Notre Dame Journal of Law, Ethics & Public Policy

Volume 31 | Issue 2 Article 3

2017

Efficiency, Enforcement, and Punishment

Jim Staihar
Robert H. Smith School of Business, University of Maryland

EFFICIENCY, ENFORCEMENT, AND PUNISHMENT

JIM STAIHAR*

ABSTRACT

The law and economics literature on punishment reveals strong reasons of efficiency to adopt an extreme enforcement policy for any type of crime as a means to promoting deterrence. Under such an extreme policy, a crime's severity of punishment would be set extremely high, but its probability of punishment would be set extremely low by minimizing the resources devoted to enforcing the law against the crime. This sort of policy applied to a moderately serious crime, such as a simple assault, would seem strongly unreasonable all things considered. However, it is not immediately obvious why such a policy would be so unreasonable on the assumption that the policy would be an efficient means of promoting deterrence. In this Essay, I argue that a novel theory of deserved punishment best explains why such an extreme enforcement policy would be so unreasonable when applied to a moderately serious crime. Although such an extreme policy might be efficient, someone who commits a moderately serious crime would not deserve the extremely severe punishment he would receive under the policy. In general, I conclude that my theory of punitive desert should play an important deontological role in constraining the means by which a state may enforce its laws against any type of crime.

INTRODUCTION

The law and economics literature on punishment reveals strong reasons of efficiency to adopt an extreme enforcement policy for any type of crime,[1] including a moderately serious crime, such as a simple

* Assistant Professor, Robert H. Smith School of Business, University of Maryland. Associate Director, Center for the Study of Business Ethics, Regulation, and Crime, University of Maryland. For financial support while working on this Essay, I thank the University of Maryland, the University of Chicago Law School, Princeton University's Law and Public Affairs Program, and the University of Michigan. For comments, I thank Elizabeth Anderson, Stephen Darwall, Allan Gibbard, and Thomas Green.

1. I assume that the essence of a crime consists in manifesting insufficient concern for the rights of others in committing a criminal act. The manifestation of such insufficient concern entails a degree of recklessness: an awareness of an unjustifiable risk that one's criminal act will violate the rights of others. *See* Larry Alexander, *Insufficient Concern: A Unified Conception of Criminal Culpability*, 88 CAL. L. REV. 931, 931 (2000); LARRY ALEXANDER & KIMBERLY KESSLER FERZAN, WITH STEPHEN J. MORSE, CRIME AND CULPABILITY: A THEORY OF CRIMINAL LAW 23–31 (2009); MODEL PENAL CODE § 2.02(2)(c) (1985). I also assume that a more serious crime manifests a worse deficiency in the offender's concern for the interests of others, and the harmful results of the crime, if any, do not bear directly on its seriousness. *See, e.g.*, ALEXANDER & FERZAN, *supra*, at 171–96 (arguing that the harmful results of a crime do not affect the culpability of the offender); Joel Feinberg, *Equal Punishment for Failed Attempts: Some Bad But Instructive Arguments Against It*, 37 ARIZ. L. REV. 117 (1995).

assault.[2] Under such an extreme policy, a crime's severity of punishment would be set extremely high, but its probability of punishment would be set extremely low.[3] Thus, under this sort of policy, a state would punish those who commit the crime extremely severely, but the rate at which the state would detect and punish those who commit the crime would be extremely low. The efficiency reasons favoring an extreme enforcement policy for any type of crime follow from four assumptions.

First, suppose deterrence is an important valuable aim of a state's criminal justice system. A criminal justice system promotes deterrence by providing people with a prudential incentive not to commit crimes out of fear of being detected and punished for committing crimes.

Second, suppose people are deterred from committing a crime by the crime's expected punishment. The expected punishment of a crime is the product of the crime's severity of punishment and the crime's probability of punishment. In other words, a crime's expected punishment is its severity of punishment discounted by the probability of detection and actually suffering the punishment threatened by the state.

Third, assume a crime's probability of punishment varies with the amount of resources that a state devotes to enforcing the law against the crime. So a state can lower a crime's probability of punishment by devoting less resources to enforcing the law against the crime. For example, the state might hire fewer police officers to search for those who commit the crime.

Fourth, suppose there is a scarcity of resources needed to satisfy people's critical needs. Hence, if one enforcement policy for a crime requires less resources to administer than another policy, the resources saved could be used to satisfy some people's critical needs that would otherwise go unsatisfied. For example, the resources saved might be devoted to meeting the healthcare needs of citizens rather than funding law enforcement.

Under these four assumptions, a state has strong reason to adopt an extreme enforcement policy for any crime because such a policy would be the most efficient means of promoting deterrence.[4] To

2. I presume that a moderately serious crime manifests a moderately bad deficiency in the criminal's concern for the rights of others.

3. The probability that someone will be detected and punished for committing a crime will vary from situation to situation. So when I refer to a crime's probability of punishment under an enforcement policy, I mean the average probability that individuals would be detected and punished for committing the crime across all the situations in which they might find themselves.

4. For more discussion about the efficiency considerations favoring extreme enforcement policies, *see, e.g.*, Gary S. Becker, *Crime and Punishment: An Economic Approach*, 76 J. POL. ECON. 169 (1968); George J. Stigler, *The Optimum Enforcement of Laws*, 78 J. POL. ECON. 526 (1970); ROBERT COOTER & THOMAS ULEN, LAW AND ECONOMICS 427–54 (3d ed. 2000); JEFFRIE G. MURPHY & JULES L. COLEMAN, PHILOSOHY OF LAW: AN INTRODUCTION TO JURISPRUDENCE 211–13 (rev. ed. 1990); MITCHELL A. POLINSKY, AN INTRODUCTION TO LAW AND ECONOMICS 75–86 (2d ed. 1989); RICHARD A. POSNER, ECONOMIC ANALYSIS OF LAW 242–50 (5th ed. 1998).

achieve optimal deterrence of any crime, a state must ensure that the crime's expected punishment equals the optimal magnitude. To set a crime's expected punishment at the optimal level, a state can always lower the crime's probability of punishment so long as it sufficiently raises the crime's severity of punishment.[5] As a consequence, an extreme enforcement policy would be the most efficient means of achieving the optimal level of deterrence of any crime because the policy would require the least resources to administer.[6]

On the one hand, an extreme enforcement policy would minimize the resources needed to set a crime's probability of punishment because such a policy would set the probability at an extremely low level. On the other hand, an extreme enforcement policy would not require any additional resources overall to set a crime's severity of punishment. Although an extreme enforcement policy would impose more severe punishments, it would also impose fewer punishments since the fewest criminals would be punished under such a policy. The resources saved under an extreme enforcement policy for a crime could be used to satisfy some people's critical needs that would otherwise go unsatisfied under a more moderate enforcement policy.[7]

In spite of the reasons of efficiency favoring an extreme enforcement policy for any type of crime, this sort of policy applied to a moderately serious crime would still seem strongly unreasonable all things considered.[8] However, it is not immediately obvious why such a policy would be so unreasonable on the assumption that the policy would be an efficient means of promoting deterrence.

In this Essay, I argue that a novel theory of deserved punishment best explains why an extreme enforcement policy would be so unreasonable when applied to a moderately serious crime.[9] Although such an extreme policy might be efficient, someone who commits a moderately serious crime would not deserve the extremely severe punishment he would receive under the policy.[10] In general, I conclude that my

5. To ensure "marginal deterrence," the severity or probability of punishment should generally be higher for more serious crimes. POSNER, *supra* note 4, at 245. Thus, considerations of marginal deterrence can affect how extreme a state should make an enforcement policy for less serious offenses.

6. As Bentham suggests, one enforcement policy might be better than another insofar as it would be a more efficient means of achieving deterrence: "The last object is, whatever the mischief be, which it is proposed to prevent, to prevent it at as *cheap* a rate as possible." JEREMY BENTHAM, AN INTRODUCTION TO THE PRINCIPLES OF MORALS AND LEGISLATION 165 (J. H. Burns & H. L. A. Hart eds., 1970) (emphasis in original).

7. Under a more moderate enforcement policy, a crime's severity of punishment would be lower and its probability of punishment would be higher than they would be under a more extreme policy.

8. *Cf.* Cass R. Sunstein, David Schkade & Daniel Kahneman, *Do People Want Optimal Deterrence?*, 29 J. LEG. STUD. 237 (2000).

9. I provide a summary of my theory of deserved punishment below. *See* text accompanying *infra* notes 27–50. For a more extensive exposition and defense of my theory of punitive desert, *see* Jim Staihar, *Proportionality and Punishment*, 100 IOWA L. REV. 1209 (2015).

10. I assume that deterrence is inevitably imperfect. So someone would inevitably commit a crime and be punished for it under any enforcement policy available.

theory of punitive desert should play an important deontological role in constraining the means by which a state may enforce its laws against any type of crime.

I. Two Principles in the Justification of Punishment

A state is justified in imposing a punishment on a criminal only if the punishment satisfies two principles. First, according to a desert principle, the criminal must deserve the punishment. If a criminal does not deserve a punishment, then imposing the punishment on the criminal would violate his rights.[11] Thus, the desert principle serves as a deontological side constraint on how much a state is morally permitted to punish an offender.[12]

Second, according to a value principle, a deserved punishment is justified only if the consequences of the punishment would be at least as good, all things considered, as the consequences of any other deserved punishment available. If the consequences of one deserved punishment would be all things considered better than the consequences of an alternative deserved punishment, then a state would have more reason to impose the former, rather than the latter, punishment on an offender.[13]

Both the desert principle and the value principle could provide different grounds on which to object to an extreme enforcement policy applied to a moderately serious crime. To determine which principle best explains why this sort of extreme policy would seem strongly unreasonable, I apply both principles to such a policy in my analysis below.

Initially, I apply the value principle. To apply the value principle, I assume for the sake of argument that someone who commits a moderately serious crime would deserve the extremely severe punishment he would receive under an extreme enforcement policy. I then argue that, on reflection, there does not seem to be a sufficiently robust reason to believe that the consequences of an extreme enforcement policy would be all things considered worse than the consequences of a more moderate policy. I conclude that the value principle does not adequately explain why an extreme enforcement policy would seem strongly unreasonable when applied to a moderately serious crime.

Next, I apply the desert principle. I argue that someone who commits a moderately serious crime does not deserve the extremely severe

11. To clarify, I assume that a criminal deserves a punishment in the negative sense that a state would not violate his rights by imposing the punishment on him against his will. *See* J. L. A. Garcia, *Two Concepts of Desert*, 5 L. & Phil. 219, 219–23 (1986) (expounding this negative sense of punitive desert). On the negative sense of punitive desert, the fact that a criminal deserves to be punished means that he has forfeited his right not to be punished. *Cf.* Christopher Heath Wellman, *The Rights Forfeiture Theory of Punishment*, 122 Ethics 371 (2012) (arguing against several objections to the claim that criminals forfeit their right not to be punished).

12. *See* Robert Nozick, Anarchy, State and Utopia 30–33 (1974) (defending a conception of rights as "side-constraints").

13. *Cf.* Staihar, *supra* note 9, at 1230–31 (arguing that in some cases, a state all things considered should not impose a deserved punishment on a wrongdoer).

punishment he would receive under an extreme enforcement policy. So such an extreme policy would be unjustified because it would violate the rights of the criminals who would be punished extremely severely under the policy. I conclude that the desert principle along with my theory of punitive desert best explains why an extreme enforcement policy would be unjustified when applied to a moderately serious crime.

II. The Value Principle

For ease of exposition, I introduce four variables. First, suppose an "m-crime" refers to a moderately serious crime, such as a simple assault. Second, assume an "m-criminal" refers to someone who commits an m-crime.

Third, suppose an "EEP" refers to an extreme enforcement policy applied to an m-crime. Under an EEP, an m-crime's severity of punishment would be set extremely high, whereas the m-crime's probability of punishment would be set extremely low.

Lastly, assume an "MEP" refers to a more moderate enforcement policy applied to an m-crime. The expected punishment of an m-crime would be the same under both an EEP and an MEP. But under an MEP, the m-crime's severity of punishment would be proportionately lower, and the m-crime's probability of punishment would be proportionately higher than they would be under an EEP.

Now assume, for the sake of argument, that m-criminals would deserve the punishments that they would receive under either an EEP or MEP. Assume also that no alternative enforcement policy would have better consequences all things considered than an EEP or MEP. Under these assumptions, to show that an EEP would violate the value principle, we must show that an EEP's consequences would be all things considered worse than an MEP's consequences.

To do so, we must identify some aspect of an EEP's consequences that would make them worse than an MEP's consequences. I consider three potential aspects below, grounded in three different conceptions of value: egalitarianism, prioritarianism, and retributivism. I argue that none of these theories of value provides a sufficiently robust reason to believe that an EEP's consequences would be all things considered worse than an MEP's.

A. *Egalitarianism*

Egalitarians assume that inequality is intrinsically bad in the sense that it would be intrinsically bad if some criminals were punished more severely than other like criminals.[14] As a corollary, egalitarians contend that it would be intrinsically bad if some criminals were punished while other like criminals were not punished at all.[15]

14. I assume like criminals have committed identical offenses.

15. In my discussion of egalitarianism, I refer only to "telic" egalitarians who claim that inequalities are objectionable because they are intrinsically bad. I do not refer to

Egalitarians might argue that an EEP would violate the value principle because an EEP's consequences would be worse than an MEP's consequences with respect to the value of equality. This is so for two reasons.[16] First, relative to an MEP, fewer m-criminals would be punished under an EEP because an m-crime's probability of punishment would be lower under an EEP. Second, the m-criminals punished under an EEP would be punished much more severely than those punished under an MEP because an m-crime's severity of punishment would be much higher under an EEP.[17]

In response, I contend that this egalitarian argument against an EEP is unpersuasive for at least three reasons. First, there is reason to doubt that inequality is ever intrinsically bad.[18] To illustrate, suppose two criminals each commit the same minor offense, such as jaywalking. A state then imposes the death penalty on one of the jaywalkers. Now the state must decide whether or not to execute the other jaywalker.

If inequality were intrinsically bad, then there would be something intrinsically good about executing the other jaywalker. For by executing the other jaywalker, the state would ensure that both receive identical punishments for their identical minor offenses. But it seems doubtful that there would be anything intrinsically good about executing the other jaywalker. Hence, it is doubtful that inequality is ever intrinsically bad.

Second, even if inequality were intrinsically bad, the egalitarian argument still does not show that the overall inequality in an EEP's overall consequences would be all things considered worse than the overall inequality in an MEP's overall consequences. This egalitarian argument focuses only on the inequality between m-criminals under an EEP and MEP. However, both policies could also affect the inequalities between innocent people, and these inequalities would likely be worse under an MEP. For the resources saved under an EEP could be used to mitigate the inequalities between innocent people by benefitting the worse off. Hence, the overall inequality in an EEP's overall consequences might be all things considered better than the overall inequality in an MEP's overall consequences, even if the inequality among m-criminals were worse under an EEP.

Third, even if an EEP's overall inequality were all things considered worse than an MEP's, this egalitarian argument against an EEP is still open to doubt. For even if inequality were intrinsically bad, it can-

"deontic" egalitarians who object to them on other grounds. Derek Parfit, *Equality or Priority?*, *in* THE IDEAL OF EQUALITY 81, 84 (Matthew Clayton & Andrew Williams eds., 2000).

16. *See* Isaac Ehrlich, *The Optimum Enforcement of Laws and the Concept of Justice: A Positive Analysis*, 2 INT'L REV. L. & ECON. 3, 10–11 (1982) (noting some of the inequalities that would result from extreme enforcement policies).

17. For detailed approaches to evaluating inequalities in different distributions of benefits or harms, *see* Larry S. Temkin, *Inequality*, 15 PHIL. & PUB. AFF. 99 (1986); LARRY S. TEMKIN, INEQUALITY (1993).

18. My argument against the intrinsic badness of inequality in the punishment of like criminals parallels "the levelling down objection" that Parfit raises against telic egalitarians more generally. *See* Parfit, *supra* note 15, at 98–99.

not be the only thing that is intrinsically bad. Suffering is also intrinsically bad, and decreases in suffering can offset increases in inequality.

An EEP would likely realize less overall suffering than an MEP for two reasons. First, both policies would realize the same overall amount of suffering among m-criminals. Because an m-crime's expected punishment would be the same under an EEP and MEP, both policies would realize the same overall amount of punishment among m-criminals.[19] Second, the resources saved under an EEP could be used to relieve the suffering of innocent people in need. Thus, even if an EEP's overall inequality were worse than an MEP's, an EEP's consequences might still be better on the whole because they would realize less overall suffering.

B. *Prioritarianism*

Prioritarians can concede that inequality between the punishments of like criminals is not intrinsically bad. Prioritarians can also concede that an EEP would realize less overall suffering than an MEP. However, prioritarians might still argue that collectively people's suffering would be worse under an EEP than under an MEP. Therefore, an EEP would be unjustified as a violation of the value principle.

According to prioritarians, for any particular amount of suffering that a person experiences, her suffering that amount is worse, the worse her life is considered as a whole.[20] So benefitting people matters more, the worse off they are. To illustrate, suppose two people suffer the same amount at some time. One person's life as a whole, though, is worse than the other person's life as a whole. According to prioritarians, the suffering of the former is worse than the suffering of the latter even though the suffering at issue in both cases is the same.

To show that collectively people's suffering would be worse under an EEP than under an MEP, prioritarians could argue that the suffering of the affected class would be worse under an EEP than under an MEP. The affected class consists of the people whose suffering stands to be affected by the choice between an EEP and MEP. Other things being equal, the affected class comprises a) the m-criminals who stand to be punished under either enforcement policy, and b) the innocent people whose suffering would be prevented by the resources saved under an EEP.

Prioritarians might assume that among the affected class, the m-criminals punished extremely severely under an EEP would have the worst lives considered as a whole. So the suffering of an m-criminal under an EEP would be worse than the comparable suffering of anyone else in the affected class under an MEP. As a consequence, even

19. The overall amount of punishment under an EEP would consist in a relatively small number of m-criminals receiving extremely severe punishments, whereas the overall amount under an MEP would consist in a larger number of m-criminals receiving moderately severe punishments. These differences should affect only the pattern of distribution, not the overall magnitude, of punishment.

20. *See* Parfit, *supra* note 15, at 101–03.

though the affected class could suffer less overall under an EEP, prioritarians might nevertheless infer that the suffering of the affected class would still be worse under an EEP.

In response, I concede that prioritarianism is a plausible theory of value. More specifically, it is a plausible view about the badness of suffering. However, this prioritarian argument against the justification of an EEP is doubtful for two reasons.

First, among the affected class, the people who stand to have the worst lives considered as a whole might not be the m-criminals who would be punished extremely severely under an EEP: they might be the innocent people under an MEP whose suffering would have been prevented by the resources saved under an EEP. In that case, the interests of the innocent who stand to benefit under an EEP should take priority over the interests of the m-criminals who stand to be punished extremely severely under an EEP.

Second, the mere fact that a person has a worse life, considered as a whole, than others does not make his suffering some amount worse than any amount of suffering that others might experience collectively or individually. The interests of the worse-off do have some priority over the interests of the better-off. But the priority is not absolute.[21]

So even if the m-criminals punished extremely severely under an EEP stand to have the worst lives as whole, the collective suffering of the affected class might still be better under an EEP than under an MEP. For the affected class might still suffer sufficiently less overall under an EEP given the benefits to the innocent from the resources saved under an EEP. Hence, I contend that the prioritarian case against an EEP is not sufficiently robust to explain adequately why an EEP would seem strongly unreasonable.

C. *Retributivism*

So far we have assumed that the suffering of an m-criminal punished under an EEP or MEP would always be intrinsically bad. Retributivists, however, contend that a criminal's suffering could be intrinsically good.[22] Depending on the seriousness of someone's crime, retributivists assume there is some optimal amount of suffering which it would be intrinsically good that the criminal experience.[23] To the extent that the criminal suffers any less or more than the optimal amount, it would be intrinsically bad.

Retributivists might presume that a moderate amount of suffering would be optimal for someone who commits an m-crime. So for three reasons, retributivists might infer that the suffering of m-criminals, con-

21. *See id.* at 116–21.

22. *See* Ehrlich, *supra* note 16, at 18–20 (noting some of the retributive costs of an extreme enforcement policy).

23. *See* Michael S. Moore, *Justifying Retributivism*, 27 ISRAEL L. REV. 15, 19–20 (1993). *Cf.* G. E. MOORE, PRINCIPIA ETHICA 82, 262–65 (Thomas Baldwin ed., rev. ed. 1993) (claiming that the organic whole of a criminal's suffering some punishment is better than the organic whole of his suffering no punishment).

sidered collectively or individually, would be worse under an EEP than under an MEP. First, fewer m-criminals would suffer the optimal amount under an EEP. Second, more m-criminals would suffer less than the optimal amount under an EEP. Third, more m-criminals would suffer more than the optimal amount under an EEP.

All three claims hold given that a) fewer m-criminals would be punished under an EEP because m-crime's probability of punishment would be lower under an EEP, and b) all the m-criminals punished under an EEP would suffer too much because they would be punished extremely severely, whereas all the m-criminals punished under an MEP would suffer only the optimal amount because they would be punished only moderately severely. Since the suffering of m-criminals would be worse under an EEP, its consequences would be all things considered worse than an MEP's. Hence, an EEP would be unjustified as a violation of the value principle.

In assessment, I contend that this retributive argument against an EEP is unpersuasive for at least two reasons. First, the core assumption of retributivism seems too controversial. Reasonable people can reject the premise that it would ever be intrinsically good that a criminal suffer any amount.[24] Reasonable people might endorse the more humane view that no one's suffering any amount could ever be intrinsically good. That is, anyone's suffering any amount under any conditions would be intrinsically bad.[25] Of course, suffering can be instrumentally good: suffering could have good effects. For example, a criminal's suffering a punishment might deter others from committing offenses in the future. But no one's suffering could be good in itself.

Second, even if retributivism were a sound doctrine, the value of ensuring that m-criminals suffer the optimal amount would be just one value among other competing values. Moreover, there seems no reason to give priority to ensuring that m-criminals suffer the optimal amount over relieving the suffering of the innocent. As a consequence, the resources saved under an EEP might be used to relieve enough suffering among the innocent to make the overall suffering of the affected class better under an EEP than under an MEP.

At this point, I conclude that the value principle does not provide a sufficiently robust explanation of why an EEP would seem strongly unreasonable. Assuming an EEP would be an efficient means of promoting deterrence, there does not seem to be a sufficiently robust reason to infer that an EEP's overall consequences would be worse than an MEP's.[26]

---

24. *See* T. M. Scanlon, What We Owe To Each Other 274 (1998).

25. *See id.*

26. Assuming that punishments under an EEP and MEP both involve a term of incapacitation, some might contend that even if an EEP and MEP have the same deterrence effect, an MEP would still prevent more crimes by incapacitating more m-criminals. But although an MEP would incapacitate more m-criminals, an MEP would presumably incapacitate m-criminals for shorter terms than an EEP. Thus, overall, it is not obvious that the additional terms of incapacitation under an MEP would result in fewer crimes than longer terms of incapacitation under an EEP.

III. The Desert Principle

On its face, the desert principle seems a more promising basis on which to object to an EEP. On reflection, someone who commits a moderately serious crime does not seem to deserve an extremely severe punishment. Such an offender would deserve no more than a moderately severe punishment. If this is the case, then an EEP is not only a violation of the desert principle, but also a particularly egregious violation of the principle. Under an EEP, m-criminals would be punished not just more severely than they deserve, but *much* more severely than they deserve.

The challenge is to explain why someone who commits a moderately serious crime deserves at most a moderately severe punishment. Ultimately, my explanation of this proportionality principle is grounded in a novel theory of deserved punishment.[27] My theory of deserved punishment is a type of unfair advantage theory of punitive desert.[28] In defending my view, I argue that a criminal incurs an obligation to undertake a punishment from committing his offense. The criminal deserves to be punished because unless he suffers a punishment, he will obtain an unfair advantage consisting in the illicit benefit of freedom from the burdens he is obligated to undertake as a consequence of committing his crime. The challenge in defending my theory of punitive desert is to explain why criminals incur an obligation to undertake certain burdens by committing their offenses.

When someone commits a crime without any exculpatory defenses, I assume he undermines his trustworthiness.[29] More precisely, he undermines the minimally acceptable degree of trustworthiness that we are warranted in demanding each other not to undermine. A person's minimally acceptable degree of trustworthiness consists in the conditions that are necessary for others' being justified in believing with a minimally acceptable credence that he is not disposed to commit crimes.

An unexcused crime undermines the offender's minimally acceptable degree of trustworthiness because it is sufficiently strong evidence of a standing deficiency in the offender's concern for the rights of others, such that he lacks a sufficiently reliable character trait not to

27. The summary of my theory of punitive desert that I provide below is based on Staihar, *supra* note 9, at 1216–23.

28. In the literature on the justification of punishment, unfair advantage theories of punitive desert are the most prevalent. *Cf.* David Boonin, The Problem of Punishment 120 (2008) (stating that "the fairness-based approach is arguably the preeminent form of retributivism in the current literature"). For other variants of an unfair advantage theory of punitive desert that have been proposed in the literature, widely discussed, and roundly criticized, *see* Staihar, *supra* note 9, at 1212–16.

29. *See, e.g.*, Susan Dimock, *Retributivism and Trust*, 16 L. & Phil. 37, 53 (1997); David A. Hoekema, *Trust and Obey: Toward a New Theory of Punishment*, 25 Isr. L. Rev. 332, 345 (1991); Jim Staihar, *Punishment as a Costly Signal of Reform*, 110 J. Phil. 282, 284 (2013). Because trustworthiness comes in degrees, even repeat offenders undermine their trustworthiness to an additional degree by committing their later crimes.

commit crimes.[30] Hence, for a range of comparably serious crimes and situations, the offender poses an unacceptably high risk of committing such crimes over a significant run of such situations.

Unless a criminal restores his trustworthiness, he unacceptably risks causing others certain especially significant harms.[31] For example, an untrustworthy criminal unacceptably risks committing a range of other offenses in the future. In addition, an untrustworthy criminal also unacceptably risks causing others to incur certain costs of insecurity, which would constitute harms to others even if the criminal were never to commit another offense.[32]

Three costs of insecurity seem especially salient. First, others might rationally need to invest in costly precautionary measures to protect themselves from an untrustworthy offender. For example, they might need to engage in costly monitoring of the criminal and to invest in costly protective services when interacting with him is unavoidable.

Second, others might rationally need to forgo pursuing some personally and socially valuable activities that would leave them too vulnerable to an untrustworthy offender. In other words, people might rationally need to reduce their activity levels in response to the criminal. Third, others might rationally experience higher levels of fear in response to the higher risk of the offender's committing crimes again.[33] Given the significance of these costs of insecurity to the lives of others, I suggest that these costs would constitute harms that others have a right against incurring.

To avoid the unacceptable risk of causing others to incur the relevant harms, a criminal incurs an obligation to restore his trustworthiness expeditiously.[34] Given this obligation of restoration, the offender is obligated to undertake any burdens necessary to restore his trustworthiness to a minimally acceptable degree. Assuming the criminal must undertake certain burdens to do so, the state may impose them on him as a punishment against his will without violating his rights. For unless

30. *Cf.* RICHARD B. BRANDT, *Blameworthiness and Obligation*, *in* ESSAYS IN MORAL PHILOSOPHY 3, 14, 16–17, 32 (A. I. Melden ed., 1958) (arguing that if someone is blameworthy for performing an act, then his performing it warrants our inferring a motivational defect in his character).

31. *See* Staihar, *supra* note 29, at 283.

32. *Cf.* THOMAS HOBBES, LEVIATHAN 86–90 (Richard Tuck ed. 1991) (noting the costs of insecurity that people rationally must incur in response to being justified in believing that others are disposed to engage in acts of aggression).

33. I assume that insofar as an offender poses a higher risk of committing further crimes, then others will rationally believe with a higher subjective probability that the offender will commit further crimes. *See* David Lewis, *A Subjectivist's Guide to Objective Chance*, *in* PHILOSOPHICAL PAPERS: VOLUME II 83 (1986) (discussing the concept of subjective probabilities and their relation to objective probabilities).

34. A criminal is obligated to restore his trustworthiness *expeditiously* because the longer he takes to restore it, the longer he will pose an unacceptable risk to others. For a statement of the importance of restoring trustworthiness in the more general context of reparations, *see* MARGARET URBAN WALKER, WHAT IS REPARATIVE JUSTICE? 25 (2010) (writing that "[t]he gesture of reparations needs to model the kind of relationship between victims and responsible parties that creates a new or renewed basis of trust for the future, precisely what was lacking in the circumstances in which the wrong was done").

the criminal suffers such burdens, he will obtain an illicit benefit consisting in his freedom from the burdens necessary to fulfill his obligation of restoration.

According to the main principle of my theory of punitive desert, a criminal deserves a punishment for his crime that is proportional to the burdens he must undertake to fulfill the obligation of restoration he incurs from committing his crime. In other words, a criminal deserves a punishment that is proportional to the burdens he is obligated to undertake to restore his trustworthiness to a minimally acceptable degree.[35] Once the criminal undertakes a punishment proportional to such burdens, he deserves no more punishment for his offense.[36]

Now I argue that a criminal must in fact undertake some burdens to restore his trustworthiness. To restore it, I suggest that the offender must signal his reform. Such a signal[37] would be a directly observable property which is sufficiently strong evidence that the criminal has rectified the deficiency in his concern for others that he manifested in committing his crime.[38] Offenders must restore their trustworthiness by signaling their reform because people are unavoidably vulnerable to each other under any acceptable system of criminal justice available.[39] No acceptable means of deterrence or incapacitation available can adequately reduce the risks that an untrustworthy criminal poses to others.[40] Moreover, there is no reliable way for others to induce reform

35. As a corollary, a criminal does not deserve a punishment for his crime that is more severe than the burdens he must undertake to fulfill the obligation of restoration he incurs from committing the crime. In other words, a criminal does not deserve to be punished for his crime more severely than the burdens he is obligated to undertake to restore his trustworthiness to a minimally acceptable degree.

36. So after the state punishes an offender in proportion to the severity of the burdens necessary to satisfy his obligation of restoration, the state is not morally permitted to punish the offender further for the relevant offense.

37. In general, a signal is a directly observable property that is strong evidence of its bearer's possessing another property that is not directly observable. *Cf.* Michael Bacharach & Diego Gambetta, *Trust in Signs*, *in* TRUST IN SOCIETY 148, 159 (Karen S. Cook ed., 2001).

38. *See* Staihar, *supra* note 29, at 287. By rectifying the deficiency in his concern for others, the offender would develop a good will. As Annette Baier states, when we trust others, we are confident they have a good will toward us; therefore, "[r]easonable trust will require good grounds for such confidence in another's good will. . . ." Annette Baier, *Trust and Antitrust*, 96 ETHICS 231, 235 (1986).

39. *See* Staihar, *supra* note 29, at 284–87. *Cf.* Jeffrie Murphy, *Marxism and Retribution*, 2 PHIL. & PUB. AFF. 217, 231–32 (1973) (noting that "our moral language presupposes . . . that we are vulnerable creatures—creatures who can harm and be harmed by each other").

40. To clarify, a system of criminal justice can be expected to have a general deterrence effect. And a state can promote this general deterrence effect more or less efficiently depending on how it enforces its criminal laws. But once someone commits a crime, no acceptable enforcement policy available can provide others with the needed assurance that this specific criminal will be deterred from committing crimes again. For discussion of well-known problems with trying to achieve deterrence, *see, e.g.*, Paul H. Robinson & John M. Darley, *Does Criminal Law Deter? A Behavioral Science Investigation*, 24 OXFORD J. LEG. STUD. 173 (2004); Paul H. Robinson & John M. Darley, *The Role of Deterrence in the Formulation of Criminal Law Rules: At Its Worst When Doing Its Best*, 91 GEO. L. J. 949, 954, 992–94 (2003); Staihar, *supra* note 29, at 284–85.

in a criminal, and none seems forthcoming.[41] Hence, there is a need for a sign from the offender himself that he has come to develop a sufficiently high degree of concern for the interests of others.[42]

To be credible, I suggest that a sign of reform must be costly.[43] A criminal cannot signal his reform through mere costless means, such as merely apologizing for his crime or pleading a change of heart. Such "cheap talk" is not credible because criminals who do not care at all about others would be willing to convey it.[44] So a credible sign of reform must be too costly for criminals who have not rectified the revealed deficiency in their concern for others.[45] In other words, a credible sign of reform must be too costly for criminals who have not come to care sufficiently about the interests of others by developing a sufficiently benevolent character. Thus, to demonstrate reform, I assume a criminal must send others a costly signal that he has developed a sufficiently benevolent character.[46] In general, benevolence is a trust warranting property that is inconsistent with the sort of insufficient concern typical of criminals.[47]

To demonstrate the development of a sufficiently benevolent character, I contend that a criminal must signal that he has acted with a sufficiently high degree of benevolence for a sufficiently long time after

41. *See, e.g.*, MICHAEL R. GOTTFREDSON & TRAVIS HIRSCHI, A GENERAL THEORY OF CRIME 268–269 (1990); Staihar, *supra* note 29, at 286–87. Absent an exculpatory defense, there is no form of clinical treatment that criminals can undergo which would provide others with the needed assurance of reform.

42. At this point, I note that my theory of punitive desert is a practical theory. That is, I seek to explain why and how much offenders deserve to be punished in the actual world given the natural facts that generally characterize the unavoidable conditions under which people actually live. So my theory presumes that there are no extraordinary means available of obtaining epistemic access to or changing an offender's disposition to commit crimes. *Cf.* JOHN RAWLS, A THEORY OF JUSTICE 157–61 (1971) (developing a theory of distributive justice with a similarly practical aim and its own presuppositions about the natural facts under which it applies).

43. *See* Staihar, *supra* note 29, at 287–28. The concept of a costly signal has wide interdisciplinary application. *See, e.g.*, A. MICHAEL SPENCE, MARKET SIGNALING (1974) (signaling theory in economics); AMOTZ & AVISHAG ZAHAVI, THE HANDICAP PRINCIPLE: A MISSING PIECE OF DARWIN'S PUZZLE (Naama Zahavi-Ely & Melvin Patrick Ely trans., 1997) (discussing signaling theory in biology); ERVING GOFFMAN, THE PRESENTATION OF SELF IN EVERYDAY LIFE (rev. ed. 1959) (discussing signaling theory in sociology); DOUGLAS G. BAIRD, ROBERT H. GERTNER & RANDAL C. PICKER, GAME THEORY AND THE LAW 122–58 (1994) (discussing signaling theory in law). For an elementary game-theoretic analysis of signaling, *see* AVINASH DIXIT & SUSAN SKEATH, GAMES OF STRATEGY 263–310 (2d. ed. 2004). For an application of signaling theory specifically to trust, *see* Bacharach & Gambetta, *supra* note 37.

44. For analyses of the credibility conditions on cheap talk, *see, e.g.*, Joseph Farrell, *Meaning and Credibility in Cheap-Talk Games*, 5 GAMES & ECON. BEHAV. 514 (1993); Joseph Farrell & Matthew Rabin, *Cheap Talk*, 10 J. ECON. PERSP. 103 (1996); Robert Stalnaker, *Saying and Meaning, Cheap Talk and Credibility*, *in* GAME THEORY AND PRAGMATICS 83 (Anton Benz, Gerhard Jäger & Robert van Rooij eds., 2006).

45. This is the "nonpooling condition" on the credibility of a costly signal of reform. Bacharach & Gambetta, *supra* note 37, at 160; Staihar, *supra* note 29, at 289.

46. *See* Staihar, *supra* note 29, at 287.

47. *See, e.g.*, Bacharach & Gambetta, *supra* note 37, at 154 (noting benevolence as a trust-warranting property); Staihar, *supra* note 29, at 287 (noting benevolence as a trust-warranting property).

committing his crime. To show that he has acted with such benevolence, the offender must sacrifice some of his sufficiently important personal interests for a sufficiently long time for the sake of benefiting others. To make such a sacrifice for others, I suggest the offender must standardly engage in labor-intensive community service, and he usually must do so under reasonable conditions of incapacitation to mitigate the risk he poses to others while the service is performed. The more service the criminal performs for the sake of benefiting others, the stronger it will serve as evidence that he has rectified the revealed deficiency in his concern for the interests of others.[48]

Hence, to fulfill his obligation of restoration, and thereby restore his trustworthiness to a minimally acceptable degree, a criminal must undertake some burdens.[49] The offender deserves to be punished in proportion to those burdens because unless he suffers in proportion to them, he will obtain an illicit benefit consisting in his freedom from the burdens necessary to fulfill the obligation of restoration he incurs from committing his crime. In summary, under my theory of punitive desert, the absolute severity of the most severe punishment that a criminal deserves for an offense corresponds to the absolute severity of the burdens that he must undertake to fulfill the obligation of restoration he incurs from committing the offense.

Now consider someone who commits a moderately serious crime, such as a simple assault. How much punishment does the offender deserve under my theory of punitive desert? A moderately serious crime reveals only a moderately bad deficiency in the criminal's concern for the rights of others. To provide others with a credible sign of his reform, I suggest that the criminal need only sacrifice some of his moderately important personal interests for a moderately long time for the sake of benefiting others.[50] Thus, I suggest that the criminal is obligated to undertake only a moderately severe burden in order to restore his trustworthiness to a minimally acceptable degree. As a consequence, under my theory of punitive desert, the criminal deserves no

48. *Cf.* KEALLY MCBRIDE, PUNISHMENT AND POLITICAL ORDER 136 (2007) (reporting that "prisoners who work in prisons are 24 percent less likely to return to prison after release"); LINDA RADZIK, MAKING AMENDS: ATONEMENT IN MORALITY, LAW, AND POLITICS 99 (2009) (writing that "[t]he greater the sacrifice that is required to make the reparation payment, the more evidence we have that the wrongdoer is remorseful for her past action"); Staihar, *supra* note 29, at 287.

49. Like other trust building or maintaining processes, the process of a criminal's restoring his trustworthiness has a multi-layered inferential structure. Bacharach & Gambetta, *supra* note 37, at 162. In addition to undertaking the required burdens, other steps might also be necessary to restore an offender's trustworthiness, such as the offender apologizing for his crime and compensating any victims. The criminal might need to undergo some form of therapy and take steps to eliminate aspects of his situation that pressure him to commit crimes, such as unemployment, corrupting social influences, and problems of addiction. Much will depend on the specifics of the case. However, because these other steps are not necessarily burdensome for the criminal, they need not be part of his punishment properly understood.

50. For a more detailed illustration of the sorts of personal sacrifices that a criminal might make to signal reform, *see* Staihar, *supra* note 9, at 1223–24.

more than a moderately severe punishment for committing his moderately serious crime.

By explaining why someone who commits a moderately serious crime would not deserve to be punished extremely severely, I suggest that my theory of punitive desert best explains why an EEP would be unjustified. For I contend that my theory best explains why an m-criminal would not deserve the extremely severe punishment that he would receive under an EEP.

IV. Critical Discussion

Drawing on the form of my theory of punitive desert, critics might attempt to construct a similar argument for why an m-criminal would in fact deserve an extremely severe punishment under an EEP. Critics might reasonably assume that an EEP would be the most efficient means of promoting deterrence. Critics might then infer that an m-criminal would incur an obligation to undertake an extremely severe punishment under an EEP as a mean to deterring others most efficiently. Assuming an m-criminal were obligated to undertake an extremely severe punishment under an EEP, he would deserve such a punishment.

In response, I deny that a criminal incurs an obligation to undertake a punishment as a means to deterring others, efficiently or otherwise. I assume that a person is specially responsible for the consequences she causes rather than merely allows to occur.[51] In general, a person has an obligation not to cause others harm. But a person does not have a general obligation not to allow others to be harmed.

So we might concede to the critics that if an m-criminal were not to undertake an extremely severe punishment under an EEP (or under any other enforcement policy), then others would commit additional crimes who would have been deterred by the m-criminal's extreme suffering. But the m-criminal in such a case would merely allow these other criminals to commit their additional offenses. The m-criminal would not cause them to commit those offenses. Thus, the m-criminal would incur no obligation to undertake a punishment as a means to deterring others.[52]

Conclusion

Our critical analysis of an extreme enforcement policy applied to a moderately serious crime illustrates the importance of the desert principle and, more specifically, my theory of punitive desert in constraining

51. *See, e.g.*, Bernard Williams, *A Critique of Utilitarianism*, *in* Utilitarianism For and Against 75, 93–95, 99 (J. J. C. Smart & Bernard Williams eds., 1973); Warren Quinn, *Actions, Intentions, and Consequences: The Doctrine of Doing and Allowing*, 98 Phil. Rev. 287 (1989); Samuel Scheffler, *Doing and Allowing*, 114 Ethics 215 (2004).

52. Under my theory of punitive desert, an m-criminal incurs an obligation to restore his trustworthiness because unless he restores it, he unjustifiably risks causing others certain harms, namely his commission of additional crimes and the costs of insecurity in response to his own untrustworthiness.

the types of enforcement policies that a state is morally permitted to adopt for any type of crime. Without the desert principle along with my theory of punitive desert, I suggest that there is no robust reason for a state not to adopt the kinds of extreme enforcement policies that could be the most efficient means of promoting deterrence even for moderately serious offenses. Thus, under the deontological constraint of punitive desert within my theory, a state might need to adopt less efficient, more costly, enforcement policies for its criminal laws. But rather than cause for complaint, these inefficiencies are simply the costs of respecting the rights of persons. As such, these additional costs are morally acceptable.