, the general nature of the absolutist moral obligation is such that it creates an odd prioritization of known, identified offenders (whom the state must punish) over yet-undiscovered offenders (toward whom no such obligation has formed). Basing the state's punishment agenda on the happenstance of which criminals it has already found, versus those it might find later, seems hard to defend in any principled way,[121] and also might lead to perverse or distorted enforcement or prosecutorial incentives.[122]

B. The Threshold Model: A Conceptual Muddle

The threshold view of retributivism, advanced most prominently by Michael Moore,[123] has been persuasively criticized by other commentators.[124] The central difficulty of the approach is the jarring and seemingly unsupportable discontinuity of having deontological concerns dominate decisionmaking below the threshold, while consequentialist concerns become suddenly paramount once the threshold point is reached. What was seen as a binding *ex ante* principled commitment below the threshold is apparently no longer binding once the threshold is reached.[125] Further, any decision about where to locate the threshold will inevitably be arbitrary on some level.[126] Finally, because the threshold is basically a

---

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

121. *See* KAPLOW & SHAVELL, *supra* note 7, at 325 n.68 (noting that distinguishing between uncaught and caught-but-unpunished offenders for purposes of determining a duty to punish seems untenable even within a retributivist perspective).

122. Such a scheme might encourage both police and prosecutors to adopt a sort of "willful ignorance," refusing to pursue leads or gather evidence about some offenders for the sake of avoiding the binding obligation to prosecute them to the full extent of their desert.

123. *See supra* notes 44–51 and accompanying text.

124. *See* Alexander, *supra* note 70; Christopher, *supra* note 5, at 151–57; Nancy Davis, *Contemporary Deontology*, *in* A COMPANION TO ETHICS 205 (Peter Singer ed., 1991); Anthony Ellis, *Deontology, Incommensurability and the Arbitrary*, 52 PHIL. & PHENOMENOLOGICAL RES. 855 (1992).

Alexander notes some additional possible criticisms beyond those I describe in the text. Alexander points to the work of John Taurek, who raises the question "whether similar harms to different people should be aggregated in considering where our moral duties lie." Alexander, *supra* note 70, at 897 (citing John Taurek, *Should the Numbers Count?*, 6 PHIL. & PUB. AFF. 293, 293 (1977)). If such aggregation is inappropriate, then it would never be acceptable to impose harm on one person to prevent similar harm to *any* number of other people. Alexander also points to complications that arise when the likelihood or timing of the relevant consequential harms varies or is uncertain. *See* Alexander, *supra* note 70, at 904–05.

125. As Moore himself asks, "Why should goodness of consequences not count at all and then, at some point, count enormously in the sense that it fully determines the rightness of action?" MOORE, *supra* note 14, at 723.

126. *See* Ellis, *supra* note 124, at 859–70.

binary switch where consequences go from "off" (irrelevant) to "on" (paramount), bizarre or untenable results are bound to occur at decision points just above and below the threshold.[127]

Moore offers a response to these critiques. He claims that below the threshold, consequences are not irrelevant but simply have not yet risen to the level where they outweigh the deontological principles. He gives the example of a dam, where water does not spill over until it reaches the top, but even the water at the bottom is contributing to the spillover.[128] As to the charge of arbitrariness in establishing any threshold level, he claims the issue amounts to what philosophers call a "sorites paradox" involving vague predicates[129]—taking the classic example, our inability to state exactly how many stones make a "heap" does not mean there is no such thing as a heap of stones.[130] Similarly, Moore argues, any disagreement

---

127. *See* Alexander, *supra* note 70, at 900 ("[O]ne might expect that moral outcomes in the neighborhood of such a radical discontinuity might seem weird or counterintuitive. And that expectation is borne out."). Suppose that *N* is the crucial threshold number—say, the number of lives to be saved—that would provide a consequentialist justification for otherwise deontologically forbidden acts such as torturing someone for information. By definition, this means that a threat to *N*–1 lives does not justify such torture, so eliminating the peril to just *one* of *N* jeopardized lives would suddenly make the prospective torture switch from morally acceptable (or even required) to morally unacceptable. On the other hand, if *N*–1 lives were in jeopardy from a potential bomb attack, enforcement authorities could deliberately move one more person into the zone of danger and thereby suddenly acquire the right to torture someone for information. *See id.* at 900▮03 (providing several examples of such possible anomalies or deliberate manipulations around the threshold).

The essential argument here is that, faced with a risk of consequential harm exceeding the threshold, all one must do is prevent just enough harm to return the total below the threshold, at which point consequential considerations are by hypothesis no longer compelling. Accordingly, all harms below the final increment that surpasses the threshold "are pure moral ballast," propping up that final increment, and can safely be ignored if the final increment is eliminated. *Id.* at 901. Accordingly, "a deontological threshold invites manipulation" around the margins of the threshold to achieve the desired result: one can eliminate just enough marginal risk so that consequentialist concerns no longer count, or *create* just enough marginal risk to get over the threshold and make consequentialist harms dominate the moral analysis. *Id.* at 903.

128. *See* MOORE, *supra* note 14, at 723 (crediting the analogy to Joseph Raz). Alexander questions the dam analogy, which implies that once the threshold level is reached, the acceptable tradeoff between harm imposed and harm avoided is one to one: for example, if one were allowed to take a single life to save 100 others, then one could take 2 lives to save 101 others, for "once the level of the dam is reached, one needs a foot of additional sandbagging to hold back each foot of water above that level." Alexander, *supra* note 70, at 899. Alexander claims (correctly, in my view, and likely in Moore's view as well) that the proper "conceptualization of thresholds is to see them as ratios" involving a geometric rather than linear relationship between the marginal harm caused and the marginal harm averted: if one may take a single life to save 100, one may only take 2 lives to save 200. *Id.*

129. *See Sorites Paradox*, *in* STANFORD ENCYCLOPEDIA OF PHILOSOPHY, *available at* http://plato.stanford.edu/entries/sorites-paradox/. "Sorites" is derived from the Greek word for "heap." *Id.*

130. *See* MOORE, *supra* note 14, at 724 ("[T]his is no more than the medieval worry of how many stones make a heap.").

about where exactly to locate the threshold does not mean it is nonsensical to recognize that consequences must matter at some point.[131]

These analogies fail to rebut the substance of the critics' claims, however. The dam metaphor merely exemplifies the peculiar focus of this approach on the final increment of harm that passes the threshold, while marginalizing other harms: we may safely ignore the buildup of water behind a 100-foot dam until it rises a single inch above the top, at which point our real concern is not with all the water, but only the top inch.[132] The sorites-paradox analogy is inapt, for the problem of identifying the threshold is not analogous to the problem of defining a heap. First, the concept of a heap exists along a continuum, with certain groups of stones seeming more or less like a heap, whereas the threshold is binary and discontinuous: consequences do not control, then they do.[133] Further, nothing of moral consequence turns on whether a certain number of stones are or are not a heap, whereas the threshold between caring and not caring about consequences is deeply important, requiring some degree of specification and justification.[134] The threshold-setting issue, therefore, is not just a vague-predicate problem but a gap in the theory.[135] The gap becomes clearer when the focus of analysis is not a general question of what conduct to criminalize, but an effort to steer the routine decisions of numerous police and prosecutors who need guidelines regarding whether and when it is acceptable to let someone off without charge, grant immunity, plea bargain, and so on.[136]

Another question about threshold deontology is whether it really collapses into full-fledged consequentialism.[137] The critique here is that by

---

131.   *See id.*

132.   *See* Alexander, *supra* note 70, at 901 ("Like the water below the level of the dam, [the] only moral role [of harms to people below the threshold] is to boost others above that level so that they become morally entitled to be rescued.").

133.   *See id.* at 908–09 ("[I]n order to make use of the sorites analogy, the threshold deontologist must reject the idea that there are discontinuities as we move from deontologically forbidden to consequentially required.").

134.   *See* Ellis, *supra* note 124, at 859; Alexander, *supra* note 70, at 909 ("Denying that there is a specific cutoff point in threshold deontology is arbitrary and troubling in a way that is disanalogous to typical sorites examples."); Christopher, *supra* note 5, at 155 n.353 ("[W]hile nothing of significance rests on whether the number of stones chosen as comprising a heap is arbitrary, the dividing line between what is morally right and wrong should not be arbitrary.").

135.   *See* Alexander, *supra* note 70, at 905–10.

136.   *See supra* Part II.B.

137.   *See, e.g.*, GLANVILLE WILLIAMS, CRIMINAL LAW—THE GENERAL PART 729 (2d ed. 1961) (noting this problem and endorsing consequentialism); MOORE, *supra* note 14, at 721 ("Many think that the [threshold version of an] agent-relative view just sketched, allowing as it does consequences to override moral absolutes when those consequences are horrendous enough, collapses into a consequentialist morality after all.").

allowing consequences to count sometimes, the threshold view allows the camel's nose under the tent. Another version of the argument suggests that the "deontological" rules of threshold deontology might really just serve as consequentialist heuristics. On this view, threshold deontologists are truly consequentialists at heart, and they are merely seeking to guide the consequentialist analysis by suggesting rules of thumb that will typically promote good results but may be ignored where they clearly would not.[138] Larry Alexander has noted such a possibility, though he does not think it accurately describes threshold theorists' understanding of their own position.[139] Michael Moore agrees that he is not a closet consequentialist and that threshold deontology is meaningfully distinct: "Even if the goodness of consequences is always relevant to the rightness of actions for a threshold deontologist, the goodness of consequences does not determine the rightness of actions as it does for a consequentialist."[140]

A truly consequentialist version of retributivism, then, provides a distinct approach from threshold deontology—and, as the next section argues, the consequentialist-retributivist view offers a superior method for bringing retributive justice into the real world.

C. Consequentialist Retributivism: A Workable Balance?

The consequentialist-retributivist (CR) view strikes a balance between adhering to retribution as a force that justifies and drives punishment and recognizing that competing real-world goals prevent an absolute, unyielding commitment to retribution. Unlike the absolutist and threshold models, the CR approach does not generate seemingly absurd results. In fact, the notion of maximizing overall (deserved) punishment seems to accord with the most natural, intuitive response to the problem of how retributive justice would or should work in practice. If the system is unable to impose all deserved punishment all the time, a natural second-best strategy is to impose as much deserved punishment as possible.[141]

---

138. *See* Alexander, *supra* note 70, at 910B11.

139. *See id.* at 911 ("In the end, however, I doubt that Moore—or Nagel or Nozick for that matter—want to grasp this lifeline. Nothing in what Moore or the others have written suggests that they are consequentialists at the level of their deepest moral beliefs.").

140. MOORE, *supra* note 14, at 722 n.132.

141. The reader may be inclined to wonder whether my support of a consequentialist perspective on retributive justice implies, or demands, that I endorse such a perspective across the board, thinking *all* principled commitments are better seen as goals to maximize *ex post*, instead of binding *ex ante* obligations.

   While that may be a sensible view, this Article's project is to focus on retributive justice specifically, and so I do not commit here to an overarching consequentialist view, nor am I convinced

As noted earlier, however, despite its powerful intuitive appeal as a strategy for achieving retributive goals in practice, the CR approach has, thus far, apparently found no adherents in the academic literature.[142] Indeed, only two commentators seem to have discussed the idea at any length: Michael Moore has described the possibility of CR,[143] though without embracing the approach himself;[144] and David Dolinko,

my position here compels me to do that. *Cf.* Joel Feinberg, *The Classic Debate*, *in* PHILOSOPHY OF LAW, *supra* note 37, at 646, 648–49 (noting that a utilitarian "in respect to punishment" need not "be a general utilitarian across the board").

    Unlike some other possible principled commitments—such as, perhaps, respect for liberty—the principle of retribution is not merely passive in its demands on the state. Rather, retributive justice requires that the state affirmatively undertake to design and maintain institutions designed to achieve or fulfill the principle. In other words, punishment is necessarily costly—and, given the limits of human knowledge, necessarily imperfect—and therefore necessarily confronts practical resource constraints in a way other principled commitments may not.

    One might conceive of a version of retributivism that imposes no such affirmative duty but only imposes a negative burden on the state to avoid *violating* the principle of deserved punishment. In fact, my elaboration of the absolutist model of retributivism, as it would apply in practice, offers such a version of the theory. As the discussion of that model in the text makes clear, though, such a scheme would (in my view) generate undesirable or even absurd results and is therefore unsatisfying in this context. However, such a conception might work for a model of, say, distributive or corrective justice—demanding in the latter context, for example, that adjudicative institutions follow certain principles in resolving private disputes without obliging the state to actively pursue corrective justice on its own. (Note that under such a view, principled commitments operate in a way that accords with their usual description as side *constraints* on action, rather than as affirmative obligations to act.) In any event, I take no position here as to the overall necessity or desirability of seeing all moral commitments through a consequentialist lens.

    142. Of course, as the text notes, some recent criminal law work by writers such as Dan Markel (as well as Douglas Husak, *see generally* Husak, *supra* note 19), and even the recent work of Michael Moore (as committed a deontologist as anyone in the criminal theory community), has effectively adopted a CR perspective or at least suggested the importance of forcing retributive theory to accommodate practical concerns. Mitchell Berman's account of retributivism also seems to have some resonance with the CR approach. *See supra* note 37.

    The general lack of outright enthusiasm for CR in the criminal theory literature is perhaps even more interesting or unusual given that another body of recent philosophical work has debated the broad question of whether it is possible (and, if so, desirable) to "consequentialize" *any* deontological or otherwise nonconsequentialist moral theory—i.e., to translate the theory's moral rules into an expressly consequentialist framework. *See, e.g.*, James Dreier, *Structures of Normative Theories*, 76 MONIST 22 (1993); Jennie Louise, *Relativity of Value and the Consequentialist Umbrella*, 54 PHIL. Q. 518 (2004); Philip Pettit, *The Consequentialist Perspective*, *in* THREE METHODS OF ETHICS 92 (Marcia Baron et al. eds., 1997); Douglas W. Portmore, *Can an Act-Consequentialist Theory Be Agent Relative?*, 38 AM. PHIL. Q. 363 (2001); Douglas W. Portmore, *Consequentializing Moral Theories*, 88 PAC. PHIL. Q. 39 (2007); Campbell Brown, Consequentialise This (Sept. 10, 2004) (unpublished manuscript, on file with author). For purposes of this article, I take no position on these broader questions, *see supra* note 141, though obviously there is much in the larger pro-consequentializing perspective that would resonate with the CR view.

    143. Moore provides a discussion in his book *Placing Blame* that reprises an account he had offered earlier in a law review article. *See* MOORE, *supra* note 14, at 155–61; Michael Moore, *Justifying Retributivism*, 27 ISR. L. REV. 15 (1993). Dolinko's article was responding to Moore's law review article.

    144. In the chapter describing CR, Moore professes agnosticism. *See* MOORE, *supra* note 14, at

responding to Moore, has offered a critique of CR.[145] Interestingly, though, Moore's most recent work appears to support a version of CR for some aspects of the criminal justice system, noting that consequentialist concerns can (and indeed must[146]) guide "those who design the general shape of legal institutions,"[147] even as binding agent-relative obligations should continue to apply to individual actors within those institutions.[148] Though clearly not a total embrace of CR, this marks a significant shift for Moore[149] and indicates the appeal or even necessity of adopting a consequentialist perspective in the face of inevitable resource constraints.

Others have tacitly indicated support for CR without committing to it decisively or providing any extended elaboration of CR per se. Louis Kaplow and Steven Shavell, in a discussion oriented toward discrediting the retributivist approach, have observed that some version of CR may hold out the best (or only) promise of implementing retributive goals:

> It also seems that, if the retributive conception of fairness were to be made more complete, it would have more of a consequentialist character and thus resemble the kind of theory that retributivists have resisted. For example, it is hard to imagine that any theory could help us to decide how much to spend to catch violators without regard to the number of additional violators that might be caught as a consequence. Likewise, one would suppose that it would make a difference whether the additional violators that would be apprehended had committed serious or relatively trivial wrongs,

---

159 ("[N]othing in the present project requires that we decide between deontological versus consequentialist retributivism."). Elsewhere, however, Moore refers to his own "non-consequentialist views on morality," *id.* at 719, and embraces the threshold-deontology approach. *See id.* at 719В24 (discussing threshold view and contrasting it with consequentialism).

145. *See* David Dolinko, *Retributivism, Consequentialism, and the Intrinsic Goodness of Punishment*, 16 LAW & PHIL. 507 (1997). Russell Christopher has also given a succinct critique of CR, relying on Dolinko's arguments. *See* Christopher, *supra* note 5, at 148В51; *id.* at 148 (noting that Dolinko's article is "perhaps the only assessment of Moore's consequentialist retributivism").

146. *See* Moore, *Reflections*, *supra* note 66, at 1552 ("So at the level of justifying our general institutions of criminal . . . law, I take us all to be consequentialists.").

147. *Id.* at 1551.

148. *See id.; see also* Moore, *Patrolling the Borders*, *supra* note 66. Moore hints that even as to individual actors, these imperatives might not truly constitute deontological commands but some form of rule-consequentialist guidelines. *See* Moore, *Reflections*, *supra* note 66, at 1552 ("But this would be at most a kind of 'government house' deontology, a 'deontology' to be followed by agents because the legal institution requiring their adherence to it will produce the best consequences overall."). Yet Moore also indicates that he retains a commitment to "deontology for individual actors," contrasting it with his "consequentialism about institutions." *Id.* at 1552 n.79.

149. *See id.* at 1552 n.79 ("For a life-long committed deontologist, this is a troublesome conclusion to reach after all these years. Yet at the level of allocating scarce social resources, I think the concession is inevitable.").

suggesting that different value weightings would be put on different wrongful acts. . . . [A] complete retributive theory would probably have to be a complex consequentialist one, much less different *in structure* from welfare economics than one would have suspected. (It would still be different *in substance*, in that some criteria other than effects on individuals' well-being would determine the value of punishing offenders.)[150]

Finally, as noted above, Dan Markel has implied that his own brand of retributivism is a version of CR, but he does not offer a sustained explanation or defense of this aspect of his theory.[151]

This Article, then, offers the first explicit advocacy of CR. My support of CR follows from this Article's premise (that retributive justice, meaning punishment according to desert, is worth pursuing[152]) and its argument thus far (that CR is the only effective, or at least worthwhile, means of pursuing retributive justice in the real world). In short, as the Kaplow and Shavell discussion suggests, CR is the only version of a retribution-oriented account of criminal justice that might accommodate the various real-world considerations necessary for a theory to work in practice.

But does CR work in theory?[153] Dolinko directs his criticisms at this question. Dolinko notes that that the CR perspective might enable retributivism to surmount some objections to its amenability to practical implementation,[154] but he objects to CR as antithetical to the broader retributivist project in the abstract.[155] (It may be worth noting that

---

150. KAPLOW & SHAVELL, *supra* note 7, at 310 n.36. Elsewhere Kaplow and Shavell cite Moore's reference to the possibility of consequentialist retributivism (and Dolinko's response). *See id.* at 313 nn.43 ▉44.

151. *See supra* notes 58–64 and accompanying text.

152. As noted earlier, this Article does not seek to defend the principle of retribution but simply posits that retribution has some value. *See supra* notes 23–26 and accompanying text. Plenty, perhaps too much, has already been written in defense of the retributivist principle. *See generally* John Cottingham, *Varieties of Retribution*, 29 PHIL. Q. 238 (1979) (identifying, in previous scholarly literature, nine distinct approaches to expressing or defending a retributivist theory).

153. Economists seem to be the most common butt of the joke involving some head-in-the-clouds intellectual type asking, "I know it works in practice, but will it work in theory"? *See* Robert P. Merges, *Contracting into Liability Rules: Intellectual Property Rights and Collective Rights Organizations*, 84 CAL. L. REV. 1293, 1391 (1996); Ira P. Robbins, *Whither (Or Wither) Habeas Corpus?: Observations on the Supreme Court's 1985 Term*, 111 F.R.D. 265 (1986); *see also* Julian Barnes, *Hate and Hedonism*, THE NEW YORKER, July 7, 2003, at 74 (presenting the joke as involving a Frenchman's rejoinder to a practical English proposal).

154. *See* Dolinko, *supra* note 145, at 515 ("[A] consequentialist version of retributivism may appear desirable, because the inevitable error-proneness of punishment institutions creates a difficulty for deontological retributivism . . . .").

155. *See id.* at 510 ("'[C]onsequentialist retributivism' is not, in fact, a coherent form of

Dolinko's objective is to argue against retributivism in all its forms, rather than to defend a deontological version against the CR approach.)[156] His central claim is that CR is conceptually incoherent because it demands adherence to two incompatible propositions. The first proposition, which Dolinko calls "the intrinsic-good claim," is that "the end-result of inflicting deserved punishment [is] a good state of affairs,"[157] a view that treats retributive justice as a good *result* or consequence of punishing and therefore enables the CR position.[158] The second proposition, which Dolinko calls "the desert claim," is that retributive justice must be the *only* basis for imposing punishment, with any other considerations excluded.[159] These two propositions are incompatible, Dolinko notes, because the intrinsic-good claim "compels us to acknowledge that the goodness or badness of *other* consequences of inflicting punishment should also be relevant to whether the punishment is justified. And this runs afoul of . . . the desert claim . . . ."[160] This leads Dolinko to conclude that CR is an untenable view.[161]

Dolinko is correct that the intrinsic-good claim and the desert claim, as he presents them, are incompatible. This incompatibility is not fatal to CR, though, for CR does not demand adherence to both claims. Rather, CR adopts the intrinsic-good claim and abandons the desert claim.[162] This feature of CR is the very feature that distinguishes it from deontological retributivism: as a consequentialist theory, CR does not view desert as a fundamental *ex ante* obligation (or restraint), but rather as one good to be balanced against other possible goods in determining whether punishment

---

retributivism at all, given a quite standard view (which Moore espouses) as to what retributivism is."); *id.* at 513 ("[T]o be both a consequentialist and a retributivist is not, after all, a tenable position; it is well-nigh impossible, given a standard account of retributivism (such as Moore himself presents).").

156. *See id.* at 508 ("This paper will argue that a consequentialist version of retributivism actually encounters grave difficulties, and that these difficulties reflect the problematic nature of retributivism itself.").

157. *Id.* at 516.

158. *See id.* ("The first proposition—the intrinsic-good claim—generates the urge to view retributivism as a consequentialist theory.").

159. *See id.* (describing the desert claim as demanding that non-desert "consequences *cannot* be relevant to the justification of punishment").

160. *Id.*; *see also id.* at 513 (noting that "no plausible consequentialism could take [desert] as the *only* good-making characteristic"); Berman, *supra* note 8, at 48 ("[W]hile the effectuating of retributive justice could fairly count *among* the practice's rational justifications, it is unclear how, except by fiat, the retributivist can identify realization of this particular consequence as a justifying aim without allowing that the aim is, accordingly, to realize net good consequences generally.").

161. *See* Dolinko, *supra* note 145, at 515 ("[A] consequentialist version of retributivism appears impossible, at least if we insist that the meting out of just deserts is the only beneficial consequence of punishment relevant to its justification.").

162. *See supra* Part I.C (discussing general features of CR perspective).

is appropriate in a specific case. At the same time, CR differs from standard-issue consequentialism in that it treats desert-based punishment as having inherent, and significant, value.

Dolinko is certainly right that many retributivist theorists, including Moore, adopt the desert claim.[163] But none of them, including Moore, has adopted CR.[164] There is thus no inconsistency (at least in this regard) in their views, nor is there any inconsistency in adopting a contrary position that accepts the intrinsic-good claim but gives up the desert claim. The CR perspective does require, in some sense, a diluted commitment to retributive goals, for it must admit retribution to be just one of various goods worth pursuing. Unlike deontological approaches, which claim that desert is always both a necessary and sufficient condition of punishment, CR might have to take the softer view that desert is *nearly always* sufficient, and *in rare cases* not necessary, as a basis for punishing. But Dolinko is incorrect to conclude that this concession is fatal to CR—either by making CR incoherent or by making it toothless.[165] It is entirely possible to support a version of the retributive principle—that imposing deserved punishment is intrinsically good and imposing undeserved punishment (or failing to impose deserved punishment) is intrinsically

---

163. *See, e.g.*, MOORE, *supra* note 14, at 153 ("Punishment is justified, for a retributivist, solely by the fact that those receiving it deserve it. Punishment of deserving offenders may produce beneficial consequences other than giving offenders their just deserts. . . . Yet for a retributivist these are a happy surplus that punishment produces and form no part of what makes punishment just . . . ."); *see also* Dolinko, *supra* note 145, at 516 n.27 (citing retributivists advancing desert claims).

164. *Cf.* Dolinko, *supra* note 142, at 507–08 ("One feature on which retributivists and their critics have generally agreed is that retributivism is very much a non-consequentialist theory.").

165. Dolinko writes:

> [A] deterrence theorist . . . [may] believe that a guilty person's receiving his deserved punishment is, in itself and regardless of its consequences, a good state of affairs—i.e., is intrinsically good. That will not be the deterrence theorist's reason for inflicting the punishment . . . but it may nevertheless be an idea he accepts. . . .
>
>   What distinguishes the retributivist from the deterrence theorist cannot, therefore, be that only the former regards it as intrinsically good for a guilty person to be punished. Rather, what is distinctive about the retributivist must be the *role* played in her theory by the intrinsic goodness of punishing the guilty. For the retributivist, this intrinsic goodness cannot be an irrelevancy or a mere happy accident. It must be either (i) the reason for engaging in the practice of punishment (its rational justification), or (ii) the reason why that practice is morally permissible (its moral justification), or both.

Id. at 517–18. It seems to me that the penultimate sentence of this quote is accurate, but the final sentence is not—at least with respect to an adherent of CR. For such a person, the intrinsic goodness of punishing the guilty is no "irrelevancy or mere happy accident" but a strongly relevant factor in the overall calculus that weighs the good and bad aspects of the decision to punish or not. This is what distinguishes the CR adherent from Dolinko's "deterrence theorist," for whom the intrinsic goodness of desert (even if acknowledged) merits no weight in that calculus. At the same time, for the CR adherent, desert is not *the* justification (rational or moral) for punishing, but merely one (highly significant) criterion contributing to the consequentialist balancing.

bad—while acknowledging that other goods or bads may outweigh application of that principle in some cases.[166]

Of course, asserting that CR seems like the only feasible way to advance retributivist goals in the real world is not the same thing as offering a full articulation of how CR would work in practice. That complicated task is beyond the scope of this Article, which seeks only to offer a generalized account of how a real-world criminal justice system *might* pursue a retributive agenda, rather than a particular account of how a given system would or should do so.

Even for present purposes, however, it is worth noting the main general implications of the CR approach and the issues any effort to employ it would confront. Foremost among these is the significant difficulty of ascertaining how much weight the "goodness" of desert should merit in the overall consequentialist analysis. The very notion of calculating the value of desert along some metric that would enable its comparison to other goods, like crime reduction or cost savings, might seem either odd or patently impossible.[167] Yet almost any cost-benefit analysis inevitably involves difficult or even unsavory apple-orange comparisons of the "how much money is a human life worth" sort. Without minimizing the complexity (and perhaps even the ineradicable moral contestability) of such judgments, they are necessarily made all the time, and recognizing the tradeoffs they involve at least serves to focus attention and debate.

Any further exploration of just *how* good desert is relative to other consequences is, again, beyond the scope of the present Article, which seeks only to provide a general framework under which retributivism can maintain any relevance as a real-world, rather than idealized, vision of criminal justice without either overwhelming other considerations entirely (as under deontological theories) or being dismissed as irrelevant (as under standard utilitarian consequentialism). One possible way, though by no means the only way, to approach the question might be to collect empirical data regarding the extent to which people tend to value achieving desert relative to other goals, such as procedural fairness. Some of this work is already being done, with interesting and promising results.[168]

---

166. *See* Joel Feinberg, *Justice and Personal Desert*, *in* DOING AND DESERVING 60 (1970) ("[A] person's desert of X is always a reason for giving X to him, but not always a conclusive reason, [as] considerations irrelevant to his desert can have overriding cogency in establishing how he ought to be treated on balance.").

167. *But cf.* ROBINSON & CAHILL, *supra* note 14, at 137–39, 183–85 (proposing an instrumentalist framework for considering the value of desert for purposes of trading it off against other values, such as procedural fairness).

168. *See, e.g.*, *id.* at 183–85 (discussing empirical studies of relative significance of procedural

A corollary to the difficulty of specifying the value of retributive justice as a good is that the CR view, while having the advantage of offering a feasible *means* for achieving retributive justice in the real world, has the drawback of no longer offering a *mandate* for doing so. By framing retribution as a moral duty, standard deontological retributive theories evade the thorny question of whether achieving retributive justice is worth the costs it imposes.[169] If retributive justice is no longer a moral imperative but merely a good like other goods, some further explanation is required as to why we must seek this good rather than others (including the good of cost savings),[170] or why the state rather than private individuals must be the sole repository of the power to punish.[171] Further, under the CR view, desert also becomes *a* reason to punish, but not *the* reason to punish; other good consequences, like deterrence, might also justify punishment in some cases or might be needed to justify punishing at all, given its costs.[172] CR provides a practical method for implementing

---

169. Some commentators have noted, however, that the justification of retributive institutions should face the same questions even if retribution is seen as a moral duty. *See* Husak, *supra* note 19, at 999 ("The difficulty I have mentioned [of needing to show that achieving retribution is worth its costs] is not resolved if we hold that society has not only the *right*, but also the *duty*, to impose deserved punishment. . . . The burden is not merely to show that the imposition of punishment is a duty, but also to show that it is a duty of sufficient weight or stringency to justify the creation of an institution with the drawbacks I have recounted.").

170. *See* DUFF, *supra* note 5, at 199 (noting the "logical gap between the claim that wrongdoers deserve to suffer—that it is good that they should suffer—and the claim that it is for us, or for the state, to ensure that they suffer"); Dolinko, *supra* note 145, at 519 ("Once one acknowledges an irreducible plurality of intrinsic goods, 'X is intrinsically good' no longer leads straightforwardly to 'There is an obligation to bring about X'—let alone to an obligation to create *official institutions* to bring X about!"); Douglas N. Husak, *Retribution in Criminal Theory*, 37 SAN DIEGO L. REV. 959, 975 (2000) ("The intrinsic goodness of realizing a principle of retributive justice would justify punishment in a possible world (such as a divine realm) . . . . Unfortunately, that possible world is not our world. In our world, we must sympathize with citizens who balk when asked to fund an institution that has the sole objective of realizing retributive justice.").

171. *See, e.g.*, Dolinko, *supra* note 145, at 522 ("Rationally justifying the institution of punishment requires showing not merely that punishment is important, but that it cannot or should not be left in private hands. And showing *that* seems to demand attention to the contrasting consequences of private vengeance and publicly-imposed punishment."); Husak, *supra* note 170, at 972 ("If I am correct, our retributive beliefs only require that culpable wrongdoers be given their just deserts by being made to suffer . . . . These beliefs do not require that culpable wrongdoers be given their just deserts by being made to suffer by the state through the imposition of punishment."). *But see* Alon Harel, Why Only the State May Inflict Criminal Sanctions (Sept. 25, 2007), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1017298; *cf.* Kenworthey Bilz, *The Puzzle of Delegated Revenge*, 87 B.U. L. REV. 1059 (2007) (providing psychological account to explain citizens' willingness to "delegate" punishment power to state).

172. *See* Husak, *supra* note 170, at 976–77 ("What is needed . . . is some additional value that punishment can be expected to attain—a value which, when added to the value of realizing retributive justice, will justify the institution of punishment. This value, I submit, is crime reduction. . . . The benefits of crime reduction, when added to the intrinsic goodness of achieving retributive justice

retributive goals, but at the considerable cost of denying those goals primacy in the conceptual justification of criminal law.[173]

A second obvious complication is the task, noted earlier, of establishing an appropriate measure of the relative seriousness of different crimes for purposes of maximizing desert-based punishment.[174] Absent the ability to weigh the gravity of one offense against another, CR cannot establish an enforcement agenda. Whether the actual punishment levels set by the legislature for different offenses serve as an appropriate proxy for offenses' relative seriousness is an unresolved, and highly debatable, proposition.[175]

One further question raised by the potential adoption of CR as a guide to enforcing the criminal law is whether, for the sake of conceptual consistency, the CR perspective should drive the substance of criminal law as well. Under such a view, criminal prohibitions, like the mechanisms for enforcing those prohibitions, would apply in an agent-neutral fashion. One possible consequence of this perspective would be the radical expansion of justification defenses to allow violation of criminal rules any time doing so would prevent an even greater violation or set of violations. Unlike the existing lesser-evils or self-defense justifications, this perspective would impose no limitation on the means chosen to minimize the moral wrong, such as a ban on initiating harm against an innocent party: A could kill B now to prevent C from killing two or more other people later.[176]

(hopefully) offset the drawbacks of punishment. Contrary to pure retributivists like Moore, a theory of criminal law and punishment must find some justificatory role for crime reduction."); KAPLOW & SHAVELL, *supra* note 7, at 313 ("[M]any retributivists might, on reflection, concede that the consequences of punishment—in terms of the number of wrongful acts and the number of instances of unfair punishment—do matter to normative evaluation. Otherwise, their position is inconsistent with their belief that the occurrence of wrongful acts and the imposition of unfair punishment each involve injustice. But if such consequences matter, then deterrence would become highly relevant for retributivists, a conclusion that runs counter to their professed position.").

173. *See generally* Dolinko, *supra* note 145, at 518–27 (describing how the conception of retributive justice as an intrinsic good fails, without more, to justify the practice of state-imposed punishment); *see also supra* notes 36–37 (describing accounts of retribution under which punishment is permissible but not mandatory).

174. *See supra* Part II.A.3.

175. *See* Paul H. Robinson, Michael T. Cahill & Usman Mohammad, *The Five Worst (and Five Best) American Criminal Codes*, 95 NW. U. L. REV. 1, 18–20, 51–59 (2000) (describing methods and results of analysis exploring the extent to which existing criminal codes grade offenses in proportion to their relative seriousness).

176. *See* MOORE, *supra* note 14, at 74–75, 679–80.

CONCLUSION

Standard retributive-justice theories offer justifications for having institutions of punishment but have surprisingly little to say about how those institutions should actually do their work. This Article has sought to describe how the retributive principle might fit into a full theory of criminal justice, rather than only a theory of punishment. The most sensible means for incorporating retributive goals into a general theory of criminal justice is to adopt the approach of "consequentialist retributivism," or CR.

The CR approach may seem a mixed blessing for many retributivists, however, for it not only giveth, but also taketh away. While potentially offering a means to expand the application of retributive principles beyond mere abstraction and into the real world, CR also undercuts the independent force of retributivism as a theory of punishment. Under CR, unlike "standard" deontological retributivist theories, desert is no longer a freestanding absolute basis for imposing or withholding punishment. Rather, the underlying justification for punishment rests on a more general balancing of its consequences, and retributive goals count as only one factor in that balancing. Retributive justice, then, may enter the real world, but in doing so, it must allow other considerations to share space on the ideal plane where criminal law's motivating principles reside.