

2003

## The Prosecutor's Dilemma: Bargains and Punishments

Russell L. Christopher

Follow this and additional works at: <https://ir.lawnet.fordham.edu/flr>



Part of the [Law Commons](#)

---

### Recommended Citation

Russell L. Christopher, *The Prosecutor's Dilemma: Bargains and Punishments*, 72 Fordham L. Rev. 93 (2003).

Available at: <https://ir.lawnet.fordham.edu/flr/vol72/iss1/4>

This Article is brought to you for free and open access by FLASH: The Fordham Law Archive of Scholarship and History. It has been accepted for inclusion in Fordham Law Review by an authorized editor of FLASH: The Fordham Law Archive of Scholarship and History. For more information, please contact [tmelnick@law.fordham.edu](mailto:tmelnick@law.fordham.edu).

# THE PROSECUTOR'S DILEMMA: BARGAINS AND PUNISHMENTS

*Russell L. Christopher\**

Unlike consequentialists, retributivists condemn bargain justice—plea bargains and immunity deals—as violating a number of conditions of just punishment. While the difficulties prosecutor-initiated bargaining poses for prisoners in a consequentialist system of criminal justice is well known under the rubric of the prisoner's dilemma, this Article introduces the difficulties prisoner-initiated bargaining poses for prosecutors in a retributivist system of criminal justice. Applying four central principles of retributivism to a prisoner-initiated proposal offering inculpatory testimony on the prisoner's confederates in exchange for immunity demonstrates that each principle paradoxically requires, and is violated by, both acceptance and nonacceptance of the offer. Retributivism's conflicting principles generate a moral dilemma—the Prosecutor's Dilemma—in which no matter what a prosecutor does in response to the offer, retributivism's principles will be violated. According to the means preferred by retributivists themselves to resolve conflicts among principles, the Prosecutor's Dilemma is to be resolved by not accepting the offer and punishing neither the prisoner nor her confederates. Though perhaps theoretically sound, the resolution is a practical disaster. It allows prisoners to self-immunize merely by tendering a dilemmatic immunity offer. While the resolution perhaps reduces retributivism to the point of absurdity by disabling retributivism from justifying punishment of either the prisoner or her confederates, without the resolution of the Prosecutor's Dilemma retributivism remains internally inconsistent. As a result, the Prosecutor's Dilemma supplies an indirect defense of bargain justice.

---

\* Assistant Professor of Law, The University of Tulsa College of Law. I am indebted to George P. Fletcher, Kent Greenawalt, Ken Levy, and Peter B. Oh for their criticisms of earlier drafts of this Article. I also benefited from discussions with Charles Adams, Martin Belsky, Montie Deer, Paul Finkelman, Janet Levit, Marla Mansfield, and Maria Pagano. Any errors are, of course, my own. A previous version of this Article was presented at a Faculty Colloquy at The University of Tulsa College of Law.

## TABLE OF CONTENTS

INTRODUCTION.....	95
I. BARGAIN JUSTICE AND THE THEORY OF PUNISHMENT .....	106
A. Bargain Justice.....	106
1. Plea Bargain .....	107
2. Immunity Deal .....	109
B. Assessing Bargain Justice via the Theories of Punishment.....	113
1. Consequentialist Justification.....	114
2. Retributivist Critique .....	118
II. ASSESSING RETRIBUTIVISM VIA BARGAIN JUSTICE.....	122
A. The Prosecutor's Dilemma: A Hypothetical Immunity Offer.....	123
B. Retributivist Principles Applicable to the Immunity Offer.....	124
1. Categorical Duty to Punish Culpable Wrongdoers ....	125
2. Giving Culpable Wrongdoers Their Just Deserts .....	127
3. Culpable Wrongdoers' Right to Punishment .....	128
4. Treating Culpable Wrongdoers Not as Mere Means But as Ends.....	129
5. Treating Culpable Wrongdoers with Equal Desert Equally .....	131
C. Three Options.....	134
1. Accepting the Offer and Punishing Only the Two ....	134
2. Not Accepting the Offer and Punishing Only the One .....	134
3. Not Accepting the Offer and Punishing None .....	135
D. Conclusion.....	135
III. INCONSISTENCY AND MORAL DILEMMA.....	135
A. Inconsistency.....	136
1. Inconsistency Within Retributivism's Principles.....	136
2. Inconsistency Among Retributivism's Principles .....	136
B. Moral Dilemma.....	137
C. Attempts to Avoid the Prosecutor's Dilemma .....	142
1. Refusal to Participate in Bargain Justice Because It Is Unjust.....	142
2. Duty to Punish Limited to Convicted Offenders.....	143
D. Conclusion.....	145
IV. ATTEMPTS TO RECONCILE RETRIBUTIVISM WITH BARGAIN JUSTICE.....	145
A. Consequentialist Retributivism .....	146
1. Deontological v. Consequentialist Retributivism .....	146

2. Untenable Form of Retributivism .....	148
3. Applicability to Offer .....	149
B. Threshold Retributivism.....	151
1. Threshold Retributivism's Relation to Threshold Deontology.....	152
2. Untenable Form of Retributivism .....	153
3. Applicability to Offer .....	156
C. Conclusion.....	157
V. RESOLVING THE PROSECUTOR'S DILEMMA .....	158
A. Prioritizing Among Retributivism's Duties: Negative v. Positive .....	158
B. Deprioritizing the Equality Principle.....	164
C. Practical Ramifications of the Resolution.....	166
CONCLUSION.....	168

## INTRODUCTION

Everybody loves a bargain, except when it concerns crime and punishment. Commentators widely decry bargain justice<sup>1</sup> as

---

1. Bargain justice means here an arrangement between a prosecutor and purported offender in which each party exchanges something of value with the other party. Rather than determining what an offender justly deserves by a trial, a prosecutor and offender bargain as to the terms of what each shall receive from the other. The two main forms of bargain justice are the plea bargain and the immunity deal. In both forms, by providing something of value to the prosecutor—a guilty plea or inculpatory information on another's criminal activity—the offender receives less punishment than if she went to trial and was convicted. For a discussion of these two practices, see *infra* Part I.A.1-2. For the roots of plea bargaining in England, dating back as far as 1587, see Albert W. Alschuler, *Plea Bargaining and Its History*, 79 Colum. L. Rev. 1, 12-16 (1979); Lawrence M. Friedman, *Plea Bargaining in Historical Perspective*, 13 Law & Soc'y Rev. 247, 247-48 (1979); John H. Langbein, *Understanding the Short History of Plea Bargaining*, 13 Law & Soc'y Rev. 261, 261-69 (1979). Though their first instances may be unknown, immunity deals apparently were common in the mid-18th century in Italy and were employed there as early as 1630 in the Italian city of Milan. Cesare Beccaria, *On Crimes and Punishments* 102 n.1 (David Young trans., 1986) (1764) (citation omitted).

For an example of the negative light in which bargain justice is typically cast, consider the following account:

For all but the most doctrinaire free market micro-economists, the very term "plea bargaining" is distasteful. Justice, critics of plea negotiations point out, is not a matter of bazaar haggling, but of thoughtful adjudication of claims. "To bargain" is different than "to present reasoned arguments," and to reach a "bargain" or even an "agreement" is different than to obtain a "judgment." The more colloquial meaning of "bargain" is even worse. A "bargain" is a discount, something obtained at a cut-rate. If judges and scholars squirm at the notion of "bargained justice," the public, especially in our fearful times, is even more unhappy about defendants who "get off cheap."

Gerard E. Lynch, *Our Administrative System of Criminal Justice*, 66 Fordham L. Rev. 2117, 2129-30 (1998). Lynch himself, however, who as a former prosecutor and defense attorney and currently a federal judge and law professor is well-situated to

“absolutely and fundamentally immoral,”<sup>2</sup> “a disaster,”<sup>3</sup> “unfair and irrational,”<sup>4</sup> “outrageous,”<sup>5</sup> and unjust.<sup>6</sup> Perhaps the single most influential source of the animus against bargain justice springs from retributivism.<sup>7</sup> Starting in the late eighteenth century with Immanuel Kant’s declaration that “justice ceases to be justice if it can be bought at any price whatsoever,”<sup>8</sup> retributivists have nearly uniformly rejected the wholesale purchase of justice.<sup>9</sup> But apparently finding

understand the practice from multiple perspectives, finds the above-quoted account not entirely apt. *Id.* at 2130.

2. Marshall J. Hartman & Marianna Koval, *The Immorality of Plea Bargaining, in Legality, Morality, and Ethics in Criminal Justice* 70, 70 (Nicholas N. Kittrie & Jackwell Susman eds., 1979).

3. Stephen J. Schulhofer, *Plea Bargaining as Disaster*, 101 Yale L.J. 1979, 2009 (1992) (plea bargaining “can be, and should be, abolished”).

4. Albert W. Alschuler, *The Changing Plea Bargain Debate*, 69 Cal. L. Rev. 652, 652-56, 668-69 (1981); Lynch, *supra* note 1 at 2130 (“no one pictures ‘plea bargaining’ as a rational way to determine guilt or innocence”).

5. Kenneth Kipnis, *Criminal Justice and the Negotiated Plea*, 86 Ethics 93, 105 (1976).

6. R.A. Duff, *Trials and Punishments* 141 (1986).

7. Simply stated, under retributivism, “punishment is justified because people deserve it.” Kent Greenawalt, *Punishment*, 74 J. Crim. L. & Criminology 343, 347 (1981). John Rawls supplies the following concise account:

What we may call the retributive view is that punishment is justified on the grounds that wrongdoing merits punishment. It is morally fitting that a person who does wrong should suffer in proportion to his wrongdoing. That a criminal should be punished follows from his guilt, and the severity of the appropriate punishment depends on the depravity of his act. The state of affairs where a wrongdoer suffers punishment is morally better than the state of affairs where he does not; and it is better irrespective of any of the consequences of punishing him.

John Rawls, *Two Concepts of Rules*, 64 Phil. Rev. 3, 4-5 (1955). The common denominator of other retributivist accounts is to locate the justification of punishment in relation to a past offense or wrongdoing. *E.g.*, George Fletcher, *Rethinking Criminal Law* 416-17 (1978) (“Retribution simply means that punishment is justified by virtue of its relationship to the offense that has been committed.”). Another approach is to explain retributivism by recourse to a “cluster of moral concepts: rights, desert, merit, moral responsibility, and justice.” Jeffrie G. Murphy & Jules L. Coleman, *Philosophy of Law: An Introduction to Jurisprudence* 121 (1990). Retributivism, “in its most complete form,” is claimed to embody the following five principles:

- (1) The moral *right* to punish is based solely on the offense committed.
- (2) The moral *duty* to punish is also grounded exclusively on the offense committed.
- (3) Punishment ought to be *proportionate* to the offense (the *lex talionis*).
- (4) Punishment is the “*annulment*” of the offense.
- (5) Punishment is a *right* of the offender.

Igor Primoratz, *Justifying Legal Punishment* 12 (1989). For further discussion of retributivism, see *infra* Part I.B.2. For a discussion of five central principles of retributivism antithetical to bargain justice, see *infra* Part II.B.

8. Immanuel Kant, *The Metaphysics of Morals* 141 (Mary Gregor ed. and trans., 1991) (1797). Kant’s declaration follows his consideration and emphatic rejection of a bargain justice proposal. *Id.* For further discussion of Kant’s view of bargain justice, see *infra* notes 150-55 and accompanying text.

9. *E.g.*, R.A. Duff, *Punishment, Communication, and Community* 121 n.49

that the price is right, the actual players in the game—prisoners, prosecutors and courts—cannot seem to get enough of bargain justice.<sup>10</sup> Today, bargain justice, and not trials, disposes of as much as ninety percent of criminal cases.<sup>11</sup> “We now live in a world of guilty pleas [and bargain justice], not trials . . . .”<sup>12</sup> With the ascent of bargain justice as the predominant mode of criminal justice surprisingly coinciding with retributivism’s rise as the leading theory of punishment,<sup>13</sup> the gap between practice and theory has rarely been wider.<sup>14</sup>

---

(2001) (“If the criminal justice system aims to do justice, it should not bribe defendants to plead guilty . . . .”); Mark Tunick, *Punishment: Theory and Practice* 141 (1992) (“[F]or the retributivist, punishment must be deserved, and so plea-bargaining misses the point of why we punish.”). For further discussion of retributivists’ condemnation of bargain justice, see *infra* Parts I.B.2 and II.B.

10. For a possible explanation, see, for example, *Blackledge v. Allison*, 431 U.S. 63, 71 (1977) (plea bargains “benefit all concerned”). For further discussion of the benefits of bargain justice, see *infra* Part I.B.1.

11. See G. Nicholas Herman, *Plea Bargaining 1* (1997) (approximately ninety percent of criminal convictions in the United States result from guilty pleas, the overwhelming number of which are through plea bargains) (citation omitted); Paul H. Robinson, *Criminal Law* 591 (1997) (“Almost 95 percent of all criminal cases are disposed of through a plea of guilty, and a large portion of these pleas are the result of plea negotiations.” (citation omitted)); Mark Tunick, *Hegel’s Political Philosophy: Interpreting the Practice of Legal Punishment* 123 (1992) (“In our own criminal justice system, a vast majority of cases never go to trial. Perhaps as many as 90 percent of all cases are ‘plea bargained.’” (citation omitted)); Ernest Van den Haag, *Punishing Criminals: Concerning a Very Old and Painful Question* 171 (1975) (“More than 90 percent of all cases are settled by striking a bargain: the defendant is allowed to plead guilty to a lesser charge than the one originally brought so that in exchange for his lower punishment the trial can be avoided.” (citation omitted)); 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.”).

12. Stephanos Bibas, *Judicial Fact-Finding and Sentence Enhancements in a World of Guilty Pleas*, 110 *Yale L.J.* 1097, 1100 (2001) (arguing for greater scholarly attention to guilty pleas and plea bargains).

13. E.g., Mirko Bagaric & Kumar Amarasekara, *The Errors of Retributivism*, 24 *Melb. U. L. Rev.* 124, 124 (2000) (“Retributivism has been the dominant theory of punishment in the Western world for the past few decades.”); David Dolinko, *Three Mistakes of Retributivism*, 39 *UCLA L. Rev.* 1623, 1623 (1992) (“It is widely acknowledged that retributivism . . . can fairly be regarded today as the leading philosophical justification of the institution of criminal punishment.” (citation omitted)); *Developments in the Law—Alternative Punishments: Resistance and Inroads*, 111 *Harv. L. Rev.* 1967, 1970 (1998) (“[A] ‘just deserts’ philosophy associated with retributivism has claimed the mantle of penological predominance.” (citation omitted)). See Sanford H. Kadish, *Fifty Years of Criminal Law: An Opinionated Review*, 87 *Cal. L. Rev.* 943, 978-81 (1999) (chronicling the rise of retributivism); Nigel Walker, *Modern Retributivism*, in *Jurisprudence: Cambridge Essays* 73, 83 (Hyman Gross & Ross Harrison eds., 1992) (noting the ascent of retributivism). But see Kyron Huigens, *The Dead End of Deterrence, and Beyond*, 41 *Wm. & Mary L. Rev.* 943, 955 (2000) (contending that consequentialism “is probably dominant over retributivism in criminal law scholarship”).

14. Because retributivism rejects bargain justice, see *supra* notes 5, 6, text accompanying note 8, note 9 and *infra* Part I.B.2, the leading theory of punishment, see *supra* note 13, and the predominant mode by which defendants are rendered subject to punishment, see *supra* note 11 and accompanying text, are incompatible.

Though the legitimacy of bargain justice is widely questioned,<sup>15</sup> the ramifications of bargain justice for the principal theories of punishment are globally ignored.<sup>16</sup> The age-old debate over the justification of punishment,<sup>17</sup> stemming largely from the impasse between consequentialism<sup>18</sup> and retributivism,<sup>19</sup> tends to presume that

For six months in 1998, the gap between theory and practice narrowed appreciably. In *United States v. Singleton*, the Tenth Circuit Court of Appeals held that a prosecutor granting leniency for testimony against a defendant's confederates constituted a lesser form of bribe in violation of the federal witness gratuity statute, see 18 U.S.C. § 201(c)(2) (1994), prohibiting exchanging something of value to obtain testimony. See 144 F.3d 1343, 1357-58 (10th Cir. 1998), *rev'd en banc*, 165 F.3d 1297 (10th Cir. 1999). The ruling triggered shock, *Justice System Thrown Off Balance by Ruling*, Omaha World Herald, July 17, 1998, at A1 (Albert Alschuler, a notable long-time critic of plea bargaining, exclaimed, "it's a bombshell, it's nuclear, it's really a big deal"), and pointed criticism. See, e.g., *United States v. Eisenhardt*, 10 F. Supp. 2d 521, 521-22 (D. Md. 1998) (referring to *Singleton* as "amazingly unsound, not to mention nonsensical" and venturing that the odds of "the Supreme Court reaching the same conclusion as the *Singleton* panel are, in this Court's judgment, about the same as discovering that the entire roster of the Baltimore Orioles consists of cleverly disguised leprechauns"). Six months later, however, an en banc panel of the Tenth Circuit overruled the three-judge panel, thereby perpetuating the gap. *Singleton*, 165 F.3d at 1298 (holding that prosecutors engaging in plea bargaining do not fall within the ambit of the federal witness gratuity statute).

15. See, e.g., Albert W. Alschuler, *Implementing the Criminal Defendant's Right to Trial: Alternatives to the Plea Bargaining System*, 50 U. Chi. L. Rev. 931, 1048 (1983) ("The impediments to implementation of a plea bargaining prohibition are not worth a fraction of the paralysis that they have prompted."); Albert W. Alschuler, *The Prosecutor's Role in Plea Bargaining*, 36 U. Chi. L. Rev. 50, 65 n.43 (1968) (plea bargains "permit situations in which it is to the apparent advantage of innocent men to plead guilty"); Hartman & Koval, *supra* note 2, at 82 (terming plea bargaining as "the major abuse in our criminal justice system"); Kipnis, *supra* note 5, at 106 ("plea bargaining falls short of the justice we expect of our legal system"); Stephen J. Schulhofer, *Is Plea Bargaining Inevitable?*, 97 Harv. L. Rev. 1037, 1050-86 (1984) (describing a study supporting the plausibility of eliminating plea bargaining); Jeffrey Standen, *Plea Bargaining in the Shadow of the Guidelines*, 81 Cal. L. Rev. 1471, 1537-38 (1993) (prosecutors' overwhelming power in plea bargaining should be weakened).

16. The one notable exception is Mark Tunick, who is the only punishment theorist to even acknowledge that a theory of punishment could be assessed by its compatibility with, its ability to justify, or its application to, forms of bargain justice. Tunick, *supra* note 11, at 117, 143-45, 185-87.

17. See, e.g., Rawls, *supra* note 7, at 4 ("The subject of punishment . . . has always been a troubling moral question." (citation omitted)); A. John Simmons et al., *Introduction, in Punishment: A Philosophy and Public Affairs Reader* i, vii (A. John Simmons et al., eds., 1995) ("The problem of justifying legal punishment has remained at the heart of legal and social philosophy from the very earliest recorded philosophical texts to the most recent.").

18. Consequentialism assesses the value of an action or course of conduct by its consequences. Berys Gaut, *Consequentialism, in The Cambridge Dictionary of Philosophy* 176, 176 (Robert Audi ed., 2d ed. 1999). A consequentialist theory justifies punishment by the good consequences generated from punishment. R.A. Duff, *Penal Communications: Recent Work in the Philosophy of Punishment*, 20 Crime & Just. 1, 5-6 (1996) ("The common feature of all 'consequentialist accounts' . . . is that they justify punishment in terms of its *contingent* or instrumental contribution to an *independently identifiable* good."). For further discussion of the consequentialist justification of punishment, see *infra* Part I.B.1.

19. The two main justifications of punishment are retributivism and

punishment is invariably imposed subsequent to a trial and conviction.<sup>20</sup> Retributivists implicitly assume that the incompatibility of bargain justice and retributivism warrants rejection of the former.<sup>21</sup> But “[r]ather than dismiss such a significant part of our actual practice [bargain justice] because it violates the retributive ideal, we might wonder whether the retributivist has got that ideal right.”<sup>22</sup> With bargaining so entrenched,<sup>23</sup> it is time to shift the focus from whether bargain justice is justified by theories of punishment to whether theories of punishment are justified as applied to bargain justice.<sup>24</sup>

---

consequentialism. See, e.g., C.L. Ten, *Crime and Punishment*, in *A Companion to Ethics* 366, 366 (Peter Singer ed., 1991) (retributivism and consequentialism are the “two main types of theories of punishment”); Primoratz, *supra* note 7, at 9-13 (same); Andrew Oldenquist, *Retributive Rationale*, in *The Philosophy of Law* 749, 749 (Christopher Berry Gray ed., 1999) (same); A.M. Quinton, *On Punishment*, in *Philosophical Perspectives on Punishment* 6, 6 (Gertrude Ezorsky ed., 1972) (retributivism and consequentialism “exhaust the possibilities” for justifying punishment).

20. For example, in Michael Moore’s impressive 800-plus-page book, *Placing Blame*, devoted to establishing retributivism as the animating principle of criminal law, Moore allows all of three paragraphs to bargain justice, more specifically immunity deals. See Michael Moore, *Placing Blame: A Theory of Criminal Law* 156-58 (1997). And even this is more than most punishment theorists have managed to muster (to Moore’s credit, he is one of the few to attempt to reconcile retributivism with bargain justice). Perhaps the United Kingdom’s foremost punishment theorist, R.A. Duff, dispensed with plea bargains in one page in his justly acclaimed, and aptly named, book, *Trials and Punishments*. See Duff, *supra* note 6, at 141. Other fine books on the theory of punishment find literally nothing to say of bargain justice. See, e.g., Primoratz, *supra* note 7; C.L. Ten, *Crime, Guilt, and Punishment: A Philosophical Introduction* (1987). Out of the forty-five excerpts presented in Gertrude Ezorsky’s excellent anthology, *Philosophical Perspectives on Punishment*, none address the issue of bargain justice. See *supra* note 19. Arguably the most important book on the theory of criminal law since its publication in 1978, George Fletcher’s *Rethinking Criminal Law*, in over 800 pages, fails to even mention bargain justice. See Fletcher, *supra* note 7.

21. Typical of this view is George Fletcher’s comment that institutions of bargain justice, “taken for granted in the United States, stand in considerable tension with basic principles of equal justice [including retributivism].” George P. Fletcher, *The Place of Victims in the Theory of Retribution*, 3 *Buff. Crim. L. Rev.* 51, 59 (1999). The implicit conclusion is that, as a result of the tension, bargain justice, and not retributivism, is illegitimate.

22. Tunick, *supra* note 11, at 143.

23. Markus Dirk Dubber, *American Plea Bargains, German Lay Judges, and the Crisis of Criminal Procedure*, 49 *Stan. L. Rev.* 547, 551-52 (1997) (“Plea bargaining has long ago replaced the trial as the main process by which the state imposes punishment.” (citation omitted)); George Fisher, *Plea Bargaining’s Triumph*, 109 *Yale L.J.* 857, 860 (2000) (“[P]lea bargaining has so fast a grip on our institutions of justice that antagonistic institutions cannot survive.”).

24. While discussion of bargain justice is virtually nil in the debate over justifying punishment, see *supra* note 20 and accompanying text, even general discussions of bargain justice suffer from a disproportional inattention paid to it by scholars. Stephanos Bibas speaks of legal academia’s “obsession” and preoccupation with trials at the expense of guilty pleas. Bibas, *supra* note 12, at 1183 (decrying the reigning orthodoxy’s “anachronistic obsession with jury trials at the expense of pleas”). Bibas reports that though there are 24 times the number of guilty pleas than jury trials, there are only 1/10 the number of articles on guilty pleas as on jury trials. *Id.* at 1149



This Article assesses retributivism by its application to bargain justice—specifically a prisoner-initiated bargain for immunity—and concludes that retributivism is internally inconsistent.

While the difficulties prosecutor-initiated bargaining poses for prisoners in a consequentialist system of criminal justice are well known under the rubric of the “prisoner’s dilemma,”<sup>25</sup> this Article

n.328. As a result, “the literature is off by a factor of 240.” *Id.*

25. The term “prisoner’s dilemma” is generally attributed to the Princeton Mathematics Professor, and RAND Corporation consultant, Albert Tucker, in a lecture to the Stanford Psychology Department in May of 1950. Tucker had coined the term “prisoner’s dilemma” to describe the dilemma in a game devised by two other RAND consultants, Merrill Flood and Melvin Dresher. The game or dilemma posed a challenge to game theory as developed by John von Neumann, Oskar Morgenstern and John Nash. For a fascinating account of the origins of the prisoner’s dilemma and game theory, see William Poundstone, *Prisoner’s Dilemma: John von Neumann, Game Theory, and the Puzzle of the Bomb* 8, 116-19 (1992).

In the canonical presentation of the prisoner’s dilemma, a prosecutor separately makes the following offer to each of two members of a criminal gang who have been arrested and are in custody for having jointly committed a criminal offense: whichever one of the prisoners confesses and inculcates the other will receive immunity. The catch, however, is that if both confess, then not only will neither receive immunity but both will receive greater punishment than if neither had confessed. The prosecutor’s offer is made under the following circumstances: (i) the prisoners are separated, (ii) the prisoners cannot communicate with each other, (iii) neither prisoner is aware of the other’s choice and, (iv) each prisoner is deciding a course of action based only on generating the optimal individual outcome—minimizing the length of sentence that he will have to serve—but is neither concerned with generating the optimal aggregate outcome nor in the morality of betraying his confederate. See Douglas G. Baird et al., *Game Theory and the Law* 33 (1994); Poundstone, *supra*, at 118-20; Lawrence C. Becker, *Prisoner’s Dilemma*, in *The Cambridge Dictionary of Philosophy* 740, 740 (Robert Audi ed., 2d ed. 1999).

Different punishments are meted out depending on whether only one, both or neither confesses. If one confesses (i.e., supplies inculpatory information as to their crime) and the other does not, the confessor will get immunity from prosecution and the hold-out will get a severe sentence. If both confess, each will get a moderate sentence. If neither confesses, each will get light sentences. See Baird, *supra*, at 33; Poundstone, *supra*, at 117-18; Becker, *supra*, at 740. For a ranking of the four individual and three aggregate outcomes, see *infra* note 37.

Rationality seems to dictate that each prisoner should confess to minimize the amount of punishment received. This is because each prisoner should reason that the other will either confess or not confess. If one confesses, then the other should also confess (so as to get the moderate rather than the severe sentence). If the other does not confess, then he should still confess (so as to get immunity rather than a moderate sentence). Thus, no matter what the other does, each prisoner should rationally decide to confess. Yet this generates the aggregate worst outcome and the second worst individual outcome—each gets a moderate sentence. See Poundstone, *supra*, at 118, 120-22; Becker, *supra*, at 740.

As a result, the dilemma, in part, is that there is a basis for each prisoner to accept the offer (because it is purportedly rational to do so) and a basis for each prisoner to reject the offer (so as to avoid the bad outcome of both confessing). The purportedly rational choice of each confessing not only leads to the worst aggregate outcome but also the second worst individual outcome (only not confessing, while the other confesses, is a worse individual outcome). At this point, each rational prisoner might begin to question the rationality of the decision to confess. Since each realizes that both of them confessing is a worse outcome than neither confessing (both

introduces the difficulties prisoner-initiated bargaining poses for prosecutors in a retributivist system of justice.<sup>26</sup> Applying retributivism's central principles to a hypothetical immunity offer—which a prosecutor paradoxically both cannot refuse and must refuse—generates the Prosecutor's Dilemma. Suppose that a prisoner offers to betray her confederates in crime (each has jointly committed the same crime with the same level of responsibility and thus each deserves the same punishment<sup>27</sup>) by supplying inculpatory information and testifying against them in exchange for immunity.<sup>28</sup> Without the inculpatory information and testimony, the prisoner's confederates can neither be prosecuted nor punished. Thus, if the offer is accepted the confederates can be punished, but the prisoner cannot. If the offer is not accepted, the prisoner can be punished, but her confederates cannot.

Retributivism imposes an absolute duty to punish culpable wrongdoers whenever the opportunity arises.<sup>29</sup> Since the prisoner's offer supplies the opportunity to punish the confederates, retributivism's duty to punish requires acceptance of the offer and punishment of the confederates. But by accepting the offer, the prosecutor foregoes the opportunity to punish the prisoner, which violates the duty to punish. Thus, the offer must not be accepted. But nonacceptance also violates retributivism's duty to punish with respect to the confederates. Paradoxically, then, a retributivist prosecutor must both accept and not accept the offer. Retributivism's duty to

---

collectively and individually), each might realize that perhaps he should not confess. But, of course, each would subsequently reason that if the other does not confess, then he should confess (in order to get immunity) which only reintroduces the above cycle of reasoning. Thus, it seems that there is some support for each prisoner to accept the offer as well as some basis for each prisoner to reject the offer. See Poundstone, *supra*, at 122 (“[I]t is as hard to justify cooperation [neither confessing] as it is to accept mutual defection [both confessing] as the logical outcome. Therein lies the paradox.”). More conventionally, the dilemma is described as residing in the fact that what seems to be the rational choice from an individual perspective leads to the worst aggregate outcome for the prisoners. See Baird, *supra*, at 312-13.

26. The prisoner's dilemma, as it involves bargaining for the amount of punishment the prisoner will receive, cannot arise within a retributivist system of criminal justice which condemns such bargaining. See *infra* notes 35, 38 and accompanying text. The prisoner's dilemma does arise in a consequentialist system of criminal justice which does justify bargaining. See *infra* notes 35-37 and accompanying text.

27. For how retributivism determines a culpable wrongdoer's desert and degree of punishment, see *infra* Part II.B.2 and note 236.

28. For the full presentation of this hypothetical prisoner-initiated immunity proposal, termed the Prosecutor's Dilemma, see *infra* Part II.A.

29. E.g., Kant, *supra* note 8, at 141 (“The Principle of punishment is a categorical imperative, and woe to him who crawls through the windings of eudaemonism in order to discover something that releases the criminal from punishment . . .”); Moore, *supra* note 20, at 156 (“[T]he retributivist regards the punishment of the guilty to be categorically imperative whenever the opportunity to give such punishment presents itself.”). For a more expansive discussion of retributivism's absolute duty to punish culpable wrongdoers, see *infra* Part II.B.1.

punish at once requires, and is violated by, both acceptance and nonacceptance of the offer. The difficulty for the prosecutor is that retributivism's duty to punish requires punishment of both the prisoner and her confederates, but the terms of the Prosecutor's Dilemma allow punishment of either only the prisoner or only her confederates.

Consideration of retributivism's equality principle—culpable wrongdoers with the same desert must be treated equally<sup>30</sup>—compounds the dilemma. Since the terms of the Prosecutor's Dilemma preclude punishment of both the prisoner and her confederates, then the only way to treat the prisoner and her confederates equally is to punish neither the prisoner nor her confederates.<sup>31</sup> As applied to the Prosecutor's Dilemma, retributivism's duty to punish—requiring at the very least that someone be punished—conflicts with retributivism's duty of equal treatment which requires that neither the prisoner nor her confederates be punished. Applying retributivism's principles to the prisoner's offer obligates a prosecutor to satisfy the following conjunctively: (i) accept the offer (thus punishing only the confederates), (ii) not accept the offer (and punish only the prisoner), and (iii) not accept the offer (and punish neither the prisoner nor her confederates). But these options can only be satisfied disjunctively. The dilemma is that retributivism's conflicting principles obligate a prosecutor to do all three when only one is possible to attain, and that

---

30. *E.g.*, C.L. Ten, *supra* note 20, at 51 (discussing retributivism's "requirement that similar cases should be treated similarly"); Fletcher, *supra* note 21, at 59 ("In view of the fact that all theories of justice are primarily concerned with equality, it makes sense to ground retributive justice as well in a commitment to bring about equality . . . among [equally culpable] offenders." (citation omitted)). For further discussion of retributivism's duty to treat equally those culpable wrongdoers with equal desert, see *infra* Part II.B.5.

31. To satisfy retributivism's equality principle, culpable offenders of equal desert must be treated equally. Since the prisoner and her confederates committed the same crime with the same *mens rea*, the prisoner and her confederates have equal desert. See *infra* notes 235-36 and accompanying text. As a result, the equality principle requires that the prisoner and her confederates be treated equally. Generally, this can be satisfied in two ways: (i) punish both, and punish them equally, or (ii) punish neither so that each equally receives no punishment. Under the terms of the Prosecutor's Dilemma, the first possibility is foreclosed—the prisoner and her confederates cannot both be punished. The only options under the Prosecutor's Dilemma are (i) accept the offer and punish only the confederates, or (ii) not accept the offer and punish only the prisoner, or (iii) not accept the offer and punish neither the prisoner nor her confederates. The first two options violate the equality principle. Punishing only the confederates treats the prisoner unequally with respect to the confederates, and punishing only the prisoner treats the confederates unequally with respect to the prisoner. The only option which is both available under the terms of the Prosecutor's Dilemma and which satisfies retributivism's equality principle is to not accept the offer and punish neither the prisoner nor her confederates.

any one option a prosecutor adopts will violate retributivism's central principles.<sup>32</sup>

Given retributivists' condemnation of bargain justice, it might seem that a retributivist prosecutor can avoid the Prosecutor's Dilemma by simply refusing to participate in such bargaining. But while a retributivist prosecutor can (and must) refrain from initiating bargaining,<sup>33</sup> the tendering of prisoner-initiated bargain justice offers cannot be prevented. Once the offer is made, it must be accepted or not. And nonacceptance of the offer violates retributivism's central principles no less than does acceptance of the offer.<sup>34</sup>

In comparison, the Prosecutor's Dilemma does, but the prisoner's dilemma<sup>35</sup> does not, undermine the system of punishment under which it arises. The prisoner's dilemma arises only under a consequentialist system because consequentialism, and not retributivism, justifies bargain justice.<sup>36</sup> Not only does the prisoner's dilemma not discredit consequentialism, it may serve to enhance the efficacy of a consequentialist system.<sup>37</sup> The Prosecutor's Dilemma, however,

---

32. For a discussion of other central principles of retributivism which are violated regardless of which option a prosecutor selects, see *infra* Part II.B.

33. Since retributivism condemns bargain justice, see *infra* Parts I.B.2 and II.B, a retributivist prosecutor must forego initiating bargain justice offers.

34. For example, the option of not accepting the offer (and punishing only the prisoner) violates retributivism's duty to punish the confederates as well as the duty to treat equally the prisoner and the confederates.

35. For further discussion of the prisoner's dilemma, see Baird, *supra* note 25, at 33, 48, 312-13; Neil Duxbury, *Random Justice: On Lotteries and Legal Decision-Making* 112-14 (1999); Robin Paul Malloy, *Law and Market Economy* 12-13 (2000); Becker, *supra* note 25, at 740; Ronald J. Gilson & Robert H. Mnookin, *Disputing Through Agents: Cooperation and Conflict Between Lawyers in Litigation*, 94 *Colum. L. Rev.* 509, 514-22 (1994); Peter Huber, *Competition, Conglomerates, and the Evolution of Cooperation*, 93 *Yale L.J.* 1147, 1150-58, 1167-68 (1984); John Shepard Wiley, Jr., *Reciprocal Altruism as a Felony: Antitrust and the Prisoner's Dilemma*, 86 *Mich. L. Rev.* 1906, 1914-20 (1988); J. Bradford DeLong, *Prisoner's Dilemma*, at [http://www.j-bradford-delong.net/economists/prisoners\\_dilemma.html](http://www.j-bradford-delong.net/economists/prisoners_dilemma.html) (last visited March 7, 2003).

36. See Tunick, *supra* note 9, at 140 ("That plea-bargaining allows tremendous savings in resources is a strong 'utilitarian' argument in its favor. By one estimate, if we took all cases to trial we would have to quadruple the number of court personnel. But there are compelling retributivist objections to plea-bargaining."); Ralph D. Ellis, *Fairness and the Etiology of Criminal Behavior*, 13 *Phil. and Soc. Criticism* 175, 181 (1987) ("[P]lea bargaining is theoretically more compatible with deterrence than with retribution.").

37. The prisoner's dilemma, whether understood explicitly or only implicitly by actual prisoners, may lead to more favorable plea negotiations for the prosecutor. The more favorable the plea negotiations for the prosecutor, the more favorable the plea bargain, which may lead to a greater number of criminals being punished for a greater number of offenses, thereby enhancing some of the deterrent and incapacitative goals of punishment under consequentialism. To the extent that prisoners and prosecutors (i.e., society) are in a zero sum game in which the worse the aggregate outcome for the prisoners the better the consequential benefits of crime prevention for society, the prisoner's dilemma promotes consequentialist goals by inducing prisoners to opt for their worst aggregate outcome. See *supra* note 25. The

arising only under a retributivist system,<sup>38</sup> renders retributivism internally inconsistent.

Bargain justice weakens retributivism as a theory of punishment in both its descriptive and prescriptive capacities. By being incompatible with plea bargaining, retributivism risks irrelevance as a theory that can describe and explain the mode through which actual legal institutions mete out the vast majority of punishments.<sup>39</sup> And prisoner-initiated immunity bargaining, in the form of the Prosecutor's Dilemma, reveals the just conditions of punishment prescribed by retributivism to be internally inconsistent.

As a result, the Prosecutor's Dilemma supplies an indirect defense

prisoners' worst aggregate outcome is thus the best collective outcome for society. The four outcomes for an individual prisoner, ranked from best to worst (from the perspective of the prisoner) are as follows:

1. You (as a prisoner) confess, and the other prisoner does not: you get immunity (0 years imprisonment).
2. Neither you nor the other prisoner confesses: you get a light sentence (e.g., 1 year).
3. Both you and the other prisoner confess: you get a moderate sentence (e.g., 2 years).
4. You do not confess, but the other prisoner does confess: you get a heavy sentence (e.g., 3 years).

The three outcomes for the aggregate of prisoners, ranked from best to worst, are as follows:

1. Neither confesses: each gets a light sentence (1+1=2 years).
2. One confesses and the other does not: the confessor gets immunity and the hold-out gets a heavy sentence (0+3=3 years).
3. Both confess: each gets a moderate sentence (2+2=4 years).

Poundstone, *supra* note 25, at 118-21; Becker, *supra* note 25, at 740. Since the prisoner's dilemma induces each prisoner to confess, which is the worst aggregate outcome for the prisoners, it leads to the best outcome for society. This argument is, of course, premised on the assumption that bargain justice does, in fact, advance consequentialist goals. For a discussion of the consequentialist justification of bargain justice, see *infra* Part I.B.1.

With respect to the ranking of the aggregate outcomes, the second outcome is better than the third and worse than the first only based on arbitrarily assigned terms of imprisonment assigned to each category of sentence. Altering the arbitrarily assigned numbers could variously yield one getting immunity and one getting a heavy sentence as either the best aggregate outcome or the worst. The only necessary, nonarbitrary ranking of the aggregate outcomes is that neither confessing and each getting a light sentence is preferable, in the aggregate, to both confessing and each getting a moderate sentence.

38. The Prosecutor's Dilemma arises only under retributivism because retributivism imposes an absolute duty to punish culpable wrongdoers and requires the equal treatment of equally culpable offenders. Lacking these principles, consequentialism supplies a clear answer to the offer: accept. See *infra* text accompanying notes 178-79. In contrast, the prisoner's dilemma arises under consequentialism because of its support and justification for bargain justice, see *infra* Part I.B.1, but would not arise under retributivism because of its condemnation of bargain justice.

39. For a discussion of retributivism's rejection of bargain justice, see *infra* Parts I.B.2 and II.B. For a discussion of the prevalence of punishments imposed subsequent to some form of bargain justice, as opposed to a trial, see *supra* note 11 and accompanying text.

of bargain justice. Ironically, that which retributivism rejects as illegitimate—bargain justice—is that which reveals retributivism as unsound. Retributivism's incompatibility with bargain justice is no longer an indictment of bargain justice, but rather is an indictment of retributivism.

After Part I sketches a brief overview of bargain justice and theories of punishment, Part II applies consequentialism and retributivism to a hypothetical prisoner-initiated offer of immunity—the Prosecutor's Dilemma. While consequentialism supplies a clear basis to accept the offer, retributivism's response is indeterminate. An analysis of four central principles of retributivism reveals that each principle paradoxically requires, and is violated by, both acceptance and nonacceptance of the offer. An additional retributivist principle—the equality principle—conflicts with the other four retributivist principles by requiring nonacceptance of the offer (and punishing neither the prisoner nor her confederates).

Part III analyzes retributivism's indeterminate response to the offer. The indeterminacy stems not only from inconsistency within, but also among, retributivism's principles. In turn, retributivism's internal inconsistency generates a moral dilemma in which any option a prosecutor selects will violate retributivism's central principles. After situating the Prosecutor's Dilemma within the typology of moral dilemmas and assessing its comparative gravity, this part examines two attempts to avoid the Prosecutor's Dilemma.

Part IV considers two alternative versions of retributivism which attempt to forge a reconciliation with bargain justice. While their embrace of consequences is conceptually questionable, these radical forms of retributivism to a varying degree do make retributivism more compatible with bargain justice. But they neither restore retributivism to consistency nor satisfactorily avoid the Prosecutor's Dilemma.

Part V canvasses possible means of restoring retributivism to consistency and resolving the Prosecutor's Dilemma. In accordance with the method of dilemma resolution preferred by retributivists themselves, the theoretically sound resolution is not accepting the offer and punishing neither the prisoner nor her confederates. In a practical sense, however, this is disastrous in that it allows prisoners to, in effect, self-immunize by the mere tender of a dilemmatic bargain justice offer. While the resolution is a practical disaster, without a resolution of the Prosecutor's Dilemma retributivism remains internally inconsistent. Either way—with or without the resolution—the incompatibility of retributivism and bargain justice now renders illegitimate not bargain justice, but rather retributivism.

## I. BARGAIN JUSTICE AND THE THEORY OF PUNISHMENT

This part supplies a brief overview of the two principal forms of bargain justice: plea bargains and immunity deals. After explicating the consequentialist<sup>40</sup> and retributivist<sup>41</sup> accounts of punishment, this part presents their justification and critique, respectively, of bargain justice.

## A. Bargain Justice

Rather than determining what the defendant justly deserves by a trial, the prosecutor and defendant may bargain, either formally or informally, as to the terms of what each shall receive from the other.<sup>42</sup> While “offers of . . . immunity are . . . a form of plea bargaining,”<sup>43</sup> they are also distinguishable. If the plea bargain<sup>44</sup> represents “condemnation without adjudication,”<sup>45</sup> the immunity deal constitutes freedom from both condemnation and adjudication.<sup>46</sup>

40. Since the focus of this Article will be the difficulties sustained by retributivism as applied to bargain justice, the account of consequentialism will be comparatively less extensive.

41. For an examination of the principles of retributivism in greater depth, see *infra* Part II.B.

42. This conception of bargain justice excludes pardons because they typically do not entail a *quid pro quo* exchange of things of value between the offender and the state. For an illuminating discussion of pardons in connection with the purposes of punishment, with particular emphasis on retributivism, see Kathleen Dean Moore, *Pardons: Justice, Mercy, and the Public Interest* (1989). For a defense of retributivism’s compatibility with various forms of mercy, see Dan Markel, *Against Mercy*, 88 *Minn. L. Rev.* (forthcoming 2004). For a discussion of the role of presidential pardons in the development of statutory immunity, see *infra* note 70 and accompanying text.

43. Standen, *supra* note 15, at 1474 n.5.

44. Albert Alschuler notes that informal accounts of plea bargaining and the preferable conception diverge:

The term *plea bargaining* is sometimes used informally to refer to every form of discussion between the prosecution and the defense in a criminal case that might lead to disposition of the case without trial or to any sort of concession on the part of the defendant. . . . [B]ut *plea bargaining* probably can best be defined as the exchange of any actual or apparent concession for a plea of guilty.

Albert Alschuler, *Plea Bargaining*, in *Encyclopedia of Crime and Justice* 829, 829 (Sanford Kadish ed., 1983). For more expansive definitions of plea bargaining, see Douglas D. Guidorizzi, *Should We Really “Ban” Plea Bargaining?: The Core Concerns of Plea Bargaining Critics*, 47 *Emory L.J.* 753, 755-57 (1998); Schulhofer, *supra* note 15, at 1037; Robert E. Scott & William J. Stuntz, *Plea Bargaining as Contract*, 101 *Yale L.J.* 1909, 1909 (1992); Standen, *supra* note 15, at 1474 n.5.

45. John H. Langbein, *Land Without Plea Bargaining: How the Germans Do It*, 78 *Mich. L. Rev.* 204, 204 (1979).

46. For a discussion of the varying degrees of protection from prosecution that a purported offender receives in an immunity deal, see *infra* notes 78-88 and accompanying text.

## 1. Plea Bargain

In perhaps the first reported domestic case addressing plea bargaining,<sup>47</sup> the court in *Commonwealth v. Battis*, in 1804, reluctantly allowed a guilty plea to a capital crime, but not until the court satisfied itself that “there had not been tampering with him, either by promises . . . [or] persuasions . . . if he would plead guilty.”<sup>48</sup> Since at least 1970,<sup>49</sup> however, with the Supreme Court’s first endorsement of plea bargaining,<sup>50</sup> prosecutors and defendants openly exchange both promises and persuasions. To obtain an accused’s guilty plea, a prosecutor offers two principal forms of consideration: sentence bargaining and charge bargaining.<sup>51</sup> In sentence bargaining, a

---

47. Though plea bargaining may have its roots in long-entrenched practices of rewarding cooperation by defendants dating back as far as 1587 in England, Fisher, *supra* note 23, at 860 n.1, plea bargaining was mostly discouraged in the United States prior to the Civil War. Alschuler, *supra* note 1, at 16-18.

48. *Commonwealth v. Battis*, 1 Mass. (1 Will.) 94, 95 (1804). In addition, the court inquired as to the sanity of the prisoner. *Id.*

49. After the Civil War, the incidence of plea bargaining began to increase, Alschuler, *supra* note 1, at 19-26, until the 1920s when it became relatively commonplace. *Id.* at 26-33. But sustained support for plea bargaining did not commence until the late 1960s and early 1970s. Albert W. Alschuler, *Plea Bargaining and Its History*, 13 Law & Soc’y Rev. 211, 236-38 (1979) (noting that in 1967, both the American Bar Association Project on Minimum Standards for Criminal Justice as well as the President’s Commission on Law Enforcement and Administration of Justice gave their blessings to plea bargaining). In 1973, however, “[p]ublic disenchantment with the practice of plea bargaining probably reached its peak.” Samuel Walker, *Taming the System: The Control of Discretion in Criminal Justice 1950-1990*, at 93 (1993). Former Vice President Spiro Agnew entered a *nolo contendere* plea to a lesser charge of federal tax evasion in exchange for federal prosecutors’ promise not to prosecute him on the more serious charges of bribery and extortion. Bart Barnes, *Nixon Vice President Spiro T. Agnew Dies*, Wash. Post, Sept. 19, 1996, at B7; Tom Stuckey, *20 Years Later, A Life of Obscurity for Spiro Agnew*, Wash. Post, Oct. 14, 1993, at Md.7. And the National Advisory Commission on Criminal Justice Standards and Goals recommended the abolishment of plea bargaining, in all its forms, no later than 1978. See Jeff Palmer, Note, *Abolishing Plea Bargaining: An End to the Same Old Song and Dance*, 26 Am. J. Crim. L. 505, 511 (1999).

50. *Brady v. United States*, 397 U.S. 742, 753 (1970) (upholding the constitutionality of a guilty plea under a statute permitting the death penalty only after a jury trial, in part, because plea bargains benefit both the prosecution and the defendant). Later that same year, in *North Carolina v. Alford*, the Court upheld a guilty plea, as part of a plea bargain, entered to avoid the possible imposition of the death penalty—notwithstanding the defendant’s claim of innocence. 400 U.S. 25, 27-29, 39 (1970).

51. Alschuler, *supra* note 44, at 829 (reducing the number or severity of the charges and reducing the sentence). A plea bargain might consist of both charge and sentencing bargaining. See, e.g., Eleanor Randolph, *A Contrite Wachtler Apologizes: Former N.Y. Judge Faces Prison Term*, Wash. Post, Apr. 1, 1993, at A3. New York State’s then-highest ranking judge, Sol Wachtler, agreed to a plea bargain in which, in exchange for his pleading guilty to a single lesser charge of making threats against his ex-lover and her daughter, prosecutors agreed to drop a five-count extortion indictment. *Id.* In addition, the prosecutor agreed to sentencing concessions. Although Wachtler faced a maximum of five years and a \$250,000 fine, the prosecutor



prosecutor promises to recommend a lenient sentence or not oppose the defendant's sentencing recommendation.<sup>52</sup> In charge bargaining, a prosecutor promises to reduce the severity of the crime charged,<sup>53</sup> reduce the number of crimes or counts charged,<sup>54</sup> or both.<sup>55</sup> Other forms of consideration include promising to "provide leniency to a defendant's accomplices,"<sup>56</sup> not prosecute a defendant's friend or relative,<sup>57</sup> recommend (or not oppose) early release from imprisonment through probation<sup>58</sup> or parole,<sup>59</sup> recommend (or not oppose) the conditions of the offender's punishment,<sup>60</sup> or to

---

recommended twelve to eighteen months. *Id.*

52. See, e.g., *United States v. Singleton*, 144 F.3d 1343, 1344 (10th Cir. 1998), *rev'd en banc*, 165 F.3d 1297 (10th Cir. 1999) (to secure an accomplice's testimony against a co-accomplice, the prosecutor promised to inform the sentencing court of the nature and extent of the cooperating accomplice's valuable assistance which presumably might influence the court to lessen his sentence); *United States v. Dailey*, 759 F.2d 192, 194 (1st Cir. 1985) (in exchange for an offender's guilty plea and testimony against a co-conspirator, the prosecutor agreed to recommend a lenient sentence of twenty years; and if the offender's testimony was particularly valuable, the prosecutor agreed to recommend a sentence of only ten years).

53. See, e.g., *United States v. McVeigh*, 153 F.3d 1166 (10th Cir. 1998) (allowing a cooperating witness to plead guilty to a lesser charge); *United States v. Fallon*, 776 F.2d 727, 729 (7th Cir. 1985) (as part of the plea bargain, the prosecutor agreed, in part, to reduced charges).

54. See, e.g., *United States v. Waterman*, 704 F.2d 1014, 1016 (8th Cir. 1983) (to obtain the offender's guilty plea to three charges in connection with an arson for profit scheme, the government agreed to drop fifteen out of the eighteen charges for which the offender was indicted).

55. Michael Milken, the junk bond titan of the 1980s, agreed to a plea bargain with federal prosecutors and the Securities and Exchange Commission in which, in partial exchange for his guilty plea to six felonies (including tax, securities and mail fraud), prosecutors agreed, in part, to drop a 98-count indictment which included the comparatively more serious charges of racketeering. See Laurie P. Cohen, *Plea Bargains Must Balance Rewards v. Risks in Settlements Such as Milken's*, Wall St. J., Apr. 24, 1990, at B1; Laurie P. Cohen, *Public Confession: Milken Pleads Guilty to Six Felony Counts and Issues an Apology*, Wall St. J., Apr. 25, 1990, at A1 [hereinafter Cohen, *Public Confession*].

For an interesting debate on the merits of charge bargaining, see Ronald Wright & Marc Miller, *The Screening/Bargaining Tradeoff*, 55 Stan. L. Rev. 29 (2002); Gerard E. Lynch, *Screening Versus Plea Bargaining: Exactly What Are We Trading Off?*, 55 Stan. L. Rev. 1399 (2003) (replying to Wright & Miller, *supra*); Ronald Wright & Marc Miller, *Honesty and Opacity in Charge Bargains*, 55 Stan. L. Rev. 1409 (2003) (replying to Lynch, *supra*).

56. Alschuler, *supra* note 44, at 829.

57. See, e.g., Cohen, *Public Confession*, *supra* note 55, at A6 (as part of Michael Milken's plea bargain, prosecutors agreed not to seek criminal charges against his brother, Lowell Milken, who was a co-defendant in the indictment).

58. See, e.g., Barnes, *supra* note 49, at B7 (in exchange for a guilty plea, the prosecutor agreed to a sentence of three years unsupervised probation for former Vice President Spiro Agnew).

59. See, e.g., *United States v. Singleton*, 144 F.3d 1343, 1344 (10th Cir. 1998), *rev'd en banc*, 165 F.3d 1297 (10th Cir. 1999) (to secure an accomplice's testimony against a co-accomplice, the prosecutor promised to inform the parole board of the nature of the cooperating accomplice's valuable assistance which presumably might lead to an earlier grant of parole).

60. See, e.g., *United States v. Fallon*, 776 F.2d 727, 729 (7th Cir. 1995) (in exchange

recommend that credit be given for time served while awaiting trial.<sup>61</sup> In addition to pleading guilty, an offender may offer further consideration including testifying against accomplices,<sup>62</sup> cooperating with other ongoing or future government investigations,<sup>63</sup> arranging for the offender's father to testify against an accomplice,<sup>64</sup> making restitution to victims,<sup>65</sup> or relinquishing a professional license that facilitated the crime.<sup>66</sup>

## 2. Immunity Deal

The early American roots of inducing reluctant witnesses to testify,<sup>67</sup> despite their assertion of the Fifth Amendment privilege against self-incrimination,<sup>68</sup> lie not in statutory grants of immunity,<sup>69</sup> but in presidential pardons.<sup>70</sup> Today, a federal prosecutor's authority

---

for pleading guilty and supplying testimony against a co-defendant, the prosecutor agreed to recommend that incarceration should not be part of the offender's punishment). In addition, a prosecutor might promise to "arrange for a defendant to be sent to a particular correctional institution." Alschuler, *supra* note 44, at 829.

61. Alschuler, *supra* note 44, at 829. Still other forms of consideration might include the promise to:

withhold damaging information from the court, influence the date of a defendant's sentencing, . . . attempt to have charges in other jurisdictions dismissed, arrange for sentencing in a particular court or by a particular judge, provide immunity for crimes not yet charged, or simply remain silent when his recommendation otherwise might be unfavorable.

*Id.* at 829-30.

62. See *supra* note 53.

63. See, e.g., Cohen, *Public Confession*, *supra* note 55, at A1 (as part of Michael Milken's plea bargain, Milken agreed to be available to federal investigators and to "fully and truthfully disclose all information with respect to Drexel-related activities [i.e., activities related to his investment firm]").

64. In return for convicted spy John Walker's testimony against his accomplice, Jerry Whitworth, the potential sentence of life imprisonment for Walker's son was reduced to 25 years. See Stephen S. Trott, *Words of Warning for Prosecutors Using Criminals as Witnesses*, 47 *Hastings L.J.* 1381, 1426 (1996) (citing Ruth Marcus, "Israeli Defense" Rejected by Whitworth Spy Jury, *Wash. Post*, July 26, 1986, at A6.).

65. As part of Michael Milken's plea bargain, Milken agreed to pay \$400 million into a victims restitution fund administered by the Securities and Exchange Commission. Cohen, *Public Confession*, *supra* note 55, at A1.

66. See, e.g., *id.* (as part of Michael Milken's settlement with the Securities and Exchange Commission, Milken, and his brother Lowell Milken, are barred from working in the securities business for life).

67. For a discussion of immunity deals in Italy dating as far back as 1630, see *supra* note 1.

68. The Fifth Amendment provides, in relevant part, as follows: "No person shall be . . . compelled in any criminal case to be a witness against himself . . ." U.S. Const. amend. V.

69. The first congressional immunity statute was enacted in 1857. See Act of Jan. 24, 1857, ch. 19, 11 Stat. 155 (1859).

70. The Constitution explicitly grants the President the power to grant pardons: "The President shall . . . have Power to grant Reprieves and Pardons for Offences against the United States, except in Cases of Impeachment." U.S. Const. art. II, § 2. For speculation that part of the basis for granting the President the power to pardon was to compel witnesses to testify despite their invocation of the Fifth Amendment

to grant statutory immunity to a witness currently stems from the 1970 federal witness immunity statutes.<sup>71</sup> If, in a prosecutor's "judgment,"<sup>72</sup> a witness's testimony "may be necessary to the public interest"<sup>73</sup> and the witness has refused, or is likely to refuse, to provide such information based on her Fifth Amendment right against self-incrimination,<sup>74</sup> the court shall, upon the prosecutor's request,<sup>75</sup> grant immunity from prosecution to the witness and order<sup>76</sup> the witness to testify.<sup>77</sup> Broader than use immunity but narrower than transactional immunity,<sup>78</sup> the statute confers "use and derivative use immunity."<sup>79</sup>

---

right against self-incrimination, see Leonard N. Sosnov, *Separation of Powers Shell Game: The Federal Witness Immunity Act*, 73 Temp. L. Rev. 171, 176 n.21 (2000). Presidents have sought to compel the testimony of witnesses with particularly valuable information. *E.g.*, *United States v. Burr (In re Willie)*, 25 F. Cas. 38, 40 (C.C.D. Va. 1807) (No. 14,692e) (compelling witness, after receiving President Thomas Jefferson's pardon, to testify against Aaron Burr despite witness's invocation of Fifth Amendment right against self-incrimination). As a tool to compel testimony, however, the pardon's effectiveness diminished when the Supreme Court, in *Burdick v. United States*, ruled that President Woodrow Wilson's pardon offer could be refused by the recipient because its acceptance "carries an imputation of guilt." 236 U.S. 79, 94 (1915). In contrast, the other source of compelling testimony, the congressional statute of 1857 authorizing immunity, see *supra* note 69, entailed no such imputation. *Burdick*, 236 U.S. at 94-95.

71. Organized Crime Control Act of 1970, 18 U.S.C. §§ 6001-05 (2000).

72. *Id.* § 6003(b).

73. *Id.* § 6003(b)(1).

74. *Id.* § 6003(b)(2).

75. Although the prosecutor must request the court to formally grant statutory immunity and order the witness to testify, the court's role is largely that of a rubber stamp. It is the prosecutor, not the court, who controls whether immunity is granted. See *United States v. Doe*, 465 U.S. 605, 616-17 (1984) ("Congress expressly left [the decision to immunize a witness] exclusively to the Justice Department."); *Pillsbury Co. v. Conboy*, 459 U.S. 248, 253-54 (1983) (acknowledging that the authority to grant immunity lies not with the courts, but with the Department of Justice); *United States v. Stewart*, 122 F.3d 625, 627 (8th Cir. 1997) ("The district court did not have the authority to grant . . . immunity because this court has consistently refused to recognize the concept of judicial immunity."); *United States v. Lugg*, 892 F.2d 101, 104 (D.C. Cir. 1989) ("The cases are legion and uniform that only the Executive can grant statutory immunity, not a court.").

76. Even without explicit bargaining, imposed immunity may still constitute an instance of bargain justice. The offender is receiving less punishment (than she might have received if tried and convicted) because she is supplying something valuable (more valuable than her conviction and punishment), that is, her inculpatory testimony, to the prosecutor. That a bargain is forced does not preclude it from being, in some sense, a bargain.

77. 18 U.S.C. §§ 6002-03.

78. The 1857 federal immunity statute, see *supra* note 69, authorized "transactional immunity" until an amendment in 1862 implemented a narrower form of immunity, "use immunity." See, e.g., Howard R. Sklamberg, *Investigation Versus Prosecution: The Constitutional Limits on Congress's Power to Immunize Witnesses*, 78 N.C. L. Rev. 153, 158-59 (1999). In 1892, however, the Court found the narrower use immunity unconstitutional by not affording adequate protection of a witness's right against self-incrimination. See *Counselman v. Hitchcock*, 142 U.S. 547, 585-86 (1892) (rejecting use immunity and requiring transactional immunity). While in transactional immunity a witness is immunized from criminal prosecution for all the transactions or conduct discussed in the testimony, in use immunity the witness only

That is, neither information provided nor derived therefrom may be used against the witness in a criminal prosecution.<sup>80</sup> The Court, in *Kastigar v. United States*, upheld the constitutionality of use and derived use immunity<sup>81</sup> by reasoning that it leaves the witness in essentially the same position as if he had relied on the privilege against self-incrimination and not testified.<sup>82</sup>

In addition to statutory immunity, a defendant may secure a prosecutor's promise not to prosecute.<sup>83</sup> While less beneficial to offenders,<sup>84</sup> such informal immunity agreements,<sup>85</sup> or "cooperation

has immunity from her testimony being used against her but may be prosecuted for the conduct referred to in the testimony based on independent evidence. *Id.* For example, Monica Lewinsky received transactional immunity from Independent Counsel Kenneth Starr in return for agreeing to supply information regarding her relationship with President Clinton. Peter Baker & Susan Schmidt, *Lewinsky Gets Immunity for Her Testimony; Sources Say Ex-Intern to Tell of Clinton Bid to Hide Affair*, Wash. Post, July 29, 1998, at A1; Ruth Marcus, *Immunity Coverage is Extensive, Lying to Grand Jury Is Only Exception*, Wash. Post, July 29, 1998, at A8. In contrast, Linda Tripp received a type of use immunity, use and derived use immunity, see *infra* notes 79-82, 84 and accompanying text, from Starr in return for furnishing the tape-recorded conversations she had with Monica Lewinsky, but Tripp was nonetheless prosecuted in the state of Maryland. See Raja Mishra & Katherine Shaver, *Immunity for Tripp Wiretap Attacked: Judge to Issue Key Ruling Today on Use of Tapes*, Wash. Post, Dec. 14, 1999, at B1.

79. See 18 U.S.C. § 6002(3).

80. *Id.* The provision does, however, carve out a number of exceptions. The testimony or information supplied by the witness may be used against the witness in a criminal prosecution "for perjury, giving a false statement, or otherwise failing to comply with the order" compelling testimony. *Id.*

81. *Kastigar v. United States*, 406 U.S. 441, 462 (1972) (holding that use and derived use immunity adequately protected a witness's Fifth Amendment right against self-incrimination because it is broader than mere use immunity).

82. *Id.* at 458-59. Use and derived use immunity, as interpreted by *Kastigar*, may be broad enough, as a practical matter, to sufficiently protect a witness's Fifth Amendment right against self-incrimination:

As a practical matter, application of the *Kastigar* rule may make any subsequent prosecution impossible. For example, Oliver North's conviction was overturned, and the government declined to retry him because of the exceedingly exacting test that would have been required to establish an independent source apart from his immunized congressional testimony.

Adam Harris Kurland, *Lurking Pitfalls of Successive Prosecutions: Immunity, Plea Agreements, Promises Not to Prosecute, and Cooperation Agreements*, 14 Crim. Just. 4, 7 (2000) (citing *United States v. North*, 910 F.2d 843, 873-74 (D.C. Cir. 1990)).

83. *E.g.*, *United States v. Singleton*, 144 F.3d 1343, 1357-58 (10th. Cir. 1998), *rev'd en banc*, 165 F.3d 1297 (10th Cir. 1999) (to secure an accomplice's testimony against a co-accomplice for cocaine distribution and money laundering, the government promised not to prosecute the cooperating accomplice for any other violations of the Drug Abuse Prevention and Control Act arising from his activities currently under investigation).

84. Such informal agreements, whether of plea or immunity, may provide no protection from a different sovereign. This concern may have led Terry Nichols to refuse to cooperate in the Oklahoma City bombing case:

[A] jury convicted Terry Nichols on several charges. U.S. District Judge Richard Matsch told Nichols he would sentence him to life in prison—the maximum possible sentence—unless Nichols provided information and

agreements,<sup>86</sup> allow a prosecutor to “maintain [more] control . . . [of the witness than a statutory, or] ‘cold’ grant of immunity.”<sup>87</sup> Not surprisingly, “we have seen a gradual, but substantial, shift in prosecutorial tactics away from *compulsion* of testimony [via statutory

---

helped resolve several lingering questions about the bombing. Nichols rejected the offer to exchange leniency for information, largely from concern that any statements made pursuant to that agreement could be used against him in the Oklahoma state murder trial, where state authorities sought the death penalty.

Kurland, *supra* note 82, at 9-10. Testimony subsequent to a promise not to prosecute from a federal prosecutor may not shield the witness from the testimony being used in a prosecution in a state jurisdiction and vice-versa. *See, e.g.*, Taylor v. Singletary, 148 F.3d 1276, 1283 n.7 (11th Cir. 1998) (noting that testimony pursuant to federal prosecutor's informal immunity agreement does not constitute compelled testimony, and thus witness cannot claim that a state used the testimony as the basis of state prosecution against the witness); United States v. Roberson, 872 F.2d 597, 611 (5th Cir. 1989) (state prosecutor's informal immunity agreement promising not to prosecute cannot bind federal prosecutor without consent); State v. Barone, 689 A.2d 132, 139 (N.J. 1997) (federal prosecutor's informal immunity agreement “cannot bind state prosecutors and vice-versa”).

Only formal statutory immunity may prevent a successive prosecution by a different sovereign utilizing the testimony against the witness. Kurland, *supra* note 82, at 7. For example, Linda Tripp entered into an informal immunity agreement with Independent Counsel Kenneth Starr, in return for Tripp turning over her tape-recorded conversations with Monica Lewinsky. Five weeks later, Tripp received statutory immunity. Despite this federal immunity, Tripp was prosecuted in the state of Maryland on two charges: (i) recording a conversation with Lewinsky without her knowledge, and (ii) disclosing its contents to a magazine. A crucial issue in the Maryland prosecution was whether the tapes could be used as evidence against Tripp. The state judge ruled that the tapes were not protected by Tripp's federal immunity because they were turned over to Starr subsequent to her informal immunity agreement, but prior to her statutory immunity. *See* Raja Mishra & Katherine Shaver, *Md. Judge Denies Immunity to Tripp; Ruling Deals Blow to Wiretap Defense*, Wash. Post, Dec. 15, 1999, at A1; Mishra and Shaver, *supra* note 78, at B1. Ultimately, the charges were dismissed because key testimony by Monica Lewinsky was suppressed on the basis that it relied on the use of information protected by Tripp's statutory immunity. *See* Mary Otto, *Md. Judge Dismisses Tripp Wiretap Case*, Wash. Post, June 1, 2000, at A10; Michael E. Ruane & Mary Otto, *Md. Case Against Tripp Weakened; Lewinsky Not Allowed to Authenticate Tape*, Wash. Post, May 23, 2000, at B1.

85. Such informal immunity agreements are treated as contracts between the prosecutor and the witness. *E.g.*, United States v. Thompson, 25 F.3d 1558, 1562 (11th Cir. 1994) (informal immunity agreements are governed by “basic principles of contract law”); United States v. Biaggi, 675 F. Supp. 790, 803-04 (S.D.N.Y. 1987) (informal immunity agreements constitute contracts which are subject to contractual remedies).

86. John G. Douglass, *Confronting the Reluctant Accomplice*, 101 Colum. L. Rev. 1797, 1811 (2001).

87. *Id.* Douglass explains the advantages of an informal cooperation agreement over a formal, statutory grant of immunity: “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 [i.e., formal] grant of immunity.” *Id.* at 1826-27 (citation omitted). Another advantage is that the immunity may be conditioned on the witness fulfilling his duty to cooperate fully or “testify truthfully.” Singleton, 144 F.3d at 1344.

immunity] and toward *purchase* of testimony [via informal, bargained agreements].”<sup>88</sup>

### B. Assessing Bargain Justice via the Theories of Punishment

Punishment<sup>89</sup> requires justification<sup>90</sup> because it involves the deliberate infliction of pain, suffering, and deprivation which is *prima facie* wrong.<sup>91</sup> Although most agree that punishment is justifiable,<sup>92</sup> the form and content of that justification is contentious.<sup>93</sup>

88. Douglass, *supra* note 86, at 1811 n.48. Douglass notes that requests from federal prosecutors for statutory immunity increased every year from 1973 to 1986, when they peaked at 2,550. Thereafter, statutory immunity requests declined every year until 1996, when there were only 1,493 requests. *Id.* (citing Bureau of Justice Statistics, U.S. Dep’t of Justice, Sourcebook of Criminal Justice Statistics 418, tbl.5.1 (1996)).

89. H.L.A. Hart supplies perhaps the most influential conceptual definition of punishment:

(i) It must involve pain or other consequences normally considered unpleasant.

(ii) It must be for an offence against legal rules.

(iii) It must be of an actual or supposed offender for his offence.

(iv) It must be intentionally administered by human beings other than the offender.

(v) It must be imposed and administered by an authority constituted by a legal system against which the offence is committed.

H.L.A. Hart, *Punishment and Responsibility: Essays in the Philosophy of Law* 4-5 (1988). For other notable examples of definitions of punishment, see Antony Flew, *The Justification of Punishment*, in *The Philosophy of Punishment: A Collection of Papers* 83, 85-87 (H.B. Acton ed., 1973); Kent Greenawalt, *Punishment*, in *Encyclopedia of Crime and Justice* 1336, 1341 (Sanford Kadish ed., 1983).

90. *E.g.*, Duff, *supra* note 6, at 1 (“It is agreed that a system of criminal punishment stands in need of some strenuous and persuasive justification. . . .”); Gertrude Ezorsky, *The Ethics of Punishment*, in *Philosophical Perspectives on Punishment* xi, xi (Gertrude Ezorsky ed., 1972) (noting that “the need to justify punishment is obvious”); Richard Wasserstrom, *Why Punish the Guilty?*, 20 *Princeton U. Mag.* 14 (1964), *reprinted in Philosophical Perspectives on Punishment* 328, 337 (Gertrude Ezorsky ed., 1972) (“Punishment is an evil, an unpleasantness; it requires that someone suffer. Its infliction demands justification.”).

91. *See, e.g.*, Primoratz, *supra* note 7, at 7 (“To punish means to inflict an evil. But to inflict evil on someone is something that, at least *prima facie* ought not to be done.”); C.L. Ten, *supra* note 20, at 3 (“We are not normally justified in depriving people of the things which they value, such as their liberty or their property.”); Jeffrie G. Murphy, *Marxism and Retribution*, 2 *Phil. & Pub. Aff.* 217 (1973), *reprinted in* Jeffrie G. Murphy, *Retribution, Justice, and Therapy* 93, 96-102 (1979) (punishment is a form of coercion by the state, which is *prima facie*, wrong).

92. *E.g.*, Rawls, *supra* note 7, at 4 (“The trouble about it [punishment] has not been that people disagree as to whether or not punishment is justifiable. Most people have held that, freed from certain abuses, it is an acceptable institution. Only a few have rejected punishment entirely. . . .”); Simmons et al., *supra* note 17, at vii (“[M]ost of us are confident that something like our current practice of punishment *can* be justified.”). For arguments that punishment should be abolished, see Duff, *supra* note 9, at 30-34 (distinguishing between contingent abolitionism which requires the abolition of punishment until certain changes in society are made and absolute abolitionism which insists that punishment be eliminated under all circumstances).

93. *E.g.*, Ezorsky, *supra* note 90, at xi (“[I]f the need to justify punishment is

Consequentialism justifies punishment, not because the offender deserves it, but because of the good consequences generated by punishment.<sup>94</sup> In contrast, retributivism justifies punishment not by recourse to the good consequences promoted, but solely because the offender deserves it.<sup>95</sup>

### 1. Consequentialist Justification

The leading influence on, and most well-known version of,<sup>96</sup> the consequentialist justification of punishment is Jeremy Bentham's utilitarianism.<sup>97</sup> Punishment is an evil<sup>98</sup> which may be justified only if the good consequences generated by punishment outweigh the evil.<sup>99</sup> That is, punishment is justified if it generates more utility, happiness, pleasure, benefit, or good consequences than disutility, suffering, pain, expense, or bad consequences.<sup>100</sup> Although the utility or good

obvious, the manner of doing so is not."); Rawls, *supra* note 7, at 4 ("The difficulty is with the justification of punishment: various arguments for it have been given by moral philosophers, but so far none of them has won any sort of general acceptance; no justification is without those who detest it."); Simmons et al., *supra* note 17, at vii ("[P]hilosophers (and other theorists) have yet to approach any general agreement about how legal punishment is to be justified.").

94. See *infra* notes 99-100 and accompanying text.

95. See *infra* notes 135-38 and accompanying text.

96. Primoratz, *supra* note 7, at 13 ("[T]he most comprehensive and thoroughly developed formulation of the theory in philosophical literature is still to be found in the writings of Bentham.").

97. For general discussions of Bentham's account of punishment, see A.C. Ewing, *The Morality of Punishment* 53-54, 59 (1929); Hart, *supra* note 89, at 18-20, 40-41; Ted Honderich, *Punishment: The Supposed Justifications* 51-52, 59 (1969); Matt Matravers, *Justice and Punishment: The Rationale of Coercion* 13-16 (2000); Primoratz, *supra* note 7, at 15-31; C.L. Ten, *supra* note 20, at 87-89, 143-46.

98. Jeremy Bentham, *An Introduction to the Principles of Morals and Legislation* 170 (Prometheus Books 1988) (1781) ("[A]ll punishment is mischief: all punishment in itself is evil.").

99. *Id.* (punishment "ought only to be admitted in as far as it promises to exclude some greater evil").

Punishment of an offender is impermissible if the cost of punishment outweighs its good consequences. According to Bentham, punishment should not be inflicted at all under either of the following four conditions, where punishment would be: (i) "groundless" because there is no crime or harm, (ii) "inefficacious" because the crime cannot be deterred, (iii) "unprofitable," or too "expensive" because the evil of the punishment would exceed the crime, and (iv) "needless" because the crime may be deterred by other means than punishment or does not require deterrence. *Id.* at 171-77, 314-23.

100. For Bentham, conduct is to be judged according to the principle of utility, *id.* at 2 ("By the principle of utility is meant that principle which approves or disapproves of every action whatsoever, according to the tendency which it appears to have to augment or diminish the happiness of the party." (citation omitted)), the maximization of pleasure or happiness and the minimization of pain, *id.* at 1 ("Nature has placed mankind under the governance of two sovereign masters, *pain* and *pleasure*. It is for them alone to point out what we ought to do, as well as to determine what we shall do."), "the greatest happiness principle," *id.* at 5 n.1, "the greatest happiness of the greatest number," *id.*, and whether a particular course of

consequences promoted by punishment generally entails crime prevention,<sup>101</sup> punishment more specifically may promote the good consequences of rehabilitating the offender so that she will not commit future crimes,<sup>102</sup> incapacitating the offender so that she cannot commit crimes during the term of imprisonment,<sup>103</sup> deterring the offender from committing future crimes (i.e., specific deterrence) and, deterring others in society from committing future crimes (i.e., general deterrence).<sup>104</sup> The most important of these consequences of punishment that prevents crime is the general deterrent effect.<sup>105</sup>

Perhaps the first punishment theorist to weigh the benefits of the state entering into bargains with defendants<sup>106</sup> was Cesare Beccaria,<sup>107</sup> a precursor to Bentham's utilitarianism.<sup>108</sup> Writing in 1764, Beccaria found both "advantages" and "disadvantages" to immunity deals.<sup>109</sup> The cited disadvantages sound quaint to the modern sensibility: immunizing a defendant for inculcating a confederate promotes

conduct attains more happiness than pain, *id.* at 3 (an action or governmental action is "conformable to or dictated by the principle of utility, when . . . the tendency which it has to augment the happiness of the community is greater than any which it has to diminish it").

101. For Bentham, "[t]he immediate principal end of punishment is to control action." *Id.* at 170 n.1. "General [crime] prevention ought to be the chief end of punishment as it is its real justification." Jeremy Bentham, *Principles of Penal Law*, in 1 *The Works of Jeremy Bentham* 396 (John Bowring ed., 1962).

102. Influence of an offender, by punishment, is exerted on his "will, in which case it is said to operate in the way of *reformation*." Bentham, *supra* note 98, at 170-71 n.1.

103. Influence of an offender, by punishment, is exerted on his "physical power, in which case it is said to operate by *disablement*." *Id.*

104. Regarding others in the community, punishment influences "their wills; in which case it is said to operate in the way of *example*." *Id.* at 171 n.1.

105. *Id.* (punishment's general deterrent effect on others in the community, for example, as he terms it, "is the most important end of all"); Murphy & Coleman, *supra* note 7, at 118 ("While incapacitation and reform may sometimes figure in the [consequentialist] justification of punishment, *deterrence* has always been the mainstay . . ."); Primoratz, *supra* note 7, at 10 ("The most important consequences of punishment are its preventive effects . . .").

106. Although commentators pay comparatively greater attention—with respect to the purposes and justifications of punishment—to plea bargaining than to immunity deals, much of the discussion of plea bargaining substantially applies to immunity deals as well (and vice-versa). This applies not only to the consequentialist justification, but also the retributivist critique. See *infra* Part I.B.2.

107. Though Beccaria's theory of punishment is considered utilitarian and a significant influence on Bentham, Beccaria sought to fuse elements of both utility and justice (i.e., consequentialism and retributivism). See, e.g., Beccaria, *supra* note 1, at 60 ("[A] punishment for a crime cannot be deemed truly just . . . unless the laws have adopted the best possible means . . . to prevent that crime."); *id.* at 43 ("[E]ven if punishments produce a happy result, they are not therefore just, for, in order to be just, they must be necessary. A useful injustice cannot be tolerated . . .").

108. H.L.A. Hart distinguishes Beccaria's theory of punishment from Bentham's: "there is in Beccaria a respect for the dignity and value of the individual person which is absent in Bentham." H.L.A. Hart, *Essays on Bentham: Studies in Jurisprudence and Political Theory* 50 (1982).

109. Beccaria, *supra* note 1, at 69-70.



“cowardly crimes”<sup>110</sup> and rewards “betrayal”<sup>111</sup> by breeding dishonor among thieves.<sup>112</sup> The advantages of immunity deals noted by Beccaria resound in the more familiar currency of the promotion of deterrence and societal security.<sup>113</sup>

For a hodgepodge of reasons, modern commentators generally maintain that consequentialism justifies bargain justice.<sup>114</sup> Plea bargaining, some argue, must produce the best state of affairs, at least for the parties who enter them, otherwise they would not enter into them.<sup>115</sup> As opposed to a trial, bargaining will reduce the expense and

110. Offenders who accept offers of immunity commit two crimes: first, the “crimes of courage,” for which the offender has been apprehended and, second, the “cowardly crimes” of testifying against their partners in crime. *Id.* at 69. As Beccaria explains, the “crime” of accepting an immunity deal supplying inculpatory evidence against one’s confederates in crime may be an even graver crime than the initial, more conventional, crime (in Beccaria’s term, the “crimes of courage”) for which one was apprehended. “[C]rimes of courage are less fatal to a nation than cowardly crimes, for courage is not common, and it needs only a benevolent guiding hand to make it work for the public good. Cowardice, however, is more common and contagious, and it is always more self-centered and self-sustaining.” *Id.*

111. *Id.* (by granting immunity “the nation authorizes betrayal, which is detestable even among scoundrels”).

112. A further disadvantage of immunity deals noted by Beccaria, which perhaps rings more true to our modern sensibility, is that, by proffering such deals, “the court shows its own uncertainty and the weakness of the law by imploring the aid of those who break it.” *Id.*

113. Beccaria explains the benefits of immunity deals:

The advantages of offering immunity from prosecution are that it prevents serious offenses and that it reassures the people who are frightened by crimes whose effects are apparent and whose perpetrators are unknown. Moreover, it helps to show that a person who breaks faith with the laws (which is to say, with public conventions) will probably not keep faith in his private dealings.

*Id.* at 69-70. These advantages may be furthered, and some of the disadvantages minimized, by a general statute authorizing immunity deals rather than courts authorizing them on a case-by-case basis:

It would seem to me that a general statute promising impunity to an accomplice who gives information about any crime whatever would be preferable to a special declaration in a particular case, for the mutual fear of taking a risk by himself that each accomplice would have would prevent association in crime. The court would not encourage the audacity of criminals by letting them see that their help was needed in a specific instance.

*Id.* at 70.

114. *E.g.*, Tunick, *supra* note 11, at 123 (“There are persuasive utilitarian reasons for plea bargaining.”). *See* Van den Haag, *supra* note 11, at 171 (“The risk, the cost, the time-wasting procedures, and the uncertainties of our legal process cause the interests of justice to be prosecuted best by avoiding trial if the accused is willing to plead guilty to a lesser charge.”); Alschuler, *supra* note 44, at 832 (“[P]rosecutors enter plea agreements primarily because these bargains seem to them to offer greater benefit to the state than the alternative of trial.”).

115. *See* Van den Haag, *supra* note 11, at 172 (plea bargaining “may well be in the best interests of all concerned”); Alschuler, *supra* note 44, at 834 (noting as a possible defense of plea bargaining that “if a plea bargain did not improve the positions of both the defendant and the state, one party or the other would insist upon a trial”);

time the state must expend to process a defendant.<sup>116</sup> “Without plea bargains . . . you’d have to multiply resources fifty fold to get the job done that’s being done today.”<sup>117</sup> Without inducing accomplices to testify against each other through immunity and plea agreements, many crimes would go unprosecuted.<sup>118</sup> Securing testimony by one accomplice against another by promising not to prosecute the cooperating accomplice may “be necessary to the public interest.”<sup>119</sup> An offender who has bargained for a guilty plea “may manifest remorse or a willingness to enter the correctional system in a frame of mind that may afford hope for rehabilitation over a shorter period of time than would otherwise be necessary.”<sup>120</sup> Exchanging the possibility of a greater, but uncertain, punishment for a lighter, but certain, punishment, achieves greater deterrence.<sup>121</sup> The idea that “half a loaf is better than none”<sup>122</sup> underscores the consequentialist premise of bargain justice: some certain punishment is better than the risk of none.

---

Frank Easterbrook, *Criminal Procedure as a Market System*, 12 J. Legal Stud. 289, 309 (1983) (“If defendants and prosecutors (representing society) both gain, the process is desirable.”).

116. Tunick, *supra* note 9, at 138-39 (plea bargaining “is a quick and easy way to avoid the costs of trials and cope with heavy caseloads” (citation omitted)); Tunick, *supra* note 11, at 123 (“Plea bargaining allows tremendous savings in resources. The prosecutor (the state) avoids the cost of trial, and the defendant gets off with a lighter, though certain, sentence.”); Van den Haag, *supra* note 11, at 171 (“Since it is difficult, chancy, expensive, and far too time-consuming to obtain conviction through trial in our courts, prosecutors must resort to plea-bargaining.”); Michael Gorr, *The Morality of Plea Bargaining*, 26 Soc. Theory & Practice 129, 146 (2000) (“[I]t must be conceded that defenders of the efficiency argument [supporting the practice of plea bargains] have a strong prima facie case for their view.”); Hartman & Koval, *supra* note 2, at 73 (“The economic efficiency of plea bargaining is the strongest argument in favor of maintaining the system.”).

117. David E. Rovella & Gail Daine Cox, *Fallout from Singleton Bribe Ruling*, Nat’l L. J., Aug. 24, 1998, at 1A (quoting former prosecutor Martin H. Belsky).

118. *See, e.g.*, *United States v. White*, 27 F. Supp. 2d 646, 649 (E.D.N.C. 1998) (without the capacity to purchase testimony through plea and immunity agreements “the government would be unable to . . . effectively proceed in the thousands of cases each year in which it relies on witnesses who testify in return for leniency”); *United States v. Reid*, 19 F. Supp. 2d 534, 537 (E.D. Va. 1998) (acknowledging the government’s reliance on bargaining for testimony to secure convictions).

119. U.S. Dep’t of Justice, *United States Attorneys’ Manual*, § 9-27.600(A) (authorizing federal prosecutors to secure testimony by offering immunity when it is otherwise unavailable and is in the public interest), *available at* <http://www.usdoj.gov> (last visited Sept. 16, 2003).

120. Alschuler, *supra* note 44, at 834 (noting this rationale as a possible defense of plea bargaining).

121. *E.g.*, Easterbrook, *supra* note 115, at 309. *Cf.* Gorr, *supra* note 116, at 141 (“The likelihood that a system of plea bargaining would result in a higher conviction rate within the class of factually guilty defendants also serves to cast doubt on the objections [to plea bargaining] based on deterrence and social protection.”).

122. Alschuler, *supra* note 44, at 832 (ascribing to prosecutors the view that in light of the significant possibility that a defendant might well be acquitted at trial, some punishment through a plea bargain is better than no punishment subsequent to a trial).

The Supreme Court supplies perhaps the most important consequentialist defense of bargain justice.<sup>123</sup> According to the Court, plea bargaining is not merely permissible, but is “essential,”<sup>124</sup> “to be encouraged,”<sup>125</sup> “highly desirable,”<sup>126</sup> and “benefit[s] all concerned.”<sup>127</sup> In utilizing plea bargains, as opposed to trials, the state achieves greater efficiency in the criminal justice system at a significantly lower cost.<sup>128</sup> Not only do defendants secure a possibly more lenient punishment, but they also “avoid extended pretrial incarceration and the anxieties and uncertainties of a trial.”<sup>129</sup> And plea bargaining “protects the public from those accused persons who are prone to continue criminal conduct even while on pretrial release.”<sup>130</sup> In addition, plea bargaining enhances the prospects for successful rehabilitation of the offender.<sup>131</sup>

## 2. Retributivist Critique

The theories of Immanuel Kant<sup>132</sup> and G.W.F. Hegel exert the greatest historical influence on retributivism.<sup>133</sup> Kant<sup>134</sup> articulates the

123. See Alan Wertheimer, *Freedom, Morality, Plea Bargaining, and the Supreme Court*, 8 Phil. & Pub. Aff. 203, 230 (1979) (“[T]he Court’s principal defense of plea bargaining is consequentialist.”).

124. *Santobello v. New York*, 404 U.S. 257, 260 (1971).

125. *Id.*

126. *Id.* at 261.

127. *Blackledge v. Allison*, 431 U.S. 63, 71 (1977). *Accord* *Bordenkircher v. Hayes*, 434 U.S. 357, 362 (1978) (citing the mutuality of benefit to both prosecutors and defendants because of their “relatively equal bargaining power” (quoting *Parker v. North Carolina*, 397 U.S. 790, 809 (1970))).

128. *Blackledge*, 431 U.S. at 71 (“Judges and prosecutors conserve vital and scarce resources.”); *Santobello*, 404 U.S. at 260 (same).

129. *Blackledge*, 431 U.S. at 71.

130. *Santobello*, 404 U.S. at 261; see *Blackledge*, 431 U.S. at 71 (“The public is protected from the risks posed by those charged with criminal offenses who are at large on bail while awaiting completion of criminal proceedings.”) (citation omitted).

131. See *Blackledge*, 431 U.S. at 71 (the defendant “gains . . . a prompt start in realizing whatever potential there may be for rehabilitation”); *Santobello*, 404 U.S. at 261 (“[B]y shortening the time between charge and disposition, it enhances whatever may be the rehabilitative prospects of the guilty when they are ultimately imprisoned.”); *Brady v. United States*, 397 U.S. 742, 753 (1970) (the willingness of the accused to enter into plea negotiations may demonstrate some “hope for success in rehabilitation”). *But see* *North Carolina v. Alford*, 400 U.S. 25, 27-29 (1970) (rejecting the rehabilitation rationale for plea bargaining).

132. See George Fletcher, *Basic Concepts of Criminal Law* 36 (1998) (referring to Kant as “the leading philosopher of retributive punishment”).

133. Primoratz, *supra* note 7, at 13 (“The most important and influential among classical retributivists are Kant and Hegel.”).

134. For general discussions of Kant’s theory of punishment, see, for example, Ewing, *supra* note 97, at 15-21; Jeffrie Murphy, *Kant: The Philosophy of Right* 140-44 (1970); Roger J. Sullivan, *Immanuel Kant’s Moral Theory* 243-44 (1989); George Fletcher, *Law and Morality: A Kantian Perspective*, 87 Colum. L. Rev. 533, 551-52 (1987) (Symposium on Kantian Legal Theory). See generally Thomas Hill, *Kant on Wrongdoing, Desert, and Punishment*, 18 Law and Phil. 407 (1999); Donald Scheid, *Kant’s Retributivism*, 93 Ethics 262 (1983).

essential retributivist thesis:<sup>135</sup> punishment “must always be inflicted on him [an offender] only because he has committed a crime.”<sup>136</sup> Punishment’s consequences must not be the source of its justification because this would disrespect the offender by treating her as a mere means whose punishment is meant to benefit society.<sup>137</sup> And even if no good consequence may be extracted from it, “punishment is a categorical imperative,”<sup>138</sup> or absolute duty.

If retributivism “is the notion that there is a mystic bond between wrong and punishment,”<sup>139</sup> Hegel’s account<sup>140</sup> supplies retribution as the logical bond.<sup>141</sup> Hegel explains that both crime and punishment are coercive which, in the abstract, is “contrary to right.”<sup>142</sup> Because crime precedes punishment, crime is the initial coercion<sup>143</sup> and thus

135. Moore, *supra* note 20, at 88 (“The distinctive aspect of retributivism is that the moral desert of an offender is a *sufficient* reason to punish him or her.”); Hugo Bedau, *Concessions to Retribution in Punishment, in Justice and Punishment* 51, 53 (J.B. Cederbloom & William L. Blizek eds., 1977) (“[A] retributivist holds that a punishment is just if and only if the offender deserves it.”); Greenawalt, *supra* note 7, at 347 (retributivism is the view that “punishment is justified because people deserve it”); Quinton, *supra* note 19, at 7 (“The essential contention of retributivism is that punishment is only justified by guilt.”). See Bedau, *supra*, at 52 (“Probably the most widely held assumption about retribution in punishment is the idea that it makes desert the central feature of just punishment.” (citation omitted)).

136. Kant, *supra* note 8, at 140 (emphasis omitted).

137. *Id.* (punishment “can never be inflicted merely as a means to promote some other good for the criminal himself or for civil society”); *id.* (“a man can never be treated merely as a means to the purposes of another”).

138. *Id.* at 141; *id.* at 145 (“[F]ailure to punish (*impunitas criminis*) is the greatest wrong against his [a sovereign’s] subjects.”).

139. Oliver W. Holmes, Jr., *The Common Law* 42 (reprint ed. 1991) (1881).

140. For general discussion of Hegel’s theory of punishment see, for example, Ewing, *supra* note 97, at 21-26; Honderich, *supra* note 97, at 45-48; Alan W. Norrie, *Law, Ideology and Punishment: Retrieval and Critique of the Liberal Ideal of Criminal Justice* 65-88 (1991); Primoratz, *supra* note 7, at 67-81. See generally Tunick, *supra* note 11; Jami L. Anderson, *Annulment Retributivism: A Hegelian Theory of Punishment*, 5 *Legal Theory* 363 (1999); Peter J. Steinberger, *Hegel on Crime and Punishment*, 77 *Am. Pol. Sci. Rev.* 858 (1983). For the view that Hegel has developed a “richer” and more “systematic” account of retributivism than Kant, see Primoratz, *supra* note 7, at 13. *Contra* Honderich, *supra* note 97, at 45 (dismissing Hegel’s account as “obscure” and “of very secondary interest”).

141. See G.W.F. Hegel, *Elements of the Philosophy of Right* 129 (Allen W. Wood ed., H.B. Nisbet, trans., Cambridge Univ. Press 1991) (1821) (“Retribution is the inner connection and the identity of two determinations [the value and character of an offender’s crime and punishment] which are different in appearance and also have a different external existence [*Existenz*] in relation to one another.” (bracketed text in original)); *id.* (“[P]unishment is merely a manifestation of the crime, i.e. it is one half which is necessarily presupposed by the other.”); *id.* (committing the criminal “deed brings its own retribution with it”); Markus Dirk Dubber, *Rediscovering Hegel’s Theory of Crime and Punishment*, 92 *Mich. L. Rev.* 1577, 1581 (1994) (reviewing Mark Tunick, *Hegel’s Political Philosophy: Interpreting the Practice of Legal Punishment* (1992)). “Hegel developed the deontological foundations of Kant’s theory into an all-encompassing theory of the logical connection between crime and punishment.” *Id.*

142. Hegel, *supra* note 141, at 120 (emphasis omitted).

143. *Id.* at 121 (“The initial use of coercion, as force employed by a free agent in

the first wrong, whereas punishment is merely (in the abstract) the second wrong.<sup>144</sup> But because punishment follows crime and "coercion is cancelled by coercion; it [punishment] is therefore not only conditionally right but necessary—namely as a second coercion which cancels an initial coercion."<sup>145</sup> Since crime is a negation of the right,<sup>146</sup> and punishment negates crime,<sup>147</sup> punishment is, in Hegel's famous phrase, "the negation of the negation."<sup>148</sup> That is, rather than two wrongs making a right, the two wrongs (crime and punishment) restore the right.<sup>149</sup>

The retributivist rejection of bargain justice perhaps begins with Kant's consideration of a hypothetical proposal to commute the death sentence of a prisoner in return for the prisoner agreeing to undergo medical experiments that might yield information beneficial to society.<sup>150</sup> Concluding "a court would reject with contempt such a proposal,"<sup>151</sup> Kant declares that "justice ceases to be justice if it can be bought for any price whatsoever."<sup>152</sup> If reductions in deserved punishment for the societal benefits of advancements in medical science were antithetical to justice, then Kant presumably would find the reductions in deserved punishment entailed by bargain justice also unjust. George Fletcher, a leading exponent of Kant among legal scholars,<sup>153</sup> concurs:

It is patently unjust, in Kant's view, to punish some offenders less because they are willing to cooperate in some way with the state. Kant would turn over in his Koenigsberg grave if he knew about the modern American and growing European practice of plea bargaining, under which the prosecution makes special deals with certain suspects in return for their providing evidence against other suspects.<sup>154</sup>

---

such a way as to infringe the existence [*Dasein*] of freedom in its concrete sense—i.e. to infringe right as right—is *crime*." (bracketed text in original).

144. Primoratz, *supra* note 7, at 69 (interpreting Hegel's view to be that "retribution is a kind of coercion").

145. Hegel, *supra* note 141, at 120 (emphasis omitted).

146. *Id.* at 122 ("Right, whose infringement is crime . . .").

147. *Id.*

148. *Id.* at 123.

149. *Id.* at 124 (referring to punishment of crime as "the restoration of right"); see also *id.* at 252 (the law "restores and thereby actualizes itself as valid through the cancellation [punishment] of the crime").

150. Kant, *supra* note 8, at 141. The proposal, set out in full, is as follows:

What, therefore, should one think of the proposal to preserve the life of a criminal sentenced to death if he agrees to let dangerous experiments be made on him and is lucky enough to survive them, so that in this way physicians learn something new of benefit to the commonwealth?

*Id.*

151. *Id.*

152. *Id.*

153. See generally Fletcher, *supra* note 134.

154. Fletcher, *supra* note 132, at 36-37.

Thus, according to Fletcher, Kant would reject bargain justice.<sup>155</sup> Though Hegel says nothing of bargain justice,<sup>156</sup> scholars construe his writings as supporting its condemnation.<sup>157</sup>

Illustrating that bargain justice is illegitimate for violating the retributivist principle of giving offenders what they deserve,<sup>158</sup> Kenneth Kipnis analogizes between “grade bargain[ing]”<sup>159</sup> and plea bargaining. Suppose a professor, facing a huge stack of exams to grade, looks at your exam very quickly and decides that you probably would receive a grade of D.<sup>160</sup> Looking to reduce her workload, the professor offers you the following bargain: if you forego your right to a careful assessment of your exam and to receiving the grade that you actually deserve (which possibly might be an A, but will more

155. *Id.*

156. Mark Tunick explains one possible reason is that Hegel’s “Prussia no doubt had no such practice [plea bargaining].” Tunick, *supra* note 11, at 124 n.82. This explanation, however, may not be entirely persuasive since Hegel cited and discussed Beccaria’s writings on punishment, Hegel, *supra* note 141, at 126-27, and Beccaria considered arguments both pro and con concerning bargain justice. See *supra*, notes 109-13 and accompanying text.

157. See Tunick, *supra* note 11, at 124 (“Hegel insists that we punish for the sake of justice, and he would strictly scrutinize a penal system that sacrificed justice for some other purpose.”). Tunick anticipates the objection Hegel might have mounted against bargain justice had the practice been prevalent in his time:

[W]hen we punish someone who cops a plea, we are punishing them for something they did not do. If we think that the point of punishing at all is to express our condemnation of an act we regard as wrong and to vindicate right, we might think plea bargaining misses the point of why we punish.

*Id.* at 123-24 (citations omitted). Hegel does, however, criticize a jury for intentionally convicting a defendant of a lesser offense not committed rather than the greater offense which the jury knew the defendant actually did commit:

The closest he [Hegel] comes [to criticizing plea bargains] is in a passage in the lectures critical of a jury that fudged the facts of a case by finding a theft that was in fact greater than forty shillings, and therefore punishable by death according to English law, to have been a theft of less than forty shillings, and therefore not subject to capital punishment. Hegel does not condone the law that imposes death for a relatively trivial crime, but neither does he condone the mixing of legal consequences with a determination of facts; the investigation of facts should be “pure.”

*Id.* at 124 (citation omitted). Since Hegel found this objectionable, we might also expect Hegel to object to defendants pleading guilty to lesser offenses not actually committed:

If Hegel is willing to condemn a jury for punishing a defendant for a lesser offense not actually committed, albeit for benevolent purposes, perhaps he would condemn a prosecutor whose plea bargain, offered not even out of benevolence but merely to clear court dockets, results in the punishment of a person for a lesser offense not actually committed.

*Id.* at 124. For a critical response to Tunick’s attempt to infer that Hegel would have opposed bargain justice, see Dubber, *supra* note 141, at 1600-01.

158. Kipnis rejects bargain justice because our criminal justice system should be “one in which persons are justly given, not what they have bargained for, but what they deserve, irrespective of their bargaining position.” Kipnis, *supra* note 5, at 104. For a critique of Kipnis’s argument, see Gorr, *supra* note 116, at 140-41.

159. Kipnis, *supra* note 5, at 104.

160. *Id.*

probably be a D), the professor will agree to give you a grade of B.<sup>161</sup> Despite the consequentialist benefits to such a bargain and to a university that systematically institutionalized such a grade bargaining procedure,<sup>162</sup> for Kipnis grade bargaining, no less than plea bargaining, is simply “outrageous.”<sup>163</sup> Both grade and plea bargaining,<sup>164</sup> Kipnis concludes, are “out of place . . . [because] grades, like punishments, should be deserved.”<sup>165</sup>

Although retributivists condemn bargain justice for a variety of other reasons,<sup>166</sup> the discussion of them here is abbreviated because the next part examines in comparative depth five central duties of retributivism antithetical to bargain justice: (i) the duty to punish culpable wrongdoers, (ii) to give them their just deserts, (iii) to honor their right to punishment, (iv) to treat them as ends rather than as mere means, and (v) to treat those with equal desert equally.<sup>167</sup>

## II. ASSESSING RETRIBUTIVISM VIA BARGAIN JUSTICE

This part applies the competing theories of punishment to a hypothetical immunity offer—the Prosecutor’s Dilemma. Given consequentialism’s justification of bargain justice,<sup>168</sup> that consequentialism justifies acceptance of the offer fails to surprise.

161. *Id.* at 104-05.

162. Kipnis suggests the following benefits for the student, the professor, and the university:

[Grade bargaining] permits more students to be processed by fewer instructors. Teachers can be freed from the burden of having to read and criticize every paper. . . . [A] quick assignment of a grade is a more effective influence on the behavior of students, . . . [and] grade bargaining is necessary to the efficient functioning of the schools. There can be no doubt that students who have negotiated a grade are more likely to accept and to understand the verdict of the instructor. . . . Finally, a quickly assigned grade enables the guidance personnel and the registrar to respond rapidly and appropriately to the student’s situation.

*Id.* at 105.

163. *Id.*

164. For a criticism of the analogy between grade and plea bargaining, see Tunick, *supra* note 9, at 141 n.66.

165. Kipnis, *supra* note 5, at 105. Kipnis explains the incongruity of bargaining and desert:

Justice in retribution, like justice in grading, does not require that the end result be acceptable to the parties. To reason that because the parties are satisfied the bargain should stand is to be seriously confused. For bargains are out of place in contexts where persons are to receive what they deserve. And the American courtroom, like the American classroom, should be such a context.

*Id.*

166. See, e.g., Tunick, *supra* note 9, at 140 (“Punishment, the purpose of which is to express our condemnation of a wrong, loses its meaning when applied to a person who cops a plea, because it fails to take seriously the nature of that wrong.”). For additional rationales, see *infra* Part II.B.

167. See *infra* Part II.B.1-5, respectively.

168. See *supra* notes 113-31 and accompanying text.

And, in light of retributivists' criticisms of bargain justice,<sup>169</sup> we would expect a retributivist prosecutor<sup>170</sup> to refuse, as would Kant, such a proposal with "contempt[,] . . . for justice ceases to be justice if it can be bought for any price whatsoever."<sup>171</sup> Paradoxically, however, retributivism requires both acceptance and nonacceptance of the offer.

#### A. *The Prosecutor's Dilemma: A Hypothetical Immunity Offer*

In the Prosecutor's Dilemma, suppose only one member of a group of three culpable wrongdoers who have committed a serious crime is apprehended by the authorities. All three deserve the same degree of punishment because each committed the same wrongdoing with the same culpability.<sup>172</sup> The one culpable wrongdoer in custody (the "One") offers to inculcate the other two (the "Two") in exchange for immunity from prosecution and punishment.<sup>173</sup> The Two cannot be apprehended (or convicted and punished) without the One's inculpatory information and testimony.<sup>174</sup> If the offer is accepted, two culpable wrongdoers will be convicted and receive their just deserts, but one will not. If the offer is not accepted, at most one culpable wrongdoer will be convicted and receive her just deserts, but at least two will not.<sup>175</sup> Under the terms of the Prosecutor's Dilemma, the only way to punish the One is by not accepting the offer; and, the only

169. See *supra* notes 150-67 and accompanying text.

170. Of course, there are probably very few, if any, *purely* retributivist prosecutors. The construct of a purely retributivist prosecutor is merely a device by which to test whether the theory of retributivism would justify acceptance or nonacceptance of the offer.

171. Kant, *supra* note 8, at 141.

172. In addition to obtaining convictions of two rather than merely one wrongdoer, a prosecutor might also wish to engage in bargain justice to obtain the convictions of an offender who has committed a more grave crime as opposed to an offender who has committed a less serious crime. See Moore, *supra* note 20, at 156-58.

173. In the prisoner's dilemma, the prosecutor makes the immunity or leniency offer. See *supra* note 25 and accompanying text. In contrast, here, in the Prosecutor's Dilemma, the prisoner initiates the offer because, given retributivists' criticisms of bargain justice, see *supra* notes 150-67 and accompanying text, a retributivist prosecutor would presumably never tender such an offer.

174. In the interests of simplicity and clarity the bargain justice offer by which retributivism will be demonstrated to sustain internal inconsistencies has been cast as an offer of immunity. But casting the hypothetical bargain justice offer as a plea bargain could, with more complications, serve the same purpose. Suppose the three culpable wrongdoers' criminal scheme renders them liable for a number of criminal charges. The One makes an offer to the prosecutor to plead guilty to all of the charges but one and to testify—providing inculpatory evidence—against the Two. In return for the One's guilty plea and his agreement to testify against the Two, the One requests immunity (or the prosecutor's promise not to charge him) for the remaining crime.

175. Not accepting the offer encompasses two distinct options: (i) punishment of the One or (ii) punishment of neither the One nor the Two. Thus, nonacceptance of the offer entails punishment of *at most* one offender.



way to punish the Two is by accepting the offer. Thus, punishing both the One and the Two is impossible. The resulting situation is that the offer presents the following three, and *only* these three, possibilities to a prosecutor: (i) accept the offer and punish only the Two, (ii) not accept the offer and punish only the One or, (iii) not accept the offer and punish neither the Two nor the One.<sup>176</sup>

Based on their differing views of bargain justice, we would expect consequentialism and retributivism to justify different responses to the offer.<sup>177</sup> Under consequentialism, the choice is clear—accept the offer so as to maximize the number of persons punished (two versus one), maximize deterrence,<sup>178</sup> and save the expense of a trial.<sup>179</sup> But, the choice is far from clear under retributivism.

### B. *Retributivist Principles Applicable to the Immunity Offer*

Whether or not retributivism justifies acceptance or nonacceptance of the offer implicates five central principles or duties<sup>180</sup> of retributivism.<sup>181</sup> Their application to the offer reveals retributivism to

176. The resulting difficulties for retributivism as applied to the Prosecutor's Dilemma arise not only under the specific terms of the Prosecutor's Dilemma but also under a number of its variations. For these variations, see *infra* text accompanying note 298, *infra* notes 299, 337.

177. See *supra* Parts I.B.1 and I.B.2, respectively.

178. See *supra* notes 121, 122 and accompanying text.

179. See *supra* notes 116, 117, 128 and accompanying text.

180. For the five principles or duties, see *supra* text accompanying note 167.

181. There are other principles of retributivism antithetical to bargain justice, in addition to the five enumerated, but these other principles do not specifically pertain to the Prosecutor's Dilemma. For example, some retributivists argue that plea bargains undermine the duty not to punish the innocent. *E.g.*, Tunick, *supra* note 9, at 143 (in some instances, "the result of the negotiated plea is to punish an innocent person, which violates the weak 'negative retributive principle'"). Some argue that plea bargains make punishment of the innocent more likely than in trials:

While the conviction of the innocent would be a problem in any system we might devise, it appears to be a greater problem under plea bargaining. . . .

Plea bargaining substantially erodes the standards for guilt and it is reasonable to assume that the sloppier we are in establishing guilt, the more likely it is that innocent persons will be convicted.

Kipnis, *supra* note 5, at 105-06; see Schulhofer, *supra* note 3, at 2007 ("[C]onvicting the innocent is unequivocally easier in a world that permits plea-bargaining."). *Contra* Bordenkircher v. Hayes, 434 U.S. 357, 362 (1978) (asserting that, given adequate counsel and existing procedural safeguards, a defendant "was unlikely to be driven to false self-condemnation"); Gorr, *supra* note 116, at 140-41 (arguing there is no evidence that more innocents plead guilty under a plea bargain than are mistakenly convicted in a trial). *Cf.* Wertheimer, *supra* note 123, at 233 ("That plea bargaining induces some innocent defendants to plead guilty is not, in itself, a decisive moral objection to the practice. Any system will occasionally punish innocent persons unintentionally."). For the argument that retributivism justifies the intentional punishment of particular, identifiable innocents, see Russell L. Christopher, *Detering Retributivism: The Injustice of "Just" Punishment*, 96 Nw. U. L. Rev. 843, 880-88 (2002).

paradoxically require, and be violated by, both acceptance and nonacceptance of the offer.

### 1. Categorical Duty to Punish Culpable Wrongdoers

One of retributivism's most important principles is the duty to punish culpable wrongdoers.<sup>182</sup> As Kant famously expounded, "[t]he Principle of punishment is a categorical imperative, and woe to him who crawls through the windings of eudaemonism in order to discover something that releases the criminal from punishment."<sup>183</sup> Hegel agrees that "it would be impossible for a society to leave a crime unpunished."<sup>184</sup> Though some modern forms of retributivism conceive punishment to be merely permissible,<sup>185</sup> the view that punishment of wrongdoers is obligatory remains the "dominant"<sup>186</sup> and "standard retributive view."<sup>187</sup> Thus, "the retributivist regards the punishment of the guilty to be categorically imperative whenever the opportunity to give such punishment presents itself."<sup>188</sup>

Michael Moore suggests that the conception of retributivism as embodying an absolute duty to punish each and every wrongdoer precludes a retributivist from engaging in immunity bargaining.<sup>189</sup> As Moore explains,

it might be thought that we rightly refuse to punish some guilty persons in order to be able to punish other, more seriously guilty persons—as when we give immunity, in order to extract testimony needed to convict the latter. How can the retributivist accommodate these practices, given that the retributivist regards the punishment of the guilty to be categorically imperative whenever the opportunity to give such punishment presents itself?<sup>190</sup>

182. Primoratz, *supra* note 7, at 12 ("[t]he moral duty to punish" is one of the five central tenets of retributivism (emphasis omitted)).

183. Kant, *supra* note 8, at 141. Kant also declared that the "failure to punish (*impunitas criminis*) is the greatest wrong against his [a sovereign's] subjects." *Id.*

184. Hegel, *supra* note 141, at 251; *see id.* at 127 ("Both the nature of crime and the criminal's own will *require* that the infringement [crime] for which he is responsible should be cancelled [punished].") (emphasis added)).

185. *See, e.g.*, H.J. McCloskey, *A Non-Utilitarian Approach to Punishment*, 8 *Inquiry* 249 (1965), *reprinted in* *Philosophical Perspectives on Punishment* 119, 132 (Gertrude Ezorsky ed., 1972). For general discussions of retributivist accounts of punishment as not obligatory, but merely permissible, *see* Richard Brandt, *Ethical Theory: The Problems of Normative and Critical Ethics* 501-03 (1959); H.J. McCloskey, *Utilitarian and Retributive Punishment*, 64 *J. Phil.* 91 (1967).

186. Primoratz, *supra* note 7, at 110.

187. Louis Kaplow & Steven Shavell, *Fairness Versus Welfare*, 114 *Harv. L. Rev.* 961, 1229 n.660 (2001).

188. Moore, *supra* note 20, at 156. For a similar view, *see* Michael Moore, *The Moral Worth of Retribution*, in *Responsibility, Character, and the Emotions* 179, 182 (Ferdinand Schoeman ed., 1987) ("For a retributivist, the moral culpability of an offender also gives society the *duty* to punish.").

189. Moore, *supra* note 20, at 156.

190. *Id.*

Thus, retributivism's absolute duty, or "categorical imperative,"<sup>191</sup> to punish seemingly constitutes sufficient reason to reject an offer of bargain justice which foregoes the opportunity to punish a culpable wrongdoer.<sup>192</sup>

If the duty to punish culpable wrongdoers at every opportunity requires nonacceptance of the offer, then it equally requires acceptance of the offer. Not accepting the offer entails foregoing punishment of the Two.<sup>193</sup> Since there is an opportunity to punish the Two,<sup>194</sup> foregoing that opportunity violates the duty to punish the Two. Thus, not accepting the offer violates the absolute duty to punish. Only acceptance of the offer affords fulfillment of the duty to punish with respect to the Two. On this basis, then, the duty to punish requires acceptance of the offer. But accepting the offer entails foregoing punishment of the One,<sup>195</sup> thereby violating the duty to punish with respect to the One. The only option affording punishment of the One is nonacceptance of the offer. On this basis, then, retributivism's duty to punish requires nonacceptance of the offer. But nonacceptance of the offer violates the duty to punish the Two. Therefore, fulfilling the duty to punish with respect to the One (not accepting the offer) violates the duty with respect to the Two; and that fulfilling the duty to punish with respect to the Two (accepting the offer) violates the duty to punish with respect to the One. The duty to punish culpable wrongdoers paradoxically requires, and is violated by, both acceptance and nonacceptance of the offer.<sup>196</sup>

191. Kant, *supra* note 8, at 141 (stating that "punishment is a categorical imperative"). See Moore, *supra* note 20, at 156.

192. Moore himself, however, argues that two special, controversial forms of retributivism, apparently not entailing a duty to punish, would justify acceptance of such an offer of bargain justice. Moore, *supra* note 20, at 156-59. For a discussion of these alternative forms of retributivism, see *infra* Part IV.A-B.

193. Nonacceptance of the offer entails either punishing the One or punishing none, but precludes punishing the Two. For a discussion of the three options which the Prosecutor's Dilemma presents to a retributivist prosecutor, and the ramifications of each option, see *supra* Part II.C.

194. Just as there is an opportunity to punish the One, so also there is an opportunity to punish the Two. The One's offer to supply information and testimony inculpating the Two presents the retributivist system of punishment with an opportunity to punish the Two. Retributivism imposes a duty to punish culpable wrongdoers whenever the opportunity arises. See *supra* notes 189-90 and accompanying text. Thus, the duty to punish applies to the Two no less than to the One.

195. Under the terms of the Prosecutor's Dilemma hypothetical, acceptance of the offer includes punishing the Two but precludes punishing the One.

196. Fulfillment of the duty requires a prosecutor to both accept and not accept the offer. But this is logically impossible. See *infra* note 249. The offer can only be accepted or not accepted. But either acceptance or nonacceptance of the offer violates the duty to punish.

## 2. Giving Culpable Wrongdoers Their Just Deserts

Integral to retributivism is the notion that punishment must be in proportion to, and to the extent of, an offender's just deserts.<sup>197</sup> Retributivism "insists that the punishment must fit the crime."<sup>198</sup> As Kant explained the *lex talionis* of an eye-for-an-eye and a tooth-for-a-tooth,<sup>199</sup> "what is done to him [the offender] in accordance with penal law is what he has perpetrated on others."<sup>200</sup> Hegel articulates a more sophisticated version of Kant's *lex talionis*.<sup>201</sup> Rather than the literal or specific equality of crime and punishment embodied in Kant's principle—"that what the criminal has done should also happen to him,"<sup>202</sup> which may easily be reduced to an absurdity<sup>203</sup>—for Hegel, punishment need only be generally equal or comparable to the crime.<sup>204</sup> Under either view, however, retributivism requires

197. See, e.g., *Tison v. Arizona*, 481 U.S. 137, 149 (1987) ("The heart of the retribution rationale is that a criminal sentence must be directly related to the personal culpability of the criminal offender."); Primoratz, *supra* note 7, at 12 (identifying the principle that "[p]unishment ought to be *proportionate* to the offense" as one of the five main tenets of retributivism).

198. Stanley I. Benn, *Punishment*, in 7 *The Encyclopedia of Philosophy* 29, 32 (Paul Edwards ed., 1972).

199. The *lex talionis* was first expressed in the Bible as follows: "If any harm follow, then thou shall give life for life, eye for eye, tooth for tooth, hand for hand, foot for foot, burning for burning, wound for wound, stripe for stripe." *Exodus* 21: 23-25.

200. Kant, *supra* note 8, at 169. As Kant further explains the principle of equality between crime and punishment,

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

*Id.* at 141.

201. See Primoratz, *supra* note 7, at 13; Dubber, *supra* note 141, at 1581. *But see* Honderich, *supra* note 97, at 45.

202. Hegel, *supra* note 141, at 127 (emphasis omitted).

203. Hegel points out the absurd ramifications of taking Kant's principle literally:

[I]t is very easy to portray the retributive aspect of punishment as an absurdity (theft as retribution for theft, robbery for robbery, an eye for an eye, and a tooth for a tooth, so that one can even imagine the miscreant as one-eyed or toothless); but the concept has nothing to do with this absurdity, for which the introduction of that [idea of] *specific equality* is alone to blame.

*Id.* at 128 (emphasis added and citation omitted).

204. Whereas the simple version of the *lex talionis* requires a specific equality between the crime and the punishment (e.g., theft as retribution for theft), Hegel requires merely that the crime and punishment be comparable in character or value. Hegel defines value as "the *inner equality* of things [*Sachen*] which, in their existence [*Existenz*], are specifically quite different." *Id.* (bracketed text in original). That is, the two must be generally equal, or comparable:

[E]quality remains merely the basic measure of the criminal's *essential* deserts, but not of the specific external shape which the retribution should

punishment to be proportional to, and reflective of, an offender's just deserts.<sup>205</sup>

Given retributivists' criticisms of bargain justice's failure to give culpable wrongdoers their just deserts,<sup>206</sup> one might expect the principle to require nonacceptance of the offer. While only nonacceptance of the offer (and punishing only the One) allows fulfillment of the principle with respect to the One, it foregoes punishment of the Two and thus violates the principle with respect to the Two. While only acceptance of the offer fulfills the principle with respect to the Two, nonetheless by foregoing punishment of the One it violates the principle with respect to the One. The resulting situation is that which enables the One to receive her just deserts (not accepting the offer) precludes the Two from receiving their just deserts; and that which enables the Two to receive their just deserts (accepting the offer) precludes the One from receiving her just deserts. The retributivist principle that offenders receive punishment proportional to their just deserts paradoxically requires, and is violated by, both acceptance of the offer and nonacceptance of the offer.<sup>207</sup>

### 3. Culpable Wrongdoers' Right to Punishment

Regarded as a central principle of retributivism,<sup>208</sup> an offender's right to punishment is, in fact, defended by only a handful of retributivists, albeit prominent ones. Hegel declared that "[t]he injury [*Verletzung*] which is inflicted on the criminal is not only just in itself . . . it is also a right for the criminal himself."<sup>209</sup> Contemporary retributivists Herbert Morris and Jeffrie Murphy (though perhaps Murphy is now only an ex-retributivist<sup>210</sup>) acknowledge Hegel's

take. It is only in terms of this specific shape that theft and robbery [on the one hand] and fines and imprisonment etc. [on the other] are completely unequal, whereas in terms of their value, i.e. their universal character as injuries [*Verletzungen*], they are *comparable*.

*Id.* at 129 (bracketed text in original).

205. Greenawalt, *supra* note 7, at 347-48 (for retributivism "the severity of punishment should be proportional to the degree of wrongdoing").

206. *See supra* notes 150-55, 157-65 and accompanying text.

207. *See supra* note 196.

208. Primoratz, *supra* note 7, at 12 (identifying the principle that "[p]unishment is a right of the offender" as one of the five principal tenets of retributivism).

209. Hegel, *supra* note 141, at 126 (emphasis omitted) (bracketed text in original). Hegel explains further how punishment is a right of the offender:

For it is implicit in his action, as that of a rational being, that it is universal in character, and that, by performing it, he has set up a law which he has recognized for himself in his action, and under which he may therefore be subsumed as under his right.

*Id.* (emphasis omitted).

210. Jeffrie Murphy, *Retribution Reconsidered: More Essays in the Philosophy of Law ix-x* (1992). Murphy explains his retreat from retributivism: "I no longer have the total confidence in either retributivism or Kantianism that I once had . . ." *Id.* at

principle by linking the right of punishment to, and as reflecting, the wrongdoer's status as an autonomous and responsible agent.<sup>211</sup>

Satisfying a culpable wrongdoer's right to punishment requires both nonacceptance of the offer and acceptance of the offer. Honoring the One's right to punishment requires nonacceptance of the offer. But not accepting the offer violates the Two's right to punishment. Honoring the Two's right to punishment requires acceptance of the offer. But accepting the offer violates the One's right to punishment. Therefore, fulfilling the One's right to punishment (not accepting the offer) violates the Two's right to punishment; and fulfilling the Two's right to punishment (accepting the offer) violates the One's right to punishment. The principle that culpable wrongdoers have a right to punishment paradoxically requires, and is violated by, acceptance of the offer as well as nonacceptance of the offer.<sup>212</sup>

#### 4. Treating Culpable Wrongdoers Not as Mere Means But as Ends

Kant demanded that one must "[a]ct so that you treat humanity, whether in your own person or that of another, always as an end and never as a means only."<sup>213</sup> Kant's maxim also applies to the punishment of offenders:

Punishment by a court . . . can never be inflicted merely as a means to promote some other good for the criminal himself or for civil society. . . . For a man can never be treated merely as a means for the purposes of another or be put among the objects of rights to things . . . .<sup>214</sup>

Though commentators consider Kant's maxim notoriously unclear,<sup>215</sup> its importance is considerable.<sup>216</sup> Using someone as a mere means is thought to belong to the same family as the concepts of

ix.

211. Herbert Morris, *On Guilt and Innocence: Essays in Legal Philosophy and Moral Psychology* 41 (1976) ("[A] person has a right to be punished, meaning by this that a person had a right to all those institutions and practices linked to punishment."); Jeffrie Murphy, *Moral Death: A Kantian Essay on Psychopathy*, 82 *Ethics* 284, in *Retribution, Justice, and Therapy* 128, 134 (1979) ("The right to be punished and regarded as a responsible agent, though sometimes painful when honored, at least leaves one's status as a moral person intact.").

212. See *supra* note 196.

213. Immanuel Kant, *Foundations of the Metaphysics of Morals* 46 (Lewis White Beck trans., 2d ed. 1990) (1785).

214. Kant, *supra* note 8, at 140-41.

215. See, e.g., Duff, *supra* note 6, at 178 (Kant's maxim is "notoriously obscure"); Honderich, *supra* note 97, at 60 ("vague"); Nancy Davis, *Using Persons and Common Sense*, 94 *Ethics* 387, 389 (1984) ("elusive"); Dolinko, *supra* note 13, at 1631 ("unclear"); Onora O'Neill, *Between Consenting Adults*, 14 *Phil. & Pub. Aff.* 252, 252 (1985) (there is "little agreement" as to its meaning).

216. See, e.g., Duff, *supra* note 6, at 178 (Kant's maxim "generates a crucial objection" to consequentialist theories); Greenawalt, *supra* note 89, at 1341 (the maxim is the source of "[t]he most fundamental objection" to consequentialism).

“manipulation, dehumanization, exploitation, and disrespect.”<sup>217</sup> Other common conceptions include treating someone as a thing,<sup>218</sup> instrument,<sup>219</sup> tool,<sup>220</sup> or animal.<sup>221</sup>

The duty to treat persons as ends and not as mere means contains both negative and positive components.<sup>222</sup> The negative duty is to not use persons as mere means; the positive duty is to affirmatively treat persons, at least in part, “as ends-in-themselves.”<sup>223</sup> According to Hegel, retributivism avoids treating an offender “simply as a harmful animal which must be rendered harmless, or punished with a view to deterring or reforming him”<sup>224</sup> (thereby satisfying the negative duty); and, by conceiving of punishment as “embodying the criminal’s own right, the criminal is honoured as a rational being” (thereby satisfying the positive duty).<sup>225</sup> Under a modern account, retributivism purportedly satisfies the negative duty by not using the punishment of a wrongdoer as a means to attain some societal benefit and satisfies the positive duty by giving offenders the punishment that they deserve.<sup>226</sup>

If treating culpable wrongdoers not as mere means, but as ends, requires nonacceptance of the offer, it equally requires acceptance of the offer. Treating the One not as a mere means, but as an end,

217. Davis, *supra* note 215, at 387 (citation omitted).

218. Kantian scholar Onora O’Neill explains how using someone as a mere means is akin to how things are treated:

To treat something as a mere means is to treat it in ways that are appropriate to things. Things, unlike persons, are neither free nor rational; they lack the capacities required for agency. They can only be props or implements, never sharers or collaborators, in any projects. Things cannot act . . . so cannot consent to or dissent from the ways in which they are used.

Onora O’Neill, *Constructions of Reason: Explorations of Kant’s Practical Philosophy* 138 (1989).

219. O’Neill, *supra* note 215, at 252 (“Making another into a tool or instrument in my project is one way of failing to treat that other as a person . . .”).

220. Robert Nozick explains the principle of using another as a mere means by the metaphor of a tool, which is apt because there are no moral constraints on how we may use a tool:

There is no side constraint on how we may use a tool, other than the moral constraints on how we may use it upon others. There are procedures to be followed to preserve it for future use (“don’t leave it out in the rain”), and there are more and less efficient ways of using it. But there is no limit on what we may do to it to best achieve our goals.

Robert Nozick, *Anarchy, State, and Utopia* 31 (1974).

221. See *infra* text accompanying note 224.

222. According to Kant, “it is not enough that he is not authorized to use either himself or others merely as means (since he could still be indifferent to them); it is in itself his duty to make man in general his end.” Kant, *supra* note 8, at 198.

223. O’Neill, *supra* note 218, at 113.

224. Hegel, *supra* note 141, at 126. For a similar view, see Jeffrie Murphy, *Retributivism and the State’s Interest in Punishment*, in *Nomos XXVII: Criminal Justice* 156, 158-59 (J. Roland Pennock & John W. Chapman eds., 1985).

225. Hegel, *supra* note 141, at 126 (emphasis omitted).

226. See Duff, *supra* note 18, at 10.

requires nonacceptance of the offer.<sup>227</sup> But nonacceptance of the offer foregoes punishment of the Two who thereby fail to receive the punishment that they deserve. By failing to receive the punishment that they deserve, the Two fail to be affirmatively treated as ends<sup>228</sup> (thereby violating the positive component of the duty with respect to the Two). Thus, not accepting the offer (and punishing only the One), violates the positive component of the duty.<sup>229</sup> But acceptance of the offer entails treating the One as a mere means thereby violating the negative component of the duty.<sup>230</sup> As a result, that which allows treating the One as an end (not accepting the offer) precludes treating the Two as ends; and that which allows treating the Two as ends (acceptance of the offer), precludes treating the One as an end. The duty to treat culpable wrongdoers as ends, and not as mere means, paradoxically requires, and is violated by, both acceptance and nonacceptance of the offer.<sup>231</sup>

### 5. Treating Culpable Wrongdoers with Equal Desert Equally

Though not always articulated as an express tenet of retributivism,<sup>232</sup> treating equally situated culpable wrongdoers equally

227. Accepting the offer constitutes using the One as a mere means to secure the punishment of the Two, thereby violating the negative component of the duty with respect to the One. In addition, by not punishing the One, the One would not be given the punishment that she deserves, thereby violating the positive component of the duty—affirmatively treating the One as an end in herself.

228. See *supra* note 226 and accompanying text.

229. Nonacceptance of the offer (and punishing only the One) fulfills the positive component of Kant's maxim with respect to the One, but not with respect to the Two.

230. Even if we were to assume *arguendo* that the One being treated as a mere means by acceptance of the offer constituted a sufficient basis to not accept the offer in the Prosecutor's Dilemma, it would not be a sufficient basis in, and would not resolve the difficulties presented by, variations of the Prosecutor's Dilemma. Suppose the situation is the same as in the Prosecutor's Dilemma except that instead of the One making the offer, an outside party makes the offer. The outside party, let us call her 'Informant,' informs the authorities that she will provide conclusive evidence of the Two's guilt in return for the prosecutor granting immunity to the One. As opposed to the One being treated as a mere means by acceptance of the One's offer, acceptance of Informant's offer does not constitute using the One as a mere means in order that the Two may be treated as ends by giving them their just deserts. If Informant's offer is accepted, the One is neither being used as a mere means nor being used at all. Therefore, even if the Kantian principle might be able to resolve the dilemma for a retributivist system as applied to some bargain justice offers, it does not provide a resolution for all such offers.

231. See *supra* note 196.

232. C.L. Ten finds that the equality of treatment among offenders to be either a form of, or a reason supporting, retributivism. C.L. Ten, *supra* note 20, at 50-51. Ten considers a hypothetical of whether a person should be punished even if no consequentialist benefit may be attained by doing so: "[Suppose] a Nazi war criminal . . . escapes to an uninhabited island where . . . he is found leading an 'idyllic existence.' While he is still unrepentant for what he has done, he has no desire to cause further harm. (Let us also assume that his punishment would have no general deterrent effect on the behavior of others.)" *Id.* at 47 (citing to the hypothetical of



is implicit in, or presupposed by, retributivism.<sup>233</sup> As Moore declares, “[l]ike cases should indeed be treated alike.”<sup>234</sup> Additionally, treating equally those offenders with equal desert may be derived from the principle that offenders must receive punishment proportionate to their desert.<sup>235</sup> Since an offender’s desert or degree of culpable wrongdoing is the only determinant for the degree of deserved punishment, then retributivism must require that offenders with the same degree of desert and culpability receive equal treatment.<sup>236</sup>

Fletcher construes Kant’s commitment to equality of treatment among offenders with equal desert as not a derivative,<sup>237</sup> but rather an

John Kleining, *Punishment and Desert* 67 (1973)). Although no good consequences may be drawn from punishing the Nazi war criminal, as the hypothetical stipulates, Ten proposes that a retributivist principle—equality of treatment of offenders—may nonetheless support punishment of the Nazi:

[Equality of treatment of offenders] supports the belief that the Nazi . . . should be punished even though his particular punishment cannot be defended on purely utilitarian grounds. This consideration rests on a comparison of the position of the Nazi and those of other offenders who are punished for committing crimes similar to, or even much less serious than, those of the Nazi. For example, an ordinary murderer is punished for his crime. His punishment, let us assume, is justified on utilitarian grounds. Similarly there are sufficient utilitarian reasons for punishing those who have committed lesser offences like assault and theft. But if all these offenders are punished, then it is unfair that the Nazi should not also be punished for a similar, or even much worse, offence. . . . Part of the appeal here is to a requirement that similar cases should be treated similarly, where the similarity is based on intrinsic features of the act, defined in terms of the intention of the agent and the natural and foreseeable consequences.

C.L. Ten, *supra* note 20, at 50-51. Ten, however, cautions that the retributivist principle of punishing similar offenders similarly does “not in [itself] dictate the punishment of the Nazi.” *Id.* at 51. It only “give[s] rise to a conditional demand that if similar, or lesser, offenders are justifiably punished, then so must the Nazi.” *Id.*

233. See Fletcher, *supra* note 21, at 58-59. Fletcher maintains that, since “all theories of justice are primarily concerned with equality, it makes sense to ground retributive justice . . . in a commitment to bring about equality . . . among offenders.” *Id.* at 59. Fletcher explains that retributive justice contains elements of both corrective and distributive justice: “The distributive dimension of punishments consists in the legal imperative to punish all offenders [of equal desert] equally.” *Id.* at 58.

234. Moore, *supra* note 20, at 90.

235. Duff derives the principle in just this way: “[T]he principle of proportionality is a principle of *relative* proportionality. Doing justice to individual offenders is therefore a matter of doing justice *between* offenders. We punish *this* offender justly by ensuring that her sentence is proportionate to her crime, relative to the sentences imposed on *other* offenders.” Duff, *supra* note 9, at 156 (citation omitted).

236. E.g., C.L. Ten, *supra* note 20, at 51 (discussing retributivism’s “requirement that similar cases should be treated similarly, where the similarity is based on intrinsic features of the act, defined in terms of the intention of the agent and the natural and foreseeable consequences”).

237. Kant’s formulation determines desert and the degree of punishment based on the specific equality between the crime and punishment, or the *lex talionis*: “what is done to him [the offender] in accordance with penal law is what he has perpetrated on others.” Kant, *supra* note 8, at 169. For Kant’s account of how the *lex talionis* determines the degree of a culpable wrongdoer’s punishment, see *supra* note 200. As

express tenet.<sup>238</sup> Fletcher grounds Kant's demand that equally situated offenders be treated equally in concerns of distributive justice.<sup>239</sup> "For this reason [distributive justice], the leading philosopher of retributive punishment, Immanuel Kant, stressed the imperative of maintaining strict equality among offenders."<sup>240</sup> As Fletcher explains, "Kant was so firm in his commitment to equality among offenders that he regarded any deviation from equality for the sake of practical advantage as illustrative of the principle: 'if justice goes, there is no longer any value in men's living on earth.'"<sup>241</sup>

Since retributivists criticize bargain justice for violating this principle,<sup>242</sup> it might seem that a retributivist prosecutor must, on this basis, not accept the offer (and punish only the One). Not accepting the offer (and punishing only the One) violates the principle and treats the One and the Two unequally.<sup>243</sup> Accepting the offer also

a result, if two offenders commit the same crime (with the same culpability), each must receive the same punishment as the other. In this way, we may derive Kant's support for equal punishment among equally culpable wrongdoers from his principle of equality between crime and punishment.

238. See George P. Fletcher, *Punishment and Responsibility*, in *A Companion to Philosophy of Law and Legal Theory* 514, 517 (Dennis Patterson ed., 1999) (interpreting Kant as criticizing utilitarianism because it "leads invariably to breaches in the principle of equality").

239. Fletcher, *supra* note 132, at 36-37. Fletcher explains how distributive justice is integral to retributivism:

The distinctive feature of [retributive] punishment is that it also has a distributive dimension. Distributive justice means that the benefits and burden[s] of living together in a society are distributed to each according to his due. Unless there is a sound basis for punishing some offenders more than others, distributive justice mandates equality in the distribution of the burden that punishment represents. Because the state is responsible for distributing the burdens of fines and imprisonment (not to mention the death penalty), it is critical that the state abide by criteria of distributive justice. This obligation implies that it may not discriminate in selecting some people and not others to suffer for their crimes.

*Id.* at 36.

240. *Id.*; see *id.* at 37 ("Kant was committed to the principle of equality among offenders.").

241. *Id.* at 37 (quoting Kant, *supra* note 8, at 141).

242. E.g., Fletcher, *supra* note 21, at 58-59 (bargain justice conflicts with "basic principles of equal justice"); Kipnis, *supra* note 5, at 105; Wertheimer, *supra* note 123, at 233 ("To the extent to which sentences are determined by factors unrelated to appropriate retributive or correctional considerations, plea bargaining entails that people in relevantly similar circumstances are treated differently only because some are prepared to waive their constitutional rights." (citation omitted)); *id.* at 234 (plea bargaining "may well involve 'unjust' discrimination between 'persons in similar circumstances'").

243. Arguably, the One and the Two are not equally situated in at least two respects. First, the Two can only be punished with the One's cooperation. The One can be punished without the Two's cooperation. Second, the One has expressed an interest in cooperating with the prosecutor, whereas the Two have not. Although these are differences between the situation of the One and the Two, nonetheless they are not morally relevant differences—at least not under retributivism. The only relevant basis on which retributivism justifies punishment is the offender's moral

violates the principle—punishing the Two but not the One treats them unequally. Both accepting and not accepting the offer violate, as they do the other four principles, the principle of treating equally culpable wrongdoers equally. Unlike the other four principles, however, there is an option which fulfills the principle—not accepting the offer and punishing neither the One nor the Two. That is, by punishing none of the three culpable wrongdoers, all three are treated equally.

### C. Three Options

The Prosecutor's Dilemma presents three options to a prosecutor: (i) accept the offer and punish only the Two, (ii) not accept the offer and punish only the One, or (iii) not accept the offer and punish none. Analysis of the five central principles of retributivism suggests that the offer, paradoxically, must be both accepted and not accepted. Given the uncertainty regarding which option a retributivist prosecutor should select, let us consider each of the three options separately.

#### 1. Accepting the Offer and Punishing Only the Two

While satisfying some of retributivism's principles with respect to the Two,<sup>244</sup> accepting the offer violates them all with respect to the One. By foregoing punishment of the One, acceptance of the offer violates the following duties: to (i) punish the One, (ii) give the One her just deserts, (iii) honor the One's right to punishment, (iv) treat the One as an end-in-herself, and (v) treat equally the One and the Two.<sup>245</sup> As a result, accepting the offer is not a viable option.

#### 2. Not Accepting the Offer and Punishing Only the One

While satisfying some of retributivism's principles with respect to the One,<sup>246</sup> nonacceptance of the offer (and punishing only the One)

---

and/or legal desert. *See supra* notes 197-205. Desert is based on the culpable commission of moral wrongdoing and/or the culpable violation of a criminal offense. *See supra* note 236. The desert of the One and Two is based on their culpable wrongdoing which constituted a serious criminal offense. It is not based on whether a wrongdoer can be successfully prosecuted independently of, or dependently on, the assistance of a confederate in crime. Similarly, a wrongdoer's willingness to bargain with the prosecutor and inculcate a partner in crime in exchange for immunity does not alter his or her desert. If cooperation did reduce the level of desert, then retributivists would have little basis to oppose plea bargaining and immunity deals. For retributivist criticisms of bargain justice, *see supra* Part I.B.2. Thus, despite some differences that might be pointed out between the respective situations of the One and the Two, the differences are not relevant under retributivism. As a result, in all *relevant* respects, the One and the Two are equally situated under retributivism.

244. By accepting the offer, a retributivist prosecutor could discharge the duty to punish the Two, give the Two their just deserts, satisfy the Two's right to punishment and, treat the Two not as mere means but as ends-in-themselves. *See supra* Part II.B.

245. *See supra* Part II.B.

246. By not accepting the offer and punishing only the One, a retributivist

violates them all with respect to the Two. By foregoing punishment of the Two, not accepting the offer violates the following duties: to (i) punish the Two, (ii) give the Two their just deserts, (iii) honor the Two's right to punishment, (iv) treat the Two as ends-in-themselves, and (v) treat equally the One and the Two.<sup>247</sup> As a result, not accepting the offer (and punishing only the One) is also not a viable option.

### 3. Not Accepting the Offer and Punishing None

Since punishing either only the One or only the Two violates retributivism's principles, perhaps the only viable option remaining is not accepting the offer (and punishing neither the One nor the Two). This option satisfies the principle of treating equally situated culpable wrongdoers equally. That is, by punishing neither the One nor the Two, a retributivist prosecutor treats equally the One and the Two. But punishing none violates, with respect to both the One and the Two, four other retributivist duties: to (i) punish them, (ii) give them their just deserts, (iii) honor their right to punishment, and (iv) treat them as ends-in-themselves.<sup>248</sup> As a result, not accepting the offer and punishing none is also not a viable option.

#### D. Conclusion

This part applied five central principles of retributivism to a hypothetical immunity offer—the Prosecutor's Dilemma. With respect to four of the five principles, each one paradoxically requires, and is violated by, both acceptance and nonacceptance of the offer. Only the fifth principle—the equality principle—can be satisfied without violating itself. As a result, no matter which option is chosen retributivist principles will be violated.

### III. INCONSISTENCY AND MORAL DILEMMA

Applying retributivism to the Prosecutor's Dilemma reveals retributivism's internal inconsistency. The inconsistency is of two types: (i) inconsistency *within* retributivism's principles and, (ii) inconsistency *among*, or between, retributivism's principles. As a result of retributivism's inconsistency, both within and among its principles, retributivism generates a moral dilemma in which any option a retributivist prosecutor selects will be wrong. Finally, this part considers two attempts to restore retributivism to consistency and avoid the Prosecutor's Dilemma.

---

prosecutor satisfies the duty to punish the One, give the One his just deserts, uphold the One's right to punishment and treat the One as an end. See *supra* Part II.B.

247. See *supra* Part II.B.1-5, respectively.

248. See *supra* Part II.B.1-4, respectively.

## A. Inconsistency

### 1. Inconsistency Within Retributivism's Principles

Each of the five retributivist principles applicable to the immunity offer, except the equality principle, are internally inconsistent. Fulfilling each one of these principles requires punishment of both the One and the Two which, in turn, requires both acceptance and nonacceptance of the offer. But this is logically impossible;<sup>249</sup> the offer can only be either accepted or not accepted, but not both. Each one of these principles is violated by acceptance of the offer as well as nonacceptance of the offer. Yet, a retributivist prosecutor is caught in a zero-sum game. The offer can only be accepted or not accepted, it cannot be neither accepted nor not accepted.<sup>250</sup> Each one of the principles paradoxically requires, and is violated by, both accepting the offer and not accepting the offer.<sup>251</sup>

### 2. Inconsistency Among Retributivism's Principles

The fifth principle—the equality principle—is internally consistent,<sup>252</sup> but nevertheless is inconsistent with the other four

249. The joint assertion that the offer must be accepted and that the offer must not be accepted violates the principle of contradiction. The principle of contradiction holds that a “statement and its negation cannot both be true.” Richard Purtill, *Principle of Contradiction*, in *The Cambridge Dictionary of Philosophy* 737, 737 (Robert Audi ed., 2d ed. 1999). For a similar account, see Simon Blackburn, *The Oxford Dictionary of Philosophy* 81 (1996) (“The conjunction of a proposition and its negation [cannot be true].”). The principle of contradiction is alternatively termed “the principle of non-contradiction,” Purtill, *supra*, at 737, or the “law of non-contradiction,” Blackburn, *supra*, at 81. Thus, the statement that the offer must be accepted and the negation of the statement—the offer must not be accepted—constitutes a contradiction which cannot be true. Thus, retributivist principles entail a contradiction, which cannot be true. As a result, retributivism is internally inconsistent. *Id.* (“The standard proof of the inconsistency of a set of propositions or sentences is to show that a contradiction may be derived from them.”).

250. The assertion that the offer must be neither accepted nor not accepted violates the logical principle of excluded middle. The principle, or law, of excluded middle holds that for any statement, either the statement or its negation must be true. See Blackburn, *supra* note 249, at 129 (“[t]he logical law asserting [that for any proposition  $p$ ]  $p$  or not- $p$ ” must be true); Richard Purtill, *Excluded Middle, Principle (or Law) of*, in *The Cambridge Dictionary of Philosophy* 738, 738 (Robert Audi ed., 2d ed. 1999) (“[T]he disjunction of any (significant) statement with its negation is always true.”). Thus, under the law of excluded middle it must be the case that the offer be accepted or not accepted. The statement that the offer must be neither accepted nor not accepted violates the law of excluded middle by denying that the offer must be accepted or not accepted. The following example exemplifies the law of excluded middle: “‘Either there is a tree over 500 feet tall or it is not the case that there is such a tree.’” *Id.* That retributivist principles seem to require that the offer should be neither accepted nor not accepted is similar (in form) to saying that there is neither a tree over 500 feet tall nor is there a tree under 500 feet tall.

251. See *supra* text accompanying notes 196, 207, 212, and 231.

252. The duty to treat equally those culpable wrongdoers who have equal desert is

principles. The equality principle requires that a prosecutor not accept the offer (and punish neither the One nor the Two).<sup>253</sup> In contrast, the other four principles of retributivism require, at the very least, punishment of either the One or the Two. In that way, the four principles may at least be partially satisfied. Punishing only the One at least satisfies the principles with respect to the One; punishing only the Two at least satisfies the principles with respect to the Two. But punishing neither the One nor the Two, as the equality principle requires, completely violates the four principles with respect to both the One and the Two. Thus, satisfaction of the four principles—to the greatest extent possible—requires punishment of (at least) either the One or the Two. But doing so entails treating the One and the Two differently. Inconsistency among retributivism's principles exists because fulfillment of the four principles (to the greatest extent possible) requires unequal treatment of the One and the Two, but the equality principle requires their equal treatment.

### B. Moral Dilemma

So what is a retributivist prosecutor to do when faced with the offer in the Prosecutor's Dilemma? As a result of retributivism's internal inconsistency, any option a retributivist prosecutor selects violates several retributivist principles. Thus, a retributivist prosecutor's choice constitutes a moral dilemma.<sup>254</sup> Because different types of moral dilemmas pose difficulties to a varying degree, understanding the gravity of the Prosecutor's Dilemma for retributivism requires situating it within the typology of moral dilemmas.

Narrowly defined moral dilemmas are more serious than those that are broadly defined.<sup>255</sup> Perhaps the broadest, and thus least serious, conception is any situation in which moral reasons conflict (whether with other moral reasons or nonmoral reasons<sup>256</sup>), or in which it is not

---

internally consistent as applied to the Prosecutor's Dilemma because it can be fulfilled without requiring incompatible courses of action. It can be fulfilled by one of the options—not accepting the offer and punishing neither the One nor the Two. That option fulfills the equality principle by treating the One and Two equally (or, rather, not treating them unequally).

253. See *supra* Part II.B.5.

254. See Christopher W. Gowans, *Innocence Lost: An Examination of Inescapable Moral Wrongdoing* 4 (1994) (defining moral dilemmas, in general, as "situations in which an agent morally ought to (and can) take one course of action and morally ought to (and can) take another course of action, even though the agent cannot take both courses of action").

255. See Walter Sinnott-Armstrong, *Moral Dilemma*, in *The Cambridge Dictionary of Philosophy* 584, 584 (Robert Audi ed., 2d ed. 1999) (supplying seven different, on a continuum from broadest to narrowest, conceptions of what constitutes a moral dilemma).

256. Arguably the prisoner's dilemma could be forced into this category. If we consider one prisoner betraying his confederate so as to obtain a lighter sentence to be a moral wrong, then the prisoner faces a moral dilemma of this broad variety. The

known what, if anything, is the morally right thing to do.<sup>257</sup> A comparatively narrower conception is when “an agent morally *ought* to do each of two acts, but cannot do both.”<sup>258</sup> In a well-known example of this type, from the existentialist philosopher Jean-Paul Sartre, a student during World War Two believes that he ought to go abroad and fight the Nazis but at the same time believes that he ought to stay home to care for his mother, but cannot do both.<sup>259</sup> This conception is still too broad to pose serious problems in that the “ought” refers to ideal actions (as in one ought to donate money to charities<sup>260</sup>) and not moral obligations.<sup>261</sup> The novel *Sophie’s Choice*<sup>262</sup> supplies perhaps the best-known recent example of a still narrower moral dilemma. A Nazi doctor in a World War Two concentration camp informs Sophie that she must choose which *one* of her children, Jan or Eva, to send to a gas chamber.<sup>263</sup> Although Sophie has an obligation to save Jan and Eva, if Sophie fails to choose,<sup>264</sup> the Nazi doctor will send both Jan and Eva to die.<sup>265</sup>

The relationship of the moral principles or obligations in conflict, and whether or not the choice of incompatible actions arises through the fault of the agent, also serve to distinguish moral dilemmas. In “symmetrical”<sup>266</sup> or equal<sup>267</sup> dilemmas, the choice of incompatible

prisoner must choose between loyalty to his confederate and self-interest. But perhaps confessing and aiding the state in the prosecution of crime—even if it means betraying one’s confederate—is the morally preferable course of action. In any event, the prisoner’s dilemma is not conventionally viewed as a moral dilemma, but rather as a dilemma of rationality or as a dilemma of self-interest versus collective interest. See Baird, *supra* note 25, at 312-13.

257. Sinnott-Armstrong, *supra* note 255, at 584.

258. *Id.*

259. Jean-Paul Sartre, *Existentialism is a Humanism*, in *Existentialism* (1947), reprinted in *Reason and Responsibility* 431, 435 (Joel Feinberg ed., 5th ed. 1981). The recent war with Iraq has generated a similar example of tragic choice. Prior to departing to Kuwait to join his troops, a U.S. commander faced the choice of staying home to care for his infant who needed a heart transplant or going to Kuwait to prepare and train his troops. *NBC Evening News with Tom Brokaw* (NBC television broadcast, Mar. 6, 2003).

260. Although one (ideally) ought to donate money to charities, the failure to do so does not violate a moral obligation. Sinnott-Armstrong, *supra* note 255, at 584.

261. *Id.*

262. William Styron, *Sophie’s Choice* (1980).

263. *Id.* at 589.

264. The reason for choosing to save Jan is equally applicable as a basis for choosing to save Eva. So, there seems to be no basis to choose one over the other. But if Sophie fails to choose, neither will be saved. While consequentialism would endorse choosing either child as opposed to making no choice and both being killed, some argue that the morally preferable course of action is to refuse to choose so as to avoid cooperating “in the evil project of another.” Gowans, *supra* note 254, at 62 n.7.

265. Styron, *supra* note 262, at 589.

266. Walter Sinnott-Armstrong, *Moral Dilemmas* 54 (1988).

267. Christopher Gowans refers to symmetrical dilemmas as “the equality argument.” Gowans, *supra* note 254, at 50.

actions is generated by the same obligation(s).<sup>268</sup> For example, in Sophie's choice, the twin incompatible duties to save Jan and to save Eva stem from the same obligation—saving her children.<sup>269</sup> In contrast, different principles giving rise to the duty to perform incompatible actions generate “incomparable” dilemmas.<sup>270</sup> For example, the different principles of aiding one's country and caring for one's mother, neither of which overrides the other, are incomparable and generate the dilemma for Sartre's student.<sup>271</sup> A dilemma which arises through no fault of the agent confronted with the dilemma is a “perplexity *simpliciter*.”<sup>272</sup> The dilemmas facing both Sophie and Sartre's student are examples of this type.<sup>273</sup> In contrast, a dilemma arising through the fault of the agent constitutes a “perplexity *secundum quid*.”<sup>274</sup> The typical, if prosaic, example of the latter type is when one makes incompatible promises. One is morally obliged to keep both promises, at least under Kantian ethics,<sup>275</sup> but one cannot because they are incompatible. The resulting dilemma was avoidable and arose only through the fault of the promisor.<sup>276</sup> A dilemma

268. In dilemmas in which the moral obligations or principles in conflict are equal or symmetrical, the same moral obligation or obligations require two (or more) incompatible courses of action (or non-action). *Id.* at 49-52; Sinnott-Armstrong, *supra* note 266, at 54-58. For an example of an equal or symmetrical dilemma, see *supra* text accompanying notes 262-63.

269. See *supra* note 263 and accompanying text.

270. Sinnott-Armstrong, *supra* note 266, at 58. Gowans refers to incomparable dilemmas as “The Incomparability Argument.” Gowans, *supra* note 254, at 52. In dilemmas in which the moral obligations are not equal or symmetrical, but rather “incomparable,” two (or more) different moral obligations, neither of which overrides the other, require two (or more) incompatible courses of action (or non-action). *Id.* at 52-57; Sinnott-Armstrong, *supra* note 266, at 58-60.

271. See *supra* text accompanying note 259.

272. Gowans, *supra* note 254, at 222; Sinnott-Armstrong, *supra* note 266, at 102. In a perplexity *simpliciter*, or more simply a dilemma *simpliciter*, the dilemma arises through no fault of the actor confronted with being obliged to perform two (or more) incompatible courses of action (or non-action). The actor did not commit avoidable wrongdoing which precipitated the dilemma. Gowans, *supra* note 254, at 221-22; Sinnott-Armstrong, *supra* note 266, at 102-06.

273. The dilemma facing Sophie arises, not through any fault of her own, but because of the Nazi doctor's order compelling her to choose. This choice is made under penalty of both children being killed if she failed to choose which child would be saved and which would not. Sartre's student is also not responsible for his dilemma. It is through no fault of his own that his country needs him to fight the Nazis and it is also not his fault that his mother needs his care.

274. Gowans, *supra* note 254, at 222; Sinnott-Armstrong, *supra* note 266, at 102. In a perplexity *secundum quid*, or more simply a dilemma *secundum quid*, the dilemma facing an agent arises because of the agent's prior, avoidable wrongdoing. Gowans, *supra* note 254, at 221-22; Sinnott-Armstrong, *supra* note 266, at 102-06.

275. For the view that Kant held that keeping one's promises was an important moral obligation, see, for example, H.J. Paton, *The Categorical Imperative: A Study in Kant's Moral Philosophy* 152-54 (1948); Sullivan, *supra* note 134, at 167-170.

276. Thomas Nagel, *Mortal Questions* 53, 74 (1979) (citing the making of incompatible promises as an example of a dilemma, or perplexity, *secundum quid* in which the dilemma arises through the fault of the actor).



*simpliciter* poses comparatively greater difficulties than a dilemma *secundum quid*.<sup>277</sup>

The Prosecutor's Dilemma is of the most serious type and particularly thorny. It satisfies the narrowest conception of a moral dilemma: "a situation where an agent has a *strong* moral obligation or requirement to adopt each of two alternatives, and neither is overridden, but the agent cannot adopt both alternatives."<sup>278</sup> And by containing both a symmetrical and an incomparable dilemma,<sup>279</sup> the Prosecutor's Dilemma is not merely a dilemma but a trilemma.<sup>280</sup>

277. Alan Donagan recognizes that dilemmas *secundum quid* are quite possible, but finds dilemmas *simpliciter* to be so troubling as to declare that such dilemmas do not exist. Alan Donagan, *The Theory of Morality* 144-45 (1977); Alan Donagan, *Moral Dilemmas, Genuine and Spurious: A Comparative Anatomy*, in *Moral Dilemmas and Moral Theory* 12, 13 (H.I. Mason ed., 1996) [hereinafter Donagan, *Moral Dilemmas, Genuine and Spurious*]; Alan Donagan, *Consistency in Rationalist Moral Systems*, 81 *J. of Phil.* 291, 305-06 (1984) [hereinafter Donagan, *Consistency*]. Although few state the case as strongly as Donagan, there is a widespread, albeit not uniform, belief among moral dilemma scholars that dilemmas, or perplexities, *simpliciter* create considerably greater difficulties for a moral theory giving rise to them than dilemmas, or perplexities, *secundum quid*. Gowans, *supra* note 254, at 221. Noting that an innocent person is in a situation in which there is no choice but to commit wrong, Gowans explains the comparatively greater gravity of a dilemma *simpliciter*:

The ideal of moral innocence explains the thought that there is an important difference between the two situations. Cases of perplexity *secundum quid* are brought about by oneself, they are not imposed by the world. When I cannot escape wrongdoing on account of previous wrongdoing I have only myself to blame. I could have acted in such a way that wrongdoing would not now be inescapable. Innocence, a life free from all wrongdoing, was a possibility for me. But I chose another path. All this seems quite compatible with the ideal of innocence, and so from this standpoint is unproblematic. On the other hand, cases of perplexity *simpliciter* are imposed by the world. Here wrongdoing is unavoidable, not because I have already done something wrong, but because of what happens to me. . . . If we are committed to this ideal [of innocence], "our intuitions rebel" at the idea of perplexity *simpliciter* in a way that they do not rebel at the idea of perplexity *secundum quid*.

*Id.* at 221-22 (quoting Nagel, *supra* note 276, at 74 (in contrast to dilemmas *secundum quid*, with respect to dilemmas *simpliciter* "[o]ur intuitions rebel" with the prospect that "the world itself . . . could face a previously innocent person with a choice between morally abominable courses of action.")).

278. Sinnott-Armstrong, *supra* note 255, at 584.

279. The Prosecutor's Dilemma combines an equal or symmetrical moral dilemma with an incomparable moral dilemma. In the symmetrical portion of the dilemma, four principles of retributivism require a prosecutor to accept the offer and punish only the Two. But the same four principles also require the prosecutor not to accept the offer and punish only the One. This portion of the dilemma is symmetrical or equal because it is the same principles which require the incompatible courses of action (or non-action). The incomparable portion of the Prosecutor's Dilemma is the conflict between the equality principle and the four principles which require incompatible courses of action (or non-action). The equality principle differs from the four principles; however, they are neither greater, weaker nor equal to the other—they are incomparable.

280. See Blackburn, *supra* note 249, at 105. Trilemmas are also referred to as

That is, a retributivist prosecutor has a strong moral obligation to adopt each of *three* alternatives—none of which are overridden—but the prosecutor cannot adopt all three (or even two) of the alternatives.<sup>281</sup> Furthermore, since the Prosecutor's Dilemma arises through no fault of the prosecutor, it constitutes the more troubling dilemma *simpliciter*.<sup>282</sup>

Theories generating serious moral dilemmas are widely, albeit not uniformly,<sup>283</sup> considered inadequate,<sup>284</sup> inconsistent,<sup>285</sup> false,<sup>286</sup> indeterminate,<sup>287</sup> and irrational.<sup>288</sup> Regardless of the differing

three-horned or three-pronged dilemmas. Julian Baggini & Peter S. Fosl, *The Philosopher's Toolkit: A Compendium of Philosophical Concepts and Methods* 93 (2003) (explaining that a dilemmatic choice situation may be described by reference to the number of incompatible choices which are referred to as horns or prongs).

281. Four retributivist principles require that only the One be punished; the same four principles require that only the Two be punished; and, the equality principle requires that neither the One nor the Two be punished. Under the terms of the Prosecutor's Dilemma, neither all three nor even two of the alternatives can be adopted by a retributivist prosecutor. Only one of the three alternatives can be adopted: (i) accept the offer and punish the Two, (ii) not accept the offer and punish the One, or (iii) not accept the offer and punish neither the One nor the Two. And no matter which alternative is chosen, strong moral obligations of retributivism will be violated.

282. Perhaps this is because a dilemma *secundum quid*, arising through the fault of the agent confronted with the dilemma, see *supra* notes 274-76 and accompanying text, may be blamed on the agent and not the theory generating it. In contrast, a dilemma *simpliciter*, arising through no fault of the agent, see *supra* notes 272-73, may only be blamed on the theory generating it. For other explanations, see *supra* note 277 and accompanying text.

283. Curiously, some commentators suggest that *perhaps* theories which generate moral dilemmas are not necessarily inferior to those which fail to produce them. *E.g.*, Sinnott-Armstrong, *supra* note 266, at 181-82 (discussing the claim that the existence of moral dilemmas is valuable in that they motivate us to avoid the situations in which they arise); Ruth Marcus, *More About Moral Dilemmas*, in *Moral Dilemmas and Moral Theory* 23, 28 (H.I. Mason ed., 1996) (the situations giving rise to moral dilemmas should be avoided and the occurrence of moral dilemmas are helpful by identifying the ways in which the world requires change to prevent their occurrence). For a critique of this view, see Terrance C. McConnell, *Moral Residue and Dilemmas*, in *Moral Dilemmas and Moral Theory* 36, 44 (H.I. Mason ed., 1996) (agents would still be motivated to avoid situations which generate moral dilemmas even if moral dilemmas were resolved). A further difficulty of this view is that it seems naive to suppose that agents can, in fact, with sufficient motivation, avoid all moral dilemmas. This might be true with dilemmas *secundum quid* which arise only through the fault of the agent, but is not true with respect to dilemmas *simpliciter*. Was it lack of motivation that precluded Sophie and Sartre's student from being able to avoid situations, like World War Two, that engendered their respective dilemmas? Should prosecutors avoid the prosecution of criminals so as to avoid dilemmatic situations like the Prosecutor's Dilemma or just avoid embracing retributivist principles?

284. Sinnott-Armstrong, *supra* note 266, at 169; McConnell, *supra* note 283, at 36.

285. Sinnott-Armstrong, *supra* note 266, at 169; Donagan, *Moral Dilemmas, Genuine and Spurious*, *supra* note 277, at 13 (regarding dilemmas *simpliciter*).

286. Donagan, *Moral Dilemmas, Genuine and Spurious*, *supra* note 277, at 13.

287. Thomas E. Hill Jr., *Moral Dilemmas, Gaps, and Residues: A Kantian Perspective*, in *Moral Dilemmas and Moral Theory* 167, 176-83 (H.I. Mason ed., 1996) (explaining that moral dilemmas create gaps and indeterminacies in moral theories in

conceptions of the significance of moral dilemmas for the theories which generate them, nearly all agree that either avoiding or resolving serious moral dilemmas is paramount.<sup>289</sup>

### C. Attempts to Avoid the Prosecutor's Dilemma

This section presents two attempts to avoid the Prosecutor's Dilemma.<sup>290</sup> First, a retributivist might choose to declare bargain offers as unjust or immoral, and refuse to participate in them. Second, narrowing the scope of the duty to punish culpable wrongdoers to only those who have been convicted might preclude the dilemma from arising. However, neither of these attempts successfully avoids the Prosecutor's Dilemma.

#### 1. Refusal to Participate in Bargain Justice Because It Is Unjust

Moore suggests that perhaps a "retributivist might simply deny the propriety of the practice [of bargain justice]."<sup>291</sup> As such, a retributivist prosecutor might simply refuse to participate in bargain justice. A retributivist system that neither initiated nor even entertained bargain offers, one might argue, avoids the grip of the Prosecutor's Dilemma.

---

which they fail to provide guidance to an agent as to the morally preferable course of action). *See id.* at 176 (noting the importance for Kant that morality always provide guidance as to the proper course of action).

288. *Id.* at 173 (interpreting Kant as maintaining that morality is based on reason and that a rational moral system cannot contain serious, unresolvable moral dilemmas).

289. *E.g.*, Donagan, *Moral Dilemmas, Genuine and Spurious*, *supra* note 277, at 15 (rationalist moral theories, like those of Kant and St. Thomas Aquinas, cannot allow moral dilemmas to stand); Hill, *supra* note 287, at 182 (noting that "what we would need to do [is] use all permissible means to prevent those situations [moral dilemmas] from arising"); Marcus, *supra* note 283, at 23-28 (as rational agents we must avoid the occurrence of moral dilemmas).

290. For other attempts to avoid the Prosecutor's Dilemma, see *infra* Part IV. For attempts to not merely avoid, but rather resolve the Prosecutor's Dilemma, see *infra* Part V.

291. Moore, *supra* note 20, at 158. Here Moore is referring to a general, or deontological, retributivist, rather than a consequentialist retributivist. For the distinction between the two types of retributivist, see *infra* notes 310, 312 and accompanying text. The reason why a (deontological) retributivist might simply reject the legitimacy of bargain justice and refuse to participate in it is that, according to Moore,

it might be thought that we rightly refuse to punish some guilty persons in order to be able to punish other, more seriously guilty persons—as when we give immunity, in order to extract testimony needed to convict the latter. How can the retributivist accommodate these practices, given that the retributivist regards the punishment of the guilty to be categorically imperative whenever the opportunity to give such punishment presents itself?

*Id.* at 156.

But denying the “propriety”<sup>292</sup> of such bargains does not allow retributivism to escape. Of course, a retributivist prosecutor could (and must) refrain from tendering offers of bargain justice. But that fails to preclude culpable wrongdoers from initiating such offers. And once the offer is made, it must be either accepted or not. Refusing to participate in, or even ignoring, offers of bargain justice nonetheless constitutes not accepting the offer. And by nonacceptance, a retributivist prosecutor is foregoing punishment of culpable wrongdoers where there is an opportunity to punish them.<sup>293</sup> Nonacceptance of the offer in the Prosecutor’s Dilemma violates, at least, four central retributivist principles.<sup>294</sup> As a result, a retributivist system cannot avoid the Prosecutor’s Dilemma by declaring that offers of bargain justice are unjust, and refusing to participate in them.

## 2. Duty to Punish Limited to Convicted Offenders

The scope of retributivism’s duty to punish (as well as its other principles), one might argue, does not extend to all culpable wrongdoers, but only to convicted (including those who plead guilty) culpable wrongdoers. Since neither the One nor the Two are convicted culpable wrongdoers, neither acceptance nor nonacceptance of the offer violates such a narrowed duty to punish. Thus, limiting the duty to punish, one might argue, avoids the Prosecutor’s Dilemma.

Limiting the scope of the duty to punish in this way, however, does not satisfactorily avoid the Prosecutor’s Dilemma. First, the argument proves too much. Such a limited duty to punish reduces to a nullity. No culpable wrongdoer, in a bargain justice situation or not, would ever need to be prosecuted. Without a concomitant duty to prosecute cases of *suspected* culpable wrongdoing, retributivism’s limited duty to punish might never need to be actually carried out.<sup>295</sup> This would frustrate retributivism itself. As Moore explains, in construing the duty to punish: “Retributivism, in other words, is truly a theory of justice such that, if it is true, we have an obligation to set up

292. *Id.* at 156.

293. By not accepting the offer, a prosecutor is foregoing the opportunity to punish the Two—the Two can only be punished if the offer is accepted. The tender of the One’s offer to inculcate the Two in return for immunity supplies the opportunity to punish the Two.

294. Nonacceptance of the offer entails two distinct options. Nonacceptance of the offer (and punishing only the One) violates five duties: the duty to punish the Two, give the Two their just deserts, honor the Two’s right to be punished, affirmatively treat the Two as ends, and treat the One and the Two equally. The other option, not accepting the offer (and punishing neither the One nor the Two), violates the former four duties. See *supra* Part II.C.2-3.

295. Until prosecuted and convicted, no culpable wrongdoer deserves punishment inflicted by the state. But without a duty to prosecute, there might not be any convicted culpable wrongdoers. And with no convicted culpable wrongdoers, there would be no one to punish. The limited duty to punish thereby reduces itself to a nullity.

institutions so that retribution is achieved.<sup>296</sup> Since retribution cannot be attained without punishment and punishment cannot be attained without prosecution,<sup>297</sup> an obligation to attain retribution entails an obligation to punish, as well as an obligation to prosecute.

Second, even assuming *arguendo* that the limitation is tenable, it fails to avoid the following variant of the Prosecutor's Dilemma.<sup>298</sup> Suppose that three culpable wrongdoers, who together committed serious wrongdoing, are apprehended, tried and convicted. After their conviction, but before their punishment, they escape and elude capture by the authorities. After an extensive search for a number of years, the authorities give up. Though safe from the reach of the law, one of the three no longer wishes to live in hiding, turns herself in and offers to disclose the location of, and inculcate, the other two in return for immunity for herself.<sup>299</sup>

Even if the duty to punish is limited to only those convicted, and even if the duty did not apply in the Prosecutor's Dilemma, in this variation the duty would apply. Here, all three are convicted culpable wrongdoers. And the variation poses the same difficulties for retributivism as the original Prosecutor's Dilemma.<sup>300</sup> Thus, even if

296. Moore, *supra* note 188, at 182; *see also* Benn, *supra* note 198, at 30 (for retributivism "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"). For an alternative grounding of the duty to prosecute under retributivism, *see* Fletcher, *supra* note 21, at 60-63.

297. If there was no duty to prosecute, the obligation to attain retribution could not be satisfied.

298. This is the difficulty with unprincipled, ad hoc arguments. By attacking an incidental feature of a problem, but not its root, they are typically subject to counterexamples. By simply eliminating the incidental feature, the problem recurs. *Cf.* Lawrence Lessig, *Social Meaning and Social Norms*, 144 U. Pa. L. Rev. 2181, 2181 n.1 (1996) (terming such arguments as "lawyers' [unprincipled] ad hocery").

299. In yet another variation, we might suppose that after conviction, but before incarceration, the three culpable wrongdoers escape to a country which does not have an extradition treaty with the United States. Though safe from the reach of the law, one of the three gets homesick, returns to America, and turns herself in to the authorities. She offers to lure the other two back into the United States and disclose their location so that they can be apprehended in return for immunity for herself. Since all three are convicted culpable wrongdoers, there is a duty to punish them even under the proposed limited duty to punish. Acceptance of the offer affords punishment of two of the three, but precludes punishment of the offeree. Nonacceptance of the offer affords punishment of the offeree, but precludes punishment of the other two. Should the offer be accepted or not? Since there are incompatible opportunities to punish all three, and all three are convicted, a retributivist prosecutor will violate the limited duty to punish no matter what option is chosen. Thus, even if limiting the duty to punish only to convicted culpable wrongdoers avoids the Prosecutor's Dilemma in its original form, it cannot avoid variants.

300. Acceptance of the offer violates all five retributivist duties with respect to the offeree. Nonacceptance of the offer (and punishing only the offeree) violates all five retributivist duties with respect to the other two culpable wrongdoers. Nonacceptance of the offer (and punishing none) violates all of the retributivist duties except the equality principle. Any option a retributivist prosecutor selects violates at

limiting the duty to punish avoids the original Prosecutor's Dilemma, it does not avoid variants of the Prosecutor's Dilemma.

#### D. Conclusion

The inconsistency among, and within, retributivism's principles reveals retributivism to be internally inconsistent. Four of retributivism's central principles require nonacceptance of the offer (and punishment of only the One); but the same four principles also require acceptance of the offer (and punishment of only the Two). A fifth principle—the equality principle—requires nonacceptance of the offer (and punishment of neither the One nor the Two). As a result, retributivism generates a serious moral dilemma. When faced with the offer, a retributivist prosecutor has a strong moral obligation to adopt each of three alternatives, none of which are overridden, but the prosecutor cannot adopt all three (or even two) of the alternatives. No matter what a retributivist prosecutor does in the Prosecutor's Dilemma, it will be wrong. Consideration of two attempts to avoid the Prosecutor's Dilemma failed to yield a satisfactory solution.<sup>301</sup>

#### IV. ATTEMPTS TO RECONCILE RETRIBUTIVISM WITH BARGAIN JUSTICE

This part considers two controversial forms of retributivism advanced, in part, as an effort to make retributivism compatible with bargain justice. Reconciling retributivism with bargain justice is an important project for retributivists because unless retributivism can accommodate the mode in which up to ninety percent of punishments are imposed,<sup>302</sup> retributivism risks irrelevance as a *descriptive* theory of punishment—one which can explain what legal institutions of punishment actually do. As evidenced by the few retributivists to undertake this project, retributivists are apparently willing to stand pat with retributivism as largely a *prescriptive* theory—one which prescribes what legal institutions of punishment should do by delineating the conditions for morally just punishment. But as seen in

---

least four of retributivism's duties.

301. Another perhaps obvious way to attempt to avoid the Prosecutor's Dilemma is by invoking Kant's dictum that ought implies can. *See infra* note 377. That is, one is morally obligated to do only that which one can do. Applying it to the Prosecutor's Dilemma, a retributivist prosecutor is only obligated to do that which she can do. Since she cannot punish both the One and the Two, she is not obligated by retributivism's principles to punish both the One and the Two. But she can punish either the One or the Two. And thus she is still obligated, even under Kant's dictum, to punish either the One or the Two. The problem, however, just as in Sophie's choice, *see supra* notes 262-65 and accompanying text, is in deciding (nonarbitrarily) which one of the two incompatible courses of action to adopt.

302. *See supra* note 11 and accompanying text.

the previous part, bargain justice, in the form of the Prosecutor's Dilemma, poses serious difficulties for retributivism even as a merely prescriptive theory.<sup>303</sup>

Since bargain justice poses a substantial threat to retributivism in both its descriptive and prescriptive capacities, then retributivists might profitably shift from the conventional form of retributivism<sup>304</sup> to some new form which is not only compatible with bargain justice but also avoids the Prosecutor's Dilemma. The two alternative versions of retributivism—"consequentialist' retributivism"<sup>305</sup> and threshold retributivism<sup>306</sup>—attempt to incorporate the consequences of punishment into a retributivist framework. Although conventional retributivism eschews (and retributivists criticize) the use of consequences to justify punishment,<sup>307</sup> the utilization of consequences allows these new versions of retributivism to be comparatively more compatible with bargain justice, thereby accruing greater descriptive power. This part demonstrates, however, that retributivists' embrace of consequences not only further undermines retributivism's prescriptive power, but also fails to avoid the Prosecutor's Dilemma.

### A. Consequentialist Retributivism

"Consequentialist' retributivism,"<sup>308</sup> might supply a clear answer to the Prosecutor's Dilemma—accept the offer.<sup>309</sup> Before considering whether consequentialist retributivism satisfactorily avoids the Prosecutor's Dilemma, a brief account and critique of this alternative version of retributivism is necessary.

#### 1. Deontological v. Consequentialist Retributivism

The conventional conception of retributivism, which Moore terms "deontological retributivism,"<sup>310</sup> demands the punishment of each and

303. It reveals the just conditions of punishment prescribed by retributivism to be internally inconsistent. *See supra* Part III.A.

304. For a discussion of the conventional form of retributivism, see *supra* Parts I.B.1 and II.B.

305. *See infra* Part IV.A.

306. *See infra* Part IV.B.

307. *See* Duff, *supra* note 6, at 2, 186 (explaining that the moral defects of consequentialism are due to its utilization of contingent consequences to justify punishment).

308. Moore, *supra* note 20, at 155.

309. According to Moore, the problems for retributivism in justifying the acceptance of offers of bargain justice might be alleviated under a consequentialist form of retributivism. *Id.* at 156 ("[I]f a retributivist may be a consequentialist, these are obviously non-problems for his theory of punishment.")

310. Moore distinguishes "deontological versus consequentialist retributivism," *id.* at 159, as follows:

The 'deontological' or 'agent-relative' retributivist regards the act of punishing the guilty as categorically demanded on each occasion, considered separately; the 'consequentialist' or 'agent-neutral' retributivist regards the

every guilty person even if doing so foregoes greater opportunities to punish the guilty.<sup>311</sup> In contrast, consequentialist retributivism seeks to maximize the number of culpable wrongdoers punished even if doing so foregoes punishment of some culpable wrongdoers.<sup>312</sup> Although both versions of retributivism value a culpable wrongdoer receiving his just deserts as an intrinsic good,<sup>313</sup> only consequentialist

state of the guilty receiving punishment as a good state to be maximized even when this means that some guilty persons are intentionally allowed to escape punishment.

*Id.* at 156. Moore concludes that “both are recognizable versions of retributivism.”

*Id.* at 159. Moore stated, deontology is a view in which morally “right action is . . . independent . . . of goodness . . . . Some principles of right and wrong, notably principles of justice and honesty, prescribe actions even though more evil than good would result from doing them.” John Deigh, *Ethics*, in *The Cambridge Dictionary of Philosophy* 284, 287 (Robert Audi ed., 2d ed. 1999). John Rawls supplies a concise account of deontology by contrasting it with teleology:

[A] deontological theory . . . [is] one that either does not specify the good independently from the right, or does not interpret the right as maximizing the good. (It should be noted that deontological theories are defined as non-teleological ones, not as views that characterize the rightness of institutions and acts independently from their consequences.)

John Rawls, *A Theory of Justice* 30 (1971). At the risk of oversimplification, deontology is concerned with the right more so than with the good, and is generally associated with a Kantian view of morality as well as retributivism. Teleology is comparatively more concerned with the good than the right and is typically associated with consequentialism and utilitarianism. See Nancy (Ann) Davis, *Contemporary Deontology*, in *A Companion to Ethics* 205, 205 (Peter Singer ed., 1991).

311. Moore, *supra* note 20, at 156.

312. *Id.* Moore illustrates the distinction between these two forms of retributivism by the following example:

Take Kant's old example of the last murderer in an island society that is about to disband and leave its island. A retributivist believes that the murderer should be punished because he deserves it, even though no other good will thereby be achieved. The guilty, in other words, must be punished. If one construes this retributivist norm as an agent-relative norm [deontological retributivism], then the members of Kant's island society are obligated to punish the last murderer even if by doing so other murderers elsewhere will go unpunished. (We might imagine, for example, that if we punish this murderer, he will not testify against other guilty murderers who will thereby escape conviction and punishment.) Alternatively, if one construes this retributivist norm as a norm describing a state of affairs that is to be maximized, then the members of Kant's island society should *not* punish the last murderer but should maximize the punishment of the guilty by their actions.

*Id.* at 155-56 (citation omitted).

313. Moore asserts that both forms of retributivism share that which is distinctive to retributivism in general:

[W]hat is distinctively retributivist is the view that the guilty receiving their just deserts is an intrinsic good. It is, in other words, not an instrumental good—good because such punishment causes other states of affairs to exist that are good. Even if punishing the guilty were without any further effects, it would be a good state to bring about, on this intrinsic view of punishing the guilty.

*Id.* at 157.



retributivism seeks to maximize the generation of that intrinsic good.<sup>314</sup>

## 2. Untenable Form of Retributivism

In perhaps the only assessment of Moore's consequentialist retributivism, David Dolinko argues that "to 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)."<sup>315</sup> According to Moore, under retributivism, "the beneficial consequences [of punishment] other than giving offenders their just deserts . . . [must] form no part of what makes punishment just."<sup>316</sup> But Dolinko contends that a consequentialist, or even a consequentialist-retributivist, must necessarily view offenders receiving their just deserts as only one among many good consequences to promote.<sup>317</sup> That is, a consequentialist or consequential-retributivist would necessarily recognize a plurality of good consequences or good states of affairs to

---

314. Moore argues that both forms of retributivism are valid and that some of the purported problems of deontological retributivism are not sufficient to warrant a shift to consequentialist retributivism. See *supra* notes 309-10 and accompanying text. But as Moore concedes, "[t]he same cannot be said about the intentional forgoing of any opportunity to punish a guilty offender in order to obtain the conviction and punishment of an even more guilty offender, which is why this common prosecutorial practice is more of a problem for the deontological version of retributivism." Moore, *supra* note 20, at 158. Though Moore acknowledges that deontological retributivism has difficulty accommodating bargain justice, nonetheless he asserts that a deontological retributivist still has a number of moves: "[T]he deontological retributivist might simply deny the propriety of the practice. More plausibly, if he is a 'threshold deontologist', as am I, he might more qualifiedly disavow the practice except when it is needed to punish some very deserving criminals." *Id.* (citation omitted). For a discussion as to why the former option does not provide a satisfactory solution to the difficulties sustained by retributivism as applied to offers of bargain justice, see *supra* Part III.C.1. For a discussion of why Moore's latter solution is unsatisfactory, see *infra* Part IV.B.3.

315. David Dolinko, *Retributivism, Consequentialism, and the Intrinsic Goodness of Punishment*, 16 *Law & Phil.* 507, 513 (1997).

316. Moore, *supra* note 20, at 153.

317. Otherwise, according to Dolinko, an absurdity would result. That is, if consequentialists or consequential retributivists recognized only one good consequence or state of affairs—offenders receiving their just deserts—the following absurdity would ensue:

What consequentialist, for example, would rank a state of affairs in which all guilty persons receive their just deserts but everyone else dies painfully by age 30 as better than one in which only one guilty person escapes deserved punishment and all non-criminals lead long and happy lives? Indeed, to insist that *only* the quantity of "the guilty receiving punishment" affects the goodness of a state of affairs implies the absurd conclusion that a state of affairs wherein no one ever commits any crime at all lacks goodness altogether.

Dolinko, *supra* note 315, at 514.

be promoted.<sup>318</sup> Since Moore's own requirement for retributivism insists on the relevance of only one beneficial consequence in justifying punishment, but consequentialism necessarily considers a plurality of good consequences, Dolinko concludes that "a consequentialist version of retributivism appears impossible."<sup>319</sup>

### 3. Applicability to Offer

Assuming *arguendo* that consequentialist retributivism is tenable, it might avoid the Prosecutor's Dilemma. Since punishing the Two maximizes the intrinsic good of retributive punishment, consequentialist retributivism provides a clear basis to accept the offer.<sup>320</sup> And because consequentialist retributivism does not feature the duty to punish a culpable wrongdoer on each occasion the opportunity presents itself,<sup>321</sup> there is no inconsistency within the duty to punish.

But accepting the offer, even under consequentialist retributivism, still yields conflicts among retributivism's central principles. By foregoing punishment of the One, acceptance of the offer fails to give the One her just deserts, violates the One's right to punishment, uses the One as a mere means, and treats the One and the Two unequally.<sup>322</sup> One might argue, however, that these central principles of retributivism, which conflict with the duty to maximize retributive punishment, are not principles of *consequentialist* retributivism. A version of retributivism lacking these principles eliminates the possibility of their conflict,<sup>323</sup> with no inconsistent principles, retributivism, albeit the consequentialist version, is restored to consistency.

Consequentialist retributivism attains consistency, however, at a heavy price. Its justification of the unequal treatment of equally culpable wrongdoers may be too much even for consequentialist

318. *Id.* ("The consequentialist-retributivist, then, presumably believes that there exist a plurality of distinct 'good-making' features of states of affairs, of which the guilty receiving deserved punishment is simply one.")

319. *Id.* at 515; *see id.* (stating that consequentialist retributivism is incoherent "if we insist [as Moore does] that the meting out of just deserts is the only beneficial consequence of punishment relevant to its justification").

320. Moore, *supra* note 20, at 157 ("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.")

321. *Id.* ("[T]he consequentialist-retributivist will intentionally refuse to punish guilty persons whenever more guilty persons (or greater guilt) will be punished thereby.")

322. Thus consequentialist retributivism's duty to maximize retributive punishment conflicts with other central values of retributivism.

323. If these principles are not constituent elements of consequentialist retributivism, then, of course, they can no longer be in conflict. The simple but effective tack of consequentialist retributivism might be simply to eliminate the inconsistent principles.

retributivism's proponent, Michael Moore.<sup>324</sup> By declaring that "[l]ike cases should indeed be treated alike,"<sup>325</sup> and criticizing another punishment theory for violating the principle of equality,<sup>326</sup> Moore himself is apparently unwilling to embrace consequentialist retributivism's unequal treatment of equally culpable wrongdoers. Thus, even a consequentialist retributivist might not be willing to abandon the equality principle. If so, accepting the offer creates a conflict within consequentialist retributivism—between the duty to maximize retributive punishment and treating equally culpable wrongdoers equally.<sup>327</sup>

Furthermore, presumably even consequentialist retributivists would be loath to abandon the duty to treat offenders as ends and not as mere means.<sup>328</sup> Elimination of the duty, to avoid inconsistency,<sup>329</sup> renders consequentialist retributivism subject to the same criticisms used to condemn consequentialism.<sup>330</sup> Retributivists themselves condemn consequentialism for this very failing—using persons as mere means—terming it the "primary,"<sup>331</sup> and "fundamental,"<sup>332</sup> problem of consequentialism. As a result, consequentialist retributivism would be unlikely to divest itself of this fundamental principle. But inclusion of this principle renders consequentialist

324. Moore does not champion it but merely recognizes it as a *possible* form of retributivism. See *supra* note 309.

325. Moore, *supra* note 20, at 90 (while subscribing to the equality principle, Moore is careful to note that retributivism should not be reduced to merely that principle).

326. Moore criticizes a victims' rights version of retributivism as untenable, in part, because it would violate the principle of equality:

In a truly victim-oriented system, if a wrongdoer has the good luck to injure one of the New Testament types, instead of one of those Old Testament types, then that wrongdoer is going to receive less punishment—because he is always going to get the turn-your-other-cheek forgiveness response from his victim. That is not equality. The propensity of a victim to forgive her transgressor is irrelevant to retributive desert. Two offenders, one of whom injures a resentful victim and the other of whom injures a forgiving victim, seem equally deserving of punishment, but a rights-based retributivism would not achieve that result.

*Id.* at 77. In critiquing a purported form of retributivism as an illegitimate form of retributivism, because it would violate the principle of equality, Moore seems to endorse the principle of equality as an integral component of a legitimate form of retributivism.

327. For a discussion of the duty to treat equally culpable wrongdoers equally, see *supra* Part II.B.5.

328. For a discussion of the duty to treat persons as ends rather than as mere means, see *supra* Part II.B.4.

329. For the conflict between the duty to maximize retributive punishment and the duty to treat persons as ends, rather than as mere means, see *supra* text accompanying note 322.

330. Greenawalt, *supra* note 89, at 1341 (noting that using offenders as mere means is the "most fundamental" difficulty with consequentialism).

331. See Morris, *supra* note 211, at 46.

332. See Duff, *supra* note 6, at 149.

retributivism internally inconsistent as applied to the Prosecutor's Dilemma.<sup>333</sup>

In conclusion, consequentialist retributivism does not satisfactorily avoid the Prosecutor's Dilemma. First, even apart from its application to bargain justice, consequentialist retributivism may be conceptually "impossible."<sup>334</sup> Second, consequentialist retributivism's duty to maximize retributive punishment conflicts with core retributive principles.<sup>335</sup> Third, while it could avoid inconsistency by divesting of most of retributivism's central principles, consequentialist retributivism would scarcely be a retributive theory at all. And it would share with consequentialism those fundamental failings pointed out by retributivists themselves.<sup>336</sup> And fourth, even if consequentialist retributivism avoids the Prosecutor's Dilemma, it cannot avoid variants of the Prosecutor's Dilemma.<sup>337</sup>

### B. Threshold Retributivism

Another alternative form of retributivism, threshold retributivism, might avoid the Prosecutor's Dilemma. After explaining the relation between this alternative form of retributivism and the broader philosophical view from which it stems, threshold deontology,<sup>338</sup> this

333. See *supra* text accompanying note 322.

334. Dolinko, *supra* note 315, at 513, 515.

335. To maximize the intrinsic good of retributive punishment, a consequentialist-retributivist prosecutor must accept the offer and punish the Two. But by doing so, the consequentialist retributivism system (i) fails to give the One his just deserts, (ii) violates the One's right to punishment, (iii) uses the One as a mere means, and (iv) fails to treat the One equally with the Two.

336. See *supra* notes 331-32 and accompanying text.

337. Assuming *arguendo* that consequentialist retributivism avoids internal inconsistency and avoids the Prosecutor's Dilemma, it cannot avoid the following variant. Suppose that all of the same facts obtain except that instead of three culpable wrongdoers, there are only two, *A* and *B*; *A* is in custody and offers to inculcate *B* in return for immunity. Unlike our original offer, there is nothing to maximize in this variation. Acceptance as well as nonacceptance of the offer only nets a consequentialist retributivist system the intrinsic good of punishing one culpable wrongdoer. From a maximizing perspective neither acceptance nor nonacceptance is preferable.

In the variation, the same internal conflicts arise as in the Prosecutor's Dilemma. By not accepting *A*'s offer (and foregoing punishment of *B*), a consequentialist retributivist system violates the duty to punish *B*, fails to give *B* her just deserts, fails to treat *B* equally with respect to the similarly situated *A*, violates *B*'s right to punishment, and fails to affirmatively treat *B* as an end in herself. Thus, it would seem that the offer should be accepted. But accepting the offer violates the very same four principles of retributivism. By accepting *A*'s offer (and foregoing punishment of *A*), a consequentialist retributivist system violates the duty to punish *A*, fails to give *A* his just deserts, fails to treat *A* equally with respect to the similarly situated *B*, violates *A*'s right to punishment and, uses *A* as a mere means in order to properly punish *B*. Thus, it would seem that the offer in the variation should be both accepted and not accepted.

338. For an account of the philosophical view of (non-threshold) deontology, see *supra* note 310.

section indirectly assesses the validity of threshold retributivism by considering criticisms of threshold deontology. Even assuming *arguendo* that threshold retributivism is tenable, it fails to dispel the Prosecutor's Dilemma.

### 1. Threshold Retributivism's Relation to Threshold Deontology

Threshold retributivism stems from threshold deontology, which combines deontology and consequentialism.<sup>339</sup> As Moore explains, "[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."<sup>340</sup> That is, unlike consequentialism, an act that is morally wrong should not be committed merely because it produces, on balance, good consequences.<sup>341</sup> And unlike deontology, the same morally wrong act may, or must, be committed if the good consequences of doing so become sufficiently great.<sup>342</sup>

339. Moore is not alone in endorsing threshold deontology. See, e.g., Donagan, Consistency, *supra* note 277, at 306-07. Charles Fried explains the special case in which the prohibition against intentionally harming innocents should be violated to avoid sufficiently adverse consequences:

[W]e can imagine extreme cases where killing an innocent person may save a whole nation. In such cases it seems fanatical to maintain the absoluteness of the judgment, to do right even if the heavens will in fact fall. And so the catastrophic may cause the absoluteness of right and wrong to yield, but even then it would be a non sequitur to argue (as consequentialists are fond of doing) that this proves that judgments of right and wrong are always a matter of degree, depending on the relative goods to be attained and harms to be avoided. I believe, on the contrary, that the concept of the catastrophic is a distinct concept just because it identifies the extreme situations in which the usual categories of judgment (including the category of right and wrong) no longer apply.

Charles Fried, *Right and Wrong* 10 (1978). Thomas Nagel suggests the difficulties of adhering to an absolutist position in which consequences never override the violation of norms:

While not every conflict between [deontological] absolutism and utilitarianism creates an insoluble dilemma, and while it seems to me certainly right to adhere to absolutist restrictions unless the utilitarian considerations favoring violation are overpoweringly weighty and extremely certain—nevertheless, when that special condition is met, it may become impossible to adhere to an absolutist position.

Nagel, *supra* note 276, at 56. See also Nozick, *supra* note 220, at 30 n.\* (acknowledging that consequentialist concerns might override deontological duties at a certain threshold).

340. Moore, *supra* note 20, at 158 n.13.

341. See *infra* note 371 and accompanying text.

342. As Moore explains,

[i]t 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. To prevent such extraordinary harms extreme actions seem to me to be justified.

Moore, *supra* note 20, at 719.

Threshold retributivism is merely threshold deontology applied to the punishment arena.<sup>343</sup> Under threshold retributivism, the good consequences of punishment are insufficient in justifying punishment until the good consequences become sufficiently great as to reach some threshold.<sup>344</sup> In other words, retributivist considerations control whether punishment is justified below the threshold level; but, at and above the threshold level, consequentialist considerations determine whether punishment is justified.<sup>345</sup>

Moore maintains that threshold retributivism, or threshold deontology, justifies offering and/or accepting immunity deals.<sup>346</sup> The conventional "retributivist might simply deny the propriety of the practice" of the intentional foregoing of any opportunity to punish a guilty offender in order to obtain the conviction of an even more guilty offender.<sup>347</sup> But if the retributivist "is a 'threshold deontologist', as am I, he might more qualifiedly disavow the practice except when it is needed to punish some very deserving criminal(s)."<sup>348</sup> That is, the duty to punish culpable wrongdoers whenever the opportunity arises precludes retributivism from justifying immunity deals.<sup>349</sup> But when a certain threshold is reached—some very deserving criminal(s) would otherwise escape punishment—then threshold retributivism would purportedly justify violating the duty to punish one offender in order to secure her testimony inculcating a criminal who is substantially more deserving of punishment.

## 2. Untenable Form of Retributivism

Before applying threshold retributivism to the Prosecutor's Dilemma, let us briefly present the serious criticisms leveled at threshold deontology.<sup>350</sup> Nancy Davis argues that though threshold

343. Though no theorist employs the term "threshold retributivism," our discussion of it will be based on Moore's threshold deontology as it applies, and as Moore applies it, to the punishment context.

344. For example, perhaps unlike consequentialism, the intentional punishment of a known innocent would not be justified merely because it would generate deterrence sufficient to prevent two innocents from being murdered. But if it would prevent the murders of, for example, 500 innocent victims, and this reached the threshold, then presumably threshold retributivism would justify the intentional punishment of a known innocent.

345. Threshold retributivism might also be stated in the negative: the bad consequences of failing to punish are insufficient to justify not punishing until they become sufficiently grave as to reach some threshold at which point the bad consequences become sufficient to justify not punishing.

346. Moore, *supra* note 20, at 158.

347. *Id.*

348. *Id.* (citation omitted).

349. In an immunity deal, the prosecutor foregoes the opportunity to punish one culpable wrongdoer to obtain the punishment of some other culpable wrongdoer(s). See *supra* Part I.A.2.

350. So far as I know, no commentator has addressed threshold retributivism. But commentators have addressed threshold deontology. The criticisms of threshold

deontology saves deontology from “fanaticism” and implausibility, utilization of the threshold undermines deontology as a consistent and sound theory.<sup>351</sup> Left unexplained, according to Davis, is why violating deontological norms below the threshold is absolutely wrong but permissible at, or above, the threshold.<sup>352</sup> Anthony Ellis argues persuasively that setting the threshold is inevitably and irretrievably arbitrary.<sup>353</sup> Larry Alexander entertainingly points out a number of

deontology may be relevant in assessing the plausibility of threshold retributivism.

351. Davis, *supra* note 310, at 215-16 (“Though allowing that we may violate deontological constraints in dire circumstances saves deontological views from the appearance of fanaticism, and thus confers greater normative plausibility on them, it may well undermine them as theories.”). Davis further explains the difficulties with threshold deontology:

It is hard to see how one can justify the view that a decision about whether or not to take the action that is necessary to save the nation (an action that would, in less extreme circumstances, be counted . . . as being wrongful) is not a *moral* decision. Such a view carries the suggestion that terrible circumstances somehow relieve us of the obligation (though surely not the need!) to act morally . . . . The view that concepts of right and wrong do not apply in extreme situations is one that encourages complacency, if not actual passivity. It is one that any responsible moral agent ought to reject.

*Id.* at 216.

352. *Id.* (“The addition of a ‘catastrophe clause’ is particularly problematic. Why do the effects of our acts on others’ weal and woe acquire relevance only at the ‘catastrophic’ level?”).

Moore has a reply to this sort of criticism. He denies that consequences are irrelevant to deontology prior to the threshold being reached. Consequences are relevant, but are outweighed by other principles, prior to the threshold. At the point of the threshold, consequences outweigh the other principles. Moore gives the following account of the role that consequences play both prior to, and at, the threshold:

[F]or a threshold deontologist consequences always ‘count’. For behaviour violating deontological constraints, however, until the threshold is reached the principle that makes such consequences count—the consequentialist principle itself—is outweighed by other moral principles. As the consequences get more and more severe, the consequentialist principle becomes of greater weight as applied to this situation, until at some point (the threshold) the consequentialist principle outweighs competing principles of morality. Even before such a threshold was reached, consequences counted but were of insufficient weight to determine the rightness of actions.

Moore, *supra* note 20, at 723. For a criticism of Moore’s reply, see Larry Alexander, *Deontology at the Threshold*, 37 San Diego L. Rev. 893, 907-10, 912 (arguing that for consequences to be relevant and outweighed below the threshold and outweigh deontological prohibitions above the threshold requires deontology and consequentialism to be commensurable; but, in fact, they are incommensurable).

353. Anthony Ellis, *Deontology, Incommensurability and the Arbitrary*, 52 Phil. & Phenomenological Res. 855, 859-70 (1992). A full exposition of Ellis’s argument is beyond the scope of this Article. In short, Ellis argues that the specification of any threshold will inevitably and necessarily be arbitrary because to locate the threshold one must locate the precise point at which the bad consequences of conforming to the deontological norm outweigh the wrongness of violating the deontological norm. But there is no scale upon which to comparatively weigh the badness of consequences versus the wrongness of violating deontological duties. According to Ellis, the badness of consequences and the wrongness of actions are incommensurable. *Id.* at

moral anomalies that would occur at the edges of the threshold<sup>354</sup> and concludes that threshold deontology is “erroneous”<sup>355</sup> and “implausible.”<sup>356</sup>

---

861-62. As a result, locating the threshold is necessarily arbitrary. *Id.*; cf. Davis, *supra* note 310, at 216 (“And to what (clear and practicable) features can agents appeal to distinguish a ‘catastrophic’ situation in which ‘right’ and ‘wrong’ do not apply from a merely dreadful one in which they do?”).

Moore anticipates this objection:

[T]he worry may be that any point we pick for a threshold beyond which consequences determine the rightness of an action may seem arbitrary. Do we need there to be 500 people in a building about to be blown up by a bomb before we may torture an innocent to find the bomb’s location, or will 450 do?

. . . [T]his worry . . . is not the basis of any powerful objection to threshold deontology as a moral theory. For this is no more than the medieval worry of how many stones make a heap. Our uncertainty whether it takes three, or four, or five, etc., does not justify us in thinking that there are no such things as heaps. Similarly, preventing the torture of two innocents does not justify my torturing one, but destruction of an entire city does.

Moore, *supra* note 20, at 724. Even if we accept that at some point sufficiently bad consequences should outweigh adherence to deontological norms, just as we accept that some sufficient number of stones will constitute a heap, nevertheless whatever level of bad consequences is chosen will be no less arbitrary than the number of stones that comprise a heap. And furthermore, while 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. Ellis, *supra* at 859.

354. In the following hypothetical Alexander asks us to assume that the threshold at which one has an affirmative duty to undertake a rescue is when twenty lives are at risk:

Twenty people are clinging to a housetop surrounded by rapidly rising floodwaters. Rescuer, who has a large boat that can easily and safely hold the threatened twenty, understands himself to have an affirmative duty to rescue them. As he pulls alongside the housetop, one of the stranded people jumps off and into Rescuer’s boat. Rescuer, realizing that there are now only nineteen people in harm’s way, turns away from the housetop and ignores the desperate pleas of the nineteen left behind.

Alexander, *supra* note 352, at 901. Alexander concludes that Rescuer violates no duty under threshold deontology in saving only the one because the nineteen at risk were under the threshold. *Id.* But what is disturbing is that “the other lives at risk [the nineteen] are pure moral ballast, serving no purpose other than to create a duty to save someone else.” *Id.* And, in effect, this seems to violate the very essence of threshold deontology:

[O]nce the single life is saved . . . Rescuer is relieved of his affirmative duty to rescue. By hypothesis, however, one may not be tortured or placed under an affirmative duty to rescue for the sake of one life. So cases at, or just above, the threshold seem paradoxical: the very deontological constraints that protect one from straight consequentialist trade-offs below the threshold disappear above it, so that one then can be killed, tortured, or forced to rescue merely to save one person.

*Id.*

355. *Id.* at 911.

356. *Id.* at 912.



### 3. Applicability to Offer

Despite the criticisms leveled at threshold deontology, let us assume *arguendo* that threshold retributivism is tenable. Does the failure to punish “some very deserving criminal(s)”<sup>357</sup> generate sufficiently bad consequences so as to reach the threshold level where threshold retributivism could justify immunity deals? Though Moore fails to supply a test as to precisely where the threshold should be set, and concedes the difficulty of doing so,<sup>358</sup> the common denominator of accounts of what clearly meets the threshold is that of a catastrophe.<sup>359</sup> Moore himself variously uses the standards of “extraordinary harms,”<sup>360</sup> “horrendous consequences,”<sup>361</sup> and “exceptionally awful consequences.”<sup>362</sup> Specific examples satisfying the threshold include the demise of a nation,<sup>363</sup> “if the heavens will in fact fall,”<sup>364</sup> “nuclear war,”<sup>365</sup> and the “destruction of a sizeable city by a terrorist nuclear device.”<sup>366</sup> For Moore, “the destruction . . . of a lifeboat or building full of people” only “perhaps” rises to the level of the threshold.<sup>367</sup> If such abhorrent consequences merely *perhaps* rise to the level of the threshold, then *a fortiori* it does not seem that foregoing punishment of a criminal who is “very deserving”<sup>368</sup> of punishment would satisfy the “horrendous consequences”<sup>369</sup> standard necessary for reaching the threshold.

Even assuming *arguendo* that failing to punish a very deserving offender meets the threshold (and thus threshold retributivism would justify at least some immunity deals), threshold retributivism would not justify a prosecutor accepting the offer in the Prosecutor’s Dilemma. There, each of the three offenders are equally culpable and equally deserving of punishment.<sup>370</sup> No offender is more deserving of punishment than the other. According to Moore, “preventing the torture of two innocents does not justify my torturing one, but destruction of an entire city does.”<sup>371</sup> Thus, if the torture of two

357. Moore, *supra* note 20, at 158.

358. *Id.* at 724 (setting the threshold “give[s] rise to quite genuine perplexity and anxiety when we make practical decisions”).

359. See Fried, *supra* note 339, at 10 (using the term “catastrophic” as a measure of the threshold); Alexander, *supra* note 352, at 897 (“moral catastrophes”); Davis, *supra* note 310, at 216 (referring to the threshold as the “catastrophe clause”).

360. Moore, *supra* note 20, at 719.

361. *Id.* at 721.

362. *Id.* at 722.

363. See Fried, *supra* note 339, at 10.

364. *Id.*

365. Moore, *supra* note 20, at 719.

366. *Id.*

367. *Id.* at 721-22.

368. *Id.* at 158.

369. See *supra* text accompanying note 361.

370. See *supra* text accompanying note 172.

371. Moore, *supra* note 20, at 724. In contrasting threshold deontology from

innocents does not meet the threshold whereby threshold deontology would justify the torture of one innocent to prevent the torture of the two, then similarly the failure to punish two offenders does not meet the threshold whereby threshold retributivism would justify foregoing punishment of one offender in order to punish two.<sup>372</sup> Since a necessary condition for threshold retributivism is not met (the threshold), threshold retributivism fails to avoid the Prosecutor's Dilemma.

### C. Conclusion

This part considered two alternative versions of retributivism—consequentialist and threshold retributivism—propounded, in part, to accommodate bargain justice. Consequentialist retributivism can justify bargain justice. But the cure may be worse than the malady. By gutting retributivism of its core values, consequentialist retributivism is a form of retributivism in name only. And it shares consequentialism's most fundamental problem (as pointed out by retributivists). While consequentialist retributivism can avoid the Prosecutor's Dilemma, it cannot avoid its variants. Threshold retributivism fares no better. Except in extreme circumstances, threshold retributivism cannot justify bargain justice and, under no circumstances, can it avoid the Prosecutor's Dilemma. Consequentialist retributivism's purchase of descriptive power comes at the very high price of eviscerating retributivism's ability to prescribe the just conditions for punishment. Threshold retributivism only gains modest descriptive power while still not avoiding the Prosecutor's Dilemma. By resolving the internal inconsistency revealed by the Prosecutor's Dilemma, retributivists might more profitably concede retributivism's lack of descriptive power in failing to accommodate bargain justice while seeking to restore its prescriptive power.

---

consequentialism, Moore further explains:

A consequentialist is committed by her moral theory to saying that torture of one person is justified whenever it is necessary to prevent the torture of two or more. The agent-relative view [i.e., deontological view], even as here modified [i.e., threshold deontology] is not committed to this proposition. To justify torturing one innocent person requires . . . a very high threshold of bad consequences . . .

*Id.* at 721-22.

<sup>372</sup> The threshold would not be met unless the equivalent of an entire city's criminals deserving punishment failed to receive their punishment. *See supra* text accompanying notes 366, 371.

## V. RESOLVING THE PROSECUTOR'S DILEMMA

Since retributivists cannot satisfactorily avoid the Prosecutor's Dilemma,<sup>373</sup> then they must resolve it.<sup>374</sup> There are two principal methods of resolution. First, retributivists might abandon entirely some duties until there are no duties remaining which conflict. This would require abandoning four of retributivism's central duties discussed above.<sup>375</sup> Shorn of these duties, the resulting theory would scarcely be retributive at all and would suffer from the same criticisms by which retributivists themselves condemn consequentialism.<sup>376</sup> Second, retributivists might downgrade the status of some duties from absolute to conditional. In the event that duties conflict, some duties would give way to other duties. This method allows retributivism to retain its duties (albeit in diluted form) but resolves conflicts by prioritizing among its duties. By effecting the most minimal modification of retributivism, prioritization is preferable to abandoning them outright.

### A. *Prioritizing Among Retributivism's Duties: Negative v. Positive*

Although there are competing conceptions of how to resolve moral dilemmas,<sup>377</sup> the preferred method among leading retributivists is the

373. Avoiding rather than resolving dilemmas (at least dilemmas *simpliciter*) is problematic and unrealistic. See *supra* note 283. The Prosecutor's Dilemma may quite easily be avoided—either cease the prosecution of criminals or eliminate retributivism completely. But these options are neither satisfactory nor necessary. For two unsuccessful attempts to avoid the Prosecutor's Dilemma, see *supra* Part III.C. Part IV considered two alternative versions of retributivism. While threshold retributivism fails to avoid the dilemma, consequentialist retributivism does avoid the dilemma, but in an unsatisfactory way. Consequentialist retributivism avoids the dilemma only by eliminating virtually all of retributivism's principles. Consequentialist retributivism is a form of retributivism in name only. And moreover, consequentialist retributivism cannot avoid variants of the Prosecutor's Dilemma. See *supra* note 337 and accompanying text.

374. For why moral dilemmas must be resolved (if not avoided), see *supra* notes 284-89 and accompanying text.

375. Under this method, all five duties except the equality principle would have to be abandoned. Eliminating any fewer than these four duties would still leave conflicts both within, and among, retributivism's duties.

376. See *supra* notes 331-32 and accompanying text.

377. According to the "The Dilemmas Thesis," "[t]here are moral conflicts in which the correct conclusion of moral deliberation" is that an agent is morally obligated to adopt each of the incompatible alternatives, despite that it is impossible to adopt them all. Gowans, *supra* note 254, at 49. Thus, for example, Sophie is morally obligated to save both Jan and Eva, despite that it is impossible, under the terms of the choice given to Sophie, to save both. And, under the "Dilemmas Thesis," our retributivist prosecutor is morally obligated to punish both the One and the Two, despite the impossibility of both accepting and not accepting the offer. Under this approach, then, resolving a moral dilemma requires eliminating the principle(s) in conflict giving rise to the dilemma. For a defense of this view, see Sinnott-Armstrong, *supra* note 266, at 53-71; McConell, *supra* note 283, at 36-44. For arguments against this view, see Gowans, *supra* note 254, at 70-75, 83-84.

prioritization of duties. Both Kant, the leading historical influence on retributivism,<sup>378</sup> and Moore, currently perhaps the leading retributivist among legal scholars,<sup>379</sup> agree that genuine moral dilemmas are impossible.<sup>380</sup> But both agree that prioritizing among the duties in

---

In contrast, under "The Options Thesis," "[f]or every moral conflict, the correct conclusion of moral deliberation" is that an agent is morally obligated to adopt only one of the conflicting alternatives. *Id.* at 49. Thus, for example, Sophie is only morally obligated to save either Jan or Eva, but not both. And a retributivist prosecutor, facing the Prosecutor's Dilemma, would be morally obligated to either (i) accept the offer and punish the Two, (ii) not accept the offer and punish the One, or (iii) not accept the offer and punish neither the One nor the Two. Under the "Options Thesis," moral dilemmas are transformed from situations in which no matter what one does is wrong to no matter what one does is, if not right, at least, permissible.

The Options Thesis is based in part on Kant's dictum that ought implies can. Although the principle is often attributed to Kant, he never used quite those words. See Sullivan, *supra* note 134, at 320 n.6 ("Kant is usually interpreted as holding that 'should' . . . implies 'can' . . . . Although he nowhere makes this claim in so many words, this is clearly his doctrine."). Sullivan cites a number of passages from which the principle has been cobbled. *Id.* (citing Immanuel Kant, *Critique of Practical Reason* 159 (Lewis White Beck ed., 1956) ("[W]e 'know that we *can* do [what the moral law commands] because our reason acknowledges [that law] as its law and says that we ought to do it")); Kant, *supra* note 8, at 380 ("man 'must judge that he *can* do what the law commands unconditionally that he *ought* to do"));). For discussions of the principle, see Michael Moore, *Law and Psychiatry: Rethinking the Relationship* 342 (1984) (discussing John Rawls and H.L.A. Hart's use of the principle); William K. Frankena, *Obligation and Ability, in Philosophical Analysis: A Collection of Essays* 148 (Max Black ed., 1950) ("[M]any moral philosophers . . . say, in one way or another, that 'ought' implies 'can.' Indeed, if there is anything on which philosophers are agreed with plain men and with each other, and goodness knows there is very little, it is Kant's dictum . . ."). For criticism of Kant's dictum, see Sinnott-Armstrong, *supra* note 266, at 113-26. According to Kant's dictum, we are under a duty to do something only if we physically can do it. If we physically cannot do something, then we are not under a duty to do it. Kant's dictum limits the application of morality's absolute dictates to that which is possible. Thus, for example, since Sophie cannot save both Jan and Eva, Sophie is not morally obligated to save both. But since Sophie can save either one or the other, she is morally obligated to save one or the other.

The difficulty with the Options Thesis is that it fails to resolve the resulting dilemma of *which* of the incompatible alternatives to adopt. *Id.* at 56. For example, must Sophie save Jan or Eva? Some Options Thesis defenders maintain that the choice of which alternative to adopt is *perhaps* simply nonmoral. See Donagan, *Consistency, supra* note 277, at 308 (the choice is not a moral, but a practical problem); Marcus, *supra* note 283, at 34 (the choice is not moral but personal). Whether described as practical or personal, however, a nonmoral decision as to which alternative to adopt—e.g., which child Sophie should save—will inevitably be arbitrary and unprincipled. And most commentators have argued against the legitimacy of arbitrary dilemma resolution. See *infra* notes 409-15 and accompanying text.

378. See *supra* note 132.

379. See Douglas Husak, *Retribution in Criminal Theory*, 37 San Diego L. Rev. 959, 960 (2000) ("No contemporary criminal theorist rivals Moore in his unqualified enthusiasm for retribution.").

380. In a famous passage, Kant declares that conflicting moral obligations or duties are inconceivable:

A conflict of duties (*collisio officiorum s. obligtionum*) would be a relation

conflict is the preferred means to resolve *apparent* moral dilemmas.<sup>381</sup> Whereas Kant prioritized perfect or narrow duties over imperfect or wide duties,<sup>382</sup> contemporary Kantians emphasize the priority of negative duties over positive duties.<sup>383</sup> Although the perfect/imperfect

between them in which one would cancel the other (wholly or in part). But since duty and obligation are concepts that express the objective practical *necessity* of certain actions and two rules opposed to each other cannot be necessary at the same time, if it is a duty to act in accordance with one rule, to act in accordance with the opposite rule is not a duty but even contrary to duty; so a *collision of duties* and obligations is inconceivable (*obligationes non colliduntur*).

Kant, *supra* note 8, at 50. Kantian scholar Roger Sullivan explains Kant's position as based on consistency as a paramount value:

[S]ince the ultimate moral norm commands us to act self-consistently, it is "inconceivable," Kant writes, that we could be obligated to act inconsistently. He therefore insists, as he must, that there *cannot* be a genuine conflict between *duties* (*ein Widerstreit der Pflichten*), a conflict such that we would have a duty to act at the same time on conflicting rules.

Sullivan, *supra* note 134, at 73 (quoting Kant, *Metaphysics of Morals, supra* note 8, at 50). For further discussion of Kant's denial of the possibility of genuine moral dilemmas, see Gowans, *supra* note 254, at 187-92; Barbara Herman, *Obligation and Performance: A Kantian Account of Moral Conflict, in Identity, Character, and Morality: Essays in Moral Psychology* 314, 314-16 (Owen Flanagan & Amelie Rorty eds., 1990); Hill, *supra* note 287, at 173, 187; Marcus, *supra* note 283, at 24.

Michael Moore agrees with Kant—there are no genuine moral dilemmas. Michael Moore, *Moral Reality Revisited*, 90 Mich. L. Rev. 2424, 2463-64 ("On that issue [the existence of genuine moral dilemmas], I have aligned myself with Kant ('a conflict of duties and obligations is inconceivable')." (quoting Kant, *supra* note 8, at 50) (citation omitted)).

381. Kantian scholar H.J. Paton explains that for Kant, perfect or narrow duties take priority over imperfect or wide positive duties:

Kant attaches great importance to the distinction between perfect and imperfect duties, but he seems nowhere to define the distinction clearly . . . Perfect duties . . . admit of no exception in favour of inclination, and this would suggest that imperfect duties do admit of such exceptions.

. . . In the case of perfect duties we are obliged to perform a definite *act*—for example, to pay precisely the . . . [specific amount] which we owe. In the case of imperfect duties we are bound to act only on a *maxim*: although we ought to act on the maxim of benevolence, it is left to our discretion to decide whom we ought to help. There is thus a 'latitude' or 'play-room' in the case of imperfect duties, which are also called 'broad' or 'meritorious,' as opposed to 'narrow' or 'strict' or 'rigorous,' duties. The last three adjectives appear to be equivalent to 'perfect' when they are applied to duties.

Paton, *supra* note 275, at 147-48. Moore also recognizes that moral obligations may be prioritized. Moore, *supra* note 380, at 2464 (at least in the case of incomparable dilemmas a resolution "may exist in terms of a lexical priority between our principles").

382. See *supra* note 381; see also Sullivan, *supra* note 134, at 51-54 (explaining Kant's prioritization of narrow, perfect or strict duties over wide and imperfect duties and the distinction between them). Simply stated, imperfect duties "permit one to choose among several possible ways of fulfilling them[, but] *perfect duties* do not allow that leeway." Bruce Russell, *Duty, in The Cambridge Dictionary of Philosophy* 248, 249 (Peter Singer ed., 1991). For another account of the distinction between perfect and imperfect duties, see Onora O'Neill, *Kantian Ethics, in A Companion to Ethics*, 175, 178 (Robert Audi ed., 2d ed. 1999).

383. For Moore, "negative duties are much stricter than merely positive duties."

duty distinction may imprecisely track the negative/positive duty distinction,<sup>384</sup> most perfect duties are negative and most imperfect duties are positive.<sup>385</sup> Stated most simply, “[n]egative duties are duties not to do certain things such as to kill or harm, while *positive duties* are duties to act in certain ways, such as to relieve suffering or bring aid.”<sup>386</sup> Since negative duties are prohibitions against committing certain acts, negative duties can be fulfilled by simply doing nothing.<sup>387</sup> In contrast, positive duties can only be fulfilled by doing something.<sup>388</sup> Thus, since negative duties take priority over positive duties, doing nothing will take priority over doing something.

Resolving the Prosecutor’s Dilemma via the prioritization rule requires first identifying which of the five retributivist principles or duties<sup>389</sup> are positive and which are negative. The following three duties are positive: punishing culpable wrongdoers, giving them their just deserts, and honoring their right to punishment.<sup>390</sup> These duties are positive since only affirmative acts fulfill them; doing nothing violates all three duties. And as positive duties, the priority rule accords their fulfillment or violation secondary importance.

The two remaining duties<sup>391</sup> contain both positive and negative

Moore, *supra* note 20, at 689 (discussing Phillipa Foot, *The Problem of Abortion and the Doctrine of Double Effect*, 5 *Oxford Rev.* 1, 5-15 (1967)). Surveying contemporary deontology, Nancy Davis finds negative obligations to be more fundamental than positive obligations. Davis, *supra* note 310, at 208-09. Despite questioning the usefulness of the distinction in resolving all moral dilemmas, Walter Sinnott-Armstrong acknowledges that Kantians regard negative duties as overriding positive duties. Sinnott-Armstrong, *supra* note 266, at 82-85.

384. Sullivan, *supra* note 134, at 51 (“Our strict [perfect] duties may be interpreted positively, such as the requirement to keep contracts, but they are fundamentally *negative* . . .”).

385. Russell, *supra* note 382, at 249 (“Most positive duties are imperfect; most negative ones, perfect.”). See Sullivan, *supra* note 134, at 73-74 (interpreting Kant’s priority rule to be that “[n]egative and strict [perfect] duties as a rule thus carry more weight than positive duties”).

386. Russell, *supra* note 382, at 249. Moore explains that “morality places *negative* duties on us not to do or cause things like torture, whereas morality places *positive* duties on us not to allow things like torture to be done where we can prevent it.” Moore, *supra* note 20, at 689. Moore suggests that the negative duty/positive duty distinction is similar to the acting/allowing distinction. *Id.* (quoting Phillipa Foot, *The Problem of Abortion and the Doctrine of Double Effect*, 5 *Oxford Rev.* 1, 26 (1967) (distinguishing “between what one does or causes and what one merely allows”). That is, doing some act may violate a negative duty, whereas not doing an act but allowing a state of affairs to occur may violate a positive duty. For a further explanation of the acting/allowing distinction, see Davis, *supra* note 310, at 216.

387. While negative duties can be fulfilled by doing nothing, they can only be violated by doing something.

388. While positive duties may only be fulfilled by doing something, they can be violated by doing nothing.

389. See *supra* Part II.B.

390. See *supra* Part II.B.1-3.

391. With three of the five duties identified as positive duties, see *supra* text accompanying note 390, there are two duties (of the five) remaining: (i) to treat persons as ends and not as mere means, and (ii) to treat persons equally. In general,

components. The duty to treat persons as ends and not as mere means<sup>392</sup> consists of the positive duty to treat persons as ends and the negative duty to not treat persons as mere means.<sup>393</sup> The duty to treat equally those culpable wrongdoers with equal desert also consists of both a positive and negative duty.<sup>394</sup> The positive duty may only be fulfilled by the affirmative acts of punishing both the One and the Two, and punishing them equally.<sup>395</sup> The negative duty of not treating them unequally can be fulfilled by doing nothing.

The second step, under the prioritization rule, is determining which, if any, of the three options<sup>396</sup> available to a retributivist prosecutor under the Prosecutor's Dilemma satisfy all (i.e., both) of the negative duties.<sup>397</sup> Thus, the rule requires elimination of any option violating a negative duty. Prioritization requires eliminating the option of accepting the offer (and punishing only the Two) since it violates both negative duties: (i) using the One as a mere means to obtain the punishment of the Two violates the negative duty to not treat persons as mere means<sup>398</sup> and, (ii) punishing the Two and not the One violates

---

the latter duty may be fulfilled either positively or negatively. *See supra* notes 252-53 and accompanying text. But under the terms of the Prosecutor's Dilemma, the duty may only be fulfilled negatively. *See supra* Part II.C.3.

392. *See supra* Part II.B.4.

393. *E.g.*, Paton, *supra* note 275, at 171-72 ("So far as we take the principle negatively it forbids us to use rational agents merely as a means . . . . But we must also take our principle positively: it bids us to act on the maxim of furthering the ends of rational agents."). For further discussion of Kant's maxim as containing both a positive and negative duty, see *supra* notes 222-26 and accompanying text.

394. *See supra* Parts II.B.5 and III.A.2.

395. But under the terms of the Prosecutor's Dilemma, the positive duty is impossible to fulfill. *See supra* Part II.A.

396. For an explanation of the three, and only three, options available to a retributivist prosecutor in the Prosecutor's Dilemma, see *supra* Part II.C.

397. According to the prioritization rule, fulfillment of negative duties over positive duties, the one option which fulfills all of the applicable negative duties is the option which must be taken even if it involves violating all of the positive duties. If there is more than one option which satisfies all of the negative duties, then fulfillment of positive duties becomes relevant under the prioritization rule. This is because the prioritization rule only makes fulfillment of positive duties of secondary value, but not completely irrelevant. *See supra* notes 381-85 and accompanying text. Thus, where more than one option satisfies all of the negative duties, then (of those options) the one which satisfies the greatest number of positive duties would be preferable. If no option satisfies all of the negative duties, then the option which satisfies the greatest number of negative duties is preferable. If there is no option which satisfies all of the negative duties, but more than one option which satisfies an equal number of negative duties, then (of those options) the one which satisfies the greatest number of positive duties is preferable.

398. Retributivism purportedly treats culpable wrongdoers as ends (thereby satisfying the positive component of the duty) by giving them the punishment that they deserve. But they are used as mere means (thereby violating the negative component of the duty) if their punishment or non-punishment is a means toward some other purpose. *See supra* text accompanying notes 224-26. Accepting the offer (and punishing only the Two) satisfies the positive component of the duty with respect to the Two by giving them the punishment that they deserve. But punishing

the negative duty to not treat persons unequally.<sup>399</sup> The option of not accepting the offer (and punishing only the One) satisfies the negative duty to not treat persons as mere means since neither the One nor the Two are used as a means to obtain the punishment of another.<sup>400</sup> But by violating the negative duty not to treat the One and Two unequally,<sup>401</sup> the option suffers elimination under the priority rule.<sup>402</sup>

The only remaining option—not accepting the offer and punish neither the One nor the Two—does satisfy both negative duties. By using neither the One nor the Two as mere means (nor in any way at all), it satisfies the negative duty to not use persons as mere means.<sup>403</sup> And by punishing neither the One nor the Two, that option satisfies the negative duty to not treat the One and the Two unequally.<sup>404</sup>

Prioritizing negative duties over positive duties resolves the Prosecutor's Dilemma.<sup>405</sup> Of the three options, the prioritization rule requires that a retributivist prosecutor not accept the offer and punish neither the One nor the Two. While that option violates all five of the positive duties,<sup>406</sup> it is the only option which fulfills all of the negative duties. Though providing a determinate answer as to which option a prosecutor must select, the resolution does so only by diluting the

---

only the Two fails to satisfy the positive component of the duty with respect to the One by failing to give the One the punishment that she deserves. And, more importantly, by accepting the offer and punishing only the Two a retributivist prosecutor uses the One as a mere means—her non-punishment is a means by which to obtain the punishment of the Two. Thus, accepting the offer and punishing only the Two violates the negative component of the duty.

399. The option of accepting the offer affords punishment of the Two, but precludes punishment of the One. By punishing the Two but not the One, the option violates the negative duty not to treat them unequally.

400. See *supra* text accompanying note 247.

401. The option of not accepting the offer (and punishing only the One) foregoes punishment of the Two. By affording punishment of the One but not the Two, the option violates the negative duty not to treat them unequally.

402. Though neither accepting the offer nor not accepting the offer (and punishing only the One) satisfies all of the negative duties, the latter option is preferable to the former under the prioritization rule. While the former option violates all of the negative duties, the latter option does satisfy one negative duty.

403. Not accepting the offer (and punishing none) uses neither the One nor the Two as mere means to either secure the punishment of another or to attain some societal benefit. Indeed, it fails to use them as means in any way at all.

404. Not accepting the offer (and punishing none) is the only option which satisfies the negative duty to not treat the One and the Two unequally. Each of the other options, in that each involves punishing either only the One or only the Two, violates the negative duty to not treat the One and the Two unequally.

405. If the negative/positive duty distinction imprecisely tracks the perfect/imperfect duty distinction, the appropriate resolution of the Prosecutor's Dilemma may vary as a function of the employment of the former or latter distinction.

406. Nonacceptance of the offer (and punishing none) violates the following five positive duties with respect to both the One and the Two: (i) punishing them, (ii) giving them their just deserts, (iii) honoring their right to punishment, (iv) treating them affirmatively as ends, and (v) affirmatively, or positively, treating them equally.



absolute or categorical status of retributivism's positive duties to that of contingent duties. While technically restoring retributivism to consistency, the resolution still fails to fulfill retributivism's positive duties.

### B. Deprioritizing the Equality Principle

Although it might lack the endorsement of retributivists and contemporary Kantians, there are other ways to resolve the Prosecutor's Dilemma. Deprioritizing the equality principle affords adoption of one of the other two options—accepting the offer and punishing the Two or not accepting the offer and punishing the One.<sup>407</sup> But deprioritizing the equality principle fails to completely resolve the Prosecutor's Dilemma. A dilemma would still remain as to whom should be punished—the One or the Two? Deprioritizing the equality principle only eliminates one option—not accepting the offer (and punishing neither the One nor the Two). It fails to provide a basis for choosing to seek punishment of the One or the Two.<sup>408</sup>

In order to decide whom to punish—if someone is to be punished at all—a retributivist prosecutor would need a principled, nonarbitrary reason why one party should be punished and the other not.<sup>409</sup> While persons may be fungible and interchangeable under utilitarianism, they are not under a Kantian framework.<sup>410</sup> Using random

407. Deprioritizing the equality principle eliminates the inconsistency between the equality principle and retributivism's other principles. If the equality principle is no longer absolute but merely conditional, it would be overridden when it conflicted with the group of four principles. A retributivist prosecutor could thereby opt to either not accept the offer (and punish only the One) or accept the offer (and punish only the Two) without violating an absolute duty to treat the One and the Two equally.

408. And by failing to supply a basis for deciding whom to punish, deprioritization of the equality principle fails to supply a basis for deciding whether to accept the offer (and punish only the Two) or not accept the offer (and punish only the One).

409. Christopher W. Gowans, *Introduction, in Moral Dilemmas* 6, 27-28 (Christopher W. Gowans ed., 1987) (an arbitrary resolution of a dilemma is an inadequate resolution); Sinnott-Armstrong, *supra* note 266, at 60-61 (stating that since persons are not interchangeable, the choice between persons for whom a duty should be fulfilled and for whom it should be violated does matter and arbitrariness is problematic). Thomas Nagel finds arbitrariness permissible when two choices are even, but not when two choices appear different, yet absolutely required:

When two choices are very evenly balanced, it does not matter which choice one makes, and arbitrariness is no problem. But when each seems right for reasons that appear decisive and sufficient, arbitrariness means the lack of reasons where reasons are needed, since either choice will mean acting against some reasons without being able to claim that they are *outweighed*.

Nagel, *supra* note 276, at 128-29. If persons are not interchangeable under retributivism, as they might be under utilitarianism, *see infra* note 410 and accompanying text, then no dilemmatic choice could be evenly balanced and arbitrariness will be problematic.

410. Rawls' views in contrasting utilitarianism with his theory of justice as fairness, which he identifies as deontological or Kantian as opposed to utilitarian, Rawls, *supra* note 310, at 30, are instructive. *Compare id.* at 3 ("Each person possesses an

procedures<sup>411</sup> such as flipping a coin,<sup>412</sup> drawing straws or conducting a lottery<sup>413</sup> to decide whom to punish would be an arbitrary<sup>414</sup> and unprincipled means of decision and one that retributivists themselves would find illegitimate.<sup>415</sup> But retributivism fails to supply a principled, nonarbitrary rationale.<sup>416</sup> The same four principles requiring punishment of the One also require punishment of the Two. And punishment of only the One violates the same four principles as does punishment of only the Two.

Since there is nothing within retributivism that affords a rationale for deciding whom to punish, then such a rationale must originate from without retributivism. But what could this supplement to retributivism be? Apart from the possibility of the differing consequences of punishing the One and the Two, a nonarbitrary basis to differentiate among equally culpable and equally blameworthy wrongdoers is lacking.<sup>417</sup> Relying on the differing consequences of punishing the One and the Two would, in effect, amount to adopting consequentialist retributivism.<sup>418</sup> Though consequentialist retributivism would resolve the dilemma as to whom should be punished—the Two<sup>419</sup>—its adoption would come at a heavy price.<sup>420</sup>

inviolability founded on justice. . . .”) with *id.* at 27 (“Utilitarianism does not take seriously the distinction between persons.”). See Sinnott-Armstrong, *supra* note 266, at 56-57 (finding the argument that an arbitrary choice between persons is permissible as utilitarian); *id.* at 76-77 (treating persons as fungible in a dilemmatic choice renders utilitarianism implausible as a moral theory).

411. Sinnott-Armstrong, *supra* note 266, at 180 (a moral theory which endorses the utilization of random procedures to resolve a dilemma renders the theory “not . . . very plausible”).

412. *Id.* at 55-56 (a coin flip to resolve a dilemma is inadequate because it fails to make one moral requirement or duty stronger than the opposing one); Marcus, *supra* note 283, at 23 (tossing a coin is not a morally justified grounds for choosing what to do in a moral dilemma situation).

413. Marcus, *supra* note 283, at 23 (conducting a lottery to determine what course of action to take when confronted with a dilemma is morally unjustified). Kantian scholar Roger Sullivan explains that Kant would have abhorred the use of nonrational, arbitrary decision-making procedures such as the use of a lottery: “the proposal is self-defeating, since it only postpones unequal treatment and does so in a way Kant would have abhorred—by a completely nonrational decision process. Luck, Kant wrote, ‘cannot become a matter of universal legislation.’” Sullivan, *supra* note 134, at 375 n.21 (citation omitted).

414. For an interesting account of a possible distinction between randomness and arbitrariness, see Duxbury, *supra* note 35, at 46-47, 100-01.

415. For a general discussion of why Kant would exclude the element of luck from moral judgments, see Daniel Statman, *Introduction, in Moral Luck* 1, 4 (Daniel Statman ed., 1993).

416. This failing applies even to a version of retributivism without an absolute equality principle.

417. Since the One and the Two are equally situated in all respects which are relevant to retributivism, whatever basis utilized to support only punishing the One would equally support punishing only the Two.

418. For an account of consequentialist retributivism, see *supra* Part IV.A.1.

419. Punishing the Two rather than the One would maximize the intrinsic good of retributive punishment. See *supra* notes 320-21 and accompanying text.

In addition, it could not avoid dilemmas arising from variants of the Prosecutor's Dilemma.<sup>421</sup> By excluding the differing consequences of punishing only the One and only the Two, retributivism lacks a principled, nonarbitrary means of deciding whom—between the One and the Two—should be punished. And even by including the differing consequences, retributivism lacks a *satisfactory* basis for deciding.<sup>422</sup> As a result, deprioritizing the equality principle is an unsatisfactory solution to the Prosecutor's Dilemma.<sup>423</sup>

### C. Practical Ramifications of the Resolution

While the theoretically sound resolution (at least according to retributivists and modern Kantians<sup>424</sup>) of the Prosecutor's Dilemma is to not accept the offer and punish neither the One nor the Two, this resolution has the worst practical consequences for a retributivist system of criminal justice. It leads retributivism into the *reductio ad absurdum* of allowing culpable wrongdoers to escape prosecution and punishment merely by tendering an immunity offer to a retributivist prosecutor which generates a moral dilemma for retributivism. The mere tender of such a dilemmatic immunity offer enables a culpable wrongdoer to, in effect, self-immunize. Moreover, the tender of a dilemmatic immunity offer allows the prisoner to not only immunize herself but also her confederates. The prospect of a retributivist system of punishment forced to withdraw from seeking punishment of

---

420. For an account of consequentialist retributivism's difficulties, see *supra* Part IV.A.2-3.

421. See *supra* note 337 and accompanying text.

422. While inclusion of the differing consequences of punishing the One and the Two may provide a nonarbitrary basis for deciding whom to punish, it is not a satisfactory basis because of the difficulties of consequentialist retributivism. See *supra* Part IV.A.2-3.

423. First, allowing the equality principle, as a fundamental precept of justice, to be overridden, is a substantial loss. Such a devaluing would allow retributivism to treat similar cases dissimilarly and dissimilar cases similarly. Second, the equality principle is the only retributivist principle to be unscathed by the Prosecutor's Dilemma. The other four principles are internally inconsistent as applied to the Prosecutor's Dilemma. If any of retributivism's principles are to remain intact, it should be the equality principle. Third, devaluing the equality principle does not even fully resolve the Prosecutor's Dilemma. Even if the option of punishing none is removed, the dilemma remains as to whom—between the One and the Two—should be punished. Without utilizing consequences, there is no principled, nonarbitrary means by which to choose. Further, the utilization of consequences by retributivism is not only seriously problematic but would not resolve variants of the Prosecutor's Dilemma. Fourth, and perhaps most importantly, resolving the dilemma in this way does not accord with the method endorsed by contemporary Kantians and retributivists to resolve moral dilemmas.

424. See *supra* note 383 for the endorsement by modern Kantians of the prioritization of negative duties over positive duties. For the application of the priority rule to the Prosecutor's Dilemma, see *supra* notes 389-406 and accompanying text.

culpable wrongdoers, merely by receipt of a dilemmatic immunity offer, recalls the plight of "Buridan's ass."<sup>425</sup>

The Prosecutor's Dilemma places into the hands of prisoners a more powerful bargaining tool than that used by prosecutors to create the prisoner's dilemma.<sup>426</sup> While the bargain justice proposals initiated by prosecutors—enmeshing prisoners in the prisoner's dilemma—lead "rational" prisoners to choose their worst aggregate outcome,<sup>427</sup> a prisoner-initiated proposal enmeshing a prosecutor into the Prosecutor's Dilemma generates the best possible outcome for the prisoners—immunity for all. But in the prisoner's dilemma, what is the worst aggregate outcome for prisoners is perhaps the best collective outcome for society and enhances the goals of the consequentialist system of criminal justice under which it arises.<sup>428</sup> In contrast, in the Prosecutor's Dilemma, what is the best outcome for the prisoners is perhaps the worst collective outcome for society<sup>429</sup> and undermines the retributivist system of criminal justice under which it arises.

Thus, punishing none might be considered a worse resolution, as a practical matter, than the other two options.<sup>430</sup> As impractical as the resolution may be for retributivism, without a resolution to the Prosecutor's Dilemma retributivism remains internally inconsistent.

425. "Buridan's ass" is a philosophical term originally used to deride French philosopher Jean Buridan's theory that freedom of action meant the "ability to defer choice in the absence of a compelling reason to act one way or the other." Jack A. Zupko, *Buridan, Jean*, in *The Cambridge Dictionary of Philosophy* 107, 108 (Robert Audi ed., 2d ed. 1999). In the example which came to be known as Buridan's ass (attributed to the philosopher Baruch Spinoza), an ass (a creature similar to a donkey), situated equidistant between two equally enticing bales of hay, starves to death because it has no rational basis to eat from one bale of hay over the other. *Id.* The term has come to be known as describing situations in which the choice of which course of conduct to pursue is deferred until reason has determined which course is preferable. *Id.*; Blackburn, *supra* note 249, at 52. The situation for a retributivist prosecutor is even worse. While Buridan's ass has two equally enticing, or good, options, a retributivist prosecutor has only bad options. No matter what option a prosecutor chooses, retributivism's principles will be violated.

426. For a discussion of the prisoner's dilemma, see *supra* notes 25-26, 35-38 and accompanying text.

427. For a ranking of the various aggregate outcomes for prisoners in the prisoner's dilemma, see *supra* note 37 and accompanying text.

428. This argument is premised on the notion that guilty offenders receiving greater aggregate punishment, as long as it is not more than what is deserved, than lesser aggregate punishment, promotes the consequentialist goals of incapacitation and deterrence. For further discussion of how the prisoner's dilemma promotes consequentialist goals, see *supra* note 37 and accompanying text.

429. The nonpunishment of guilty offenders is an obvious detriment to societal interests. Even under a retributivist perspective, the nonpunishment of guilty offenders is problematic. It violates at least four retributivist principles. See *supra* Part II.C.3.

430. But unlike consequentialism, it is unclear how these adverse practical consequences could legitimately enter into the calculus of retributivist decision-making.

Either way—with or without the resolution—the incompatibility of retributivism and bargain justice is no longer an indictment of the latter, but rather is now an indictment of the former.

### CONCLUSION

Bargain justice places retributivism at the crossroads. With its condemnation of bargain justice, retributivism is unable to accommodate and account for the mechanisms by which the vast majority of defendants are rendered subject to punishment. Retributivism risks irrelevance as a theory which can explain what legal institutions of punishment actually do. Attempts to enhance its descriptive power by forging a reconciliation with bargain justice only serve to undermine retributivism's prescriptive power in delineating the just conditions for the imposition of punishment.

The Prosecutor's Dilemma compounds retributivism's difficulties with bargain justice. Despite withholding justification from bargaining for punishments, retributivism cannot preclude prisoner-initiated bargaining. A particular form of which—the Prosecutor's Dilemma—reveals inconsistencies both within, and among, five of retributivism's central principles. Although this dilemma, and retributivism's internal inconsistency, may be resolved, the resolution affords prisoners a greater advantage (the ability to, in effect, self-immunize) over prosecutors than the advantages which prosecutors enjoy over prisoners via the prisoner's dilemma. The resolution perhaps reduces retributivism to the absurd—the mere tender of a prisoner-initiated dilemmatic immunity offer forecloses retributivism from justifying the punishment of either the prisoner or the confederates the prisoner sought to betray. Without resolving the Prosecutor's Dilemma, however, retributivism's prescription of the just conditions for punishment is internally inconsistent.

As a result, the Prosecutor's Dilemma supplies an indirect defense of bargain justice. The Prosecutor's Dilemma exposes as internally inconsistent the very principles of retributivism which bargain justice violates, and by which retributivists condemn bargain justice. Either way—with or without the resolution of the Prosecutor's Dilemma—the incompatibility of retributivism and bargain justice now renders illegitimate not bargain justice, but rather retributivism.