Efficiency and Fairness in Criminal Law: The Case for a Criminal Law Principle of Comparative Fault

Alon Harel†

Whether based on deterrence, retribution, or rehabilitation, the traditional justifications of criminal law penalties focus exclusively on the actual or potential criminal. This Article explores the theoretical foundations of criminal law and argues that to focus solely on the criminal is misguided, both as a normative matter and as an accurate description of the criminal law. The author posits that the alleged absence of the comparative fault principle of tort and contract law in criminal law is a puzzle in need of clarification. He proposes that criminal law should, and in fact does in some instances, incorporate a principle of comparative fault. According to a criminal law principle of comparative fault, criminals who act against careless victims would be exculpated or would have their punishment mitigated. The author argues that introducing a criminal law principle of comparative fault will promote both efficiency and fairness. It will promote efficiency by providing victims an incentive to take precautions against crime. It will promote fairness because it will lead to equality in the resources invested in the protection of cautious and careless victims. Thus,

---

† Ben Gurion Scholar in Residence, University of Southern California Law Center; Lecturer, Faculty of Law, Hebrew University, Jerusalem, Israel; LL.B. 1983, Hebrew University; LL.M. 1985, Hebrew University; D. Phil. 1989, Oxford; LL.M. 1989, Yale University. I wish to thank Peter Arenella, Dan Bain, Maya Bar Hillel, Eyal Benvenisti, Meir Dan Cohen, Celia Fassberg, Malcolm Feeley, Richard Friedman, Ruth Gavison, Steve Goldstein, Zohar Goshen, Miri Gur Arye, Itir Haviv, David Heyd, Mark Kelman, Michael Koren, Motta Kremnizer, Nicola Lacey, Andrei Marmor, David Nelkin, Ariel Porat, Richard Posner, Uriel Procaccia, Yoram Schachar, Leslie Sebba, Mike Sinclair, Keith Smith, and Alex Stein for their important comments. Special thanks to Dick Craswell for invaluable criticisms of Part III and Lauren Underkuffler for her important contributions on numerous drafts. I am also grateful to the legal workshops at the University of Pennsylvania School of Law, Yeshiva University Cardozo School of Law, the University of Southern California Law Center, the Duke University Legal Workshop, the Israeli Interdisciplinary Forum at Van Leer Institute, the Hebrew University Legal Workshop, the Haifa University Legal Workshop, the New York Law School Legal Workshop and the UCLA Legal Workshop. I benefitted enormously from the discussions in these forums. At a late stage, I benefited from comments made by several participants of the USC workshop including Scott Altman, Ed McCaffrey, Mike Shapiro, and, in particular, the commentator at the USC workshop—Jennifer Arlen. Those comments could not be integrated into the Articles, but will influence any of my future work on this subject in the future. Finally, Yuval Tal and Joshua Peleg provided excellent research assistance.
the efficient provision and fair distribution of protection require a selective introduction of a principle of comparative fault into criminal law.

It is not the mouse that is the thief but the hole.¹

INTRODUCTION

Criminal law is traditionally described as directing its injunctions exclusively to actual or potential criminals. Under this description, criminal law norms are aimed at influencing the behavior of criminals or potential criminals, but not that of victims or potential victims of crime. This Article will argue that the traditional view is both normatively unjustified (on efficiency and fairness grounds) and descriptively misleading. It is normatively unjustified because providing protection in an efficient manner and distributing protection in a fair manner require the criminal law system to address its injunctions to victims as well as criminals. It is descriptively misleading because, despite the traditional perception which regards criminal law norms as oriented solely towards influencing the behavior of criminals, various criminal law doctrines in fact can be more accurately interpreted as aimed at influencing the behavior of potential victims of crime.²

Before examining and criticizing criminal law’s traditional focus on criminals, one needs to explore the rationales behind criminal sanctions. Two of the traditional justifications for the imposition of criminal sanctions, deterrence and incapacitation, can be better explained as derivations of a single, more fundamental justification: the societal effort to guarantee the efficient production of protection, and its fair distribution among potential victims.³ Using a different terminology, we may state that the aim of criminal sanctions is to produce and distribute protection to potential victims.⁴

¹. See Nashim, Gitin, in The Babylonian Talmud 45a (Rabbi Dr. I. Epstein ed. & Maurice Simon trans., 1936).
². This argument attempts to restore symmetry between criminal law and other fields of the law. Contract law issues directives to victims of breached contracts as well as to those who have committed breaches. See infra text accompanying note 11. Tort law treats both tort victims and tortfeasors as autonomous agents whose actions or omissions may have contributed to the injury at issue. The legal doctrine which best expresses this principle in tort law is the doctrine of comparative fault. See infra text accompanying note 12.
³. Admittedly, my analysis cannot explain two important traditional rationales of punishment, namely retribution and rehabilitation. Although those rationales are not necessarily incompatible with my analysis, they are not easily subsumed or explained within its framework. I attempt to demonstrate the compatibility of my conclusions with retributive motivations in Part V.
⁴. An individual’s degree of protection can be defined as the expected harm from crime for that individual, that is, the probability of harm multiplied by its size. See infra text accompanying note 59.
Criminal sanctions become much less mysterious when they are described in the same manner as the public provision of health services or educational services: namely, as a means by which the state produces and distributes a good. I shall call this perspective of criminal law the productive-distributive paradigm.

The productive-distributive paradigm may seem innocuous, but it invites us to explore two fundamental issues that traditional terminology obscures. The first concerns the best, that is, the most efficient means of providing protection. The second concerns the fair, that is, equal distribution of protection among different potential victims.

To disregard the victim’s conduct in determining the sanctions of criminals, it will be argued, is both inefficient and unfair. It is inefficient because dismissing the behavior of the victim as irrelevant to the concerns of the criminal justice system does not provide optimal incentives for victims to take precautions against crime. In contrast, under a criminal law principle of comparative fault, according to which criminals who act against careless victims would be exculpated or have their punishment mitigated, criminal law would increase the risks to careless (potential) victims by inducing criminals to act against them, and consequently provide such victims with additional incentives to take precautions. Furthermore, it is unfair to disregard the victim’s conduct in criminal cases because given the greater likelihood that careless potential victims (relative to cautious ones) will become actual victims of crime, the expected costs of protecting careless victims are higher than the expected costs of protecting cautious ones. Hence, under the current system, cautious victims are exploited for the sake of protecting careless ones.5

The normative conclusions of the inquiry into the role of victims in criminal law provide an incentive to explore whether the official position of criminal law—denying the relevance of the victim’s behavior—accurately reflects the realities of criminal law. Select criminal law doctrines can be reinterpreted in a way that reveals the role played by the victim’s conduct in determining the punishment imposed on the criminal. For example, the

5. While distributive justice concerns require us to explore the victim’s conduct in determining criminal sanctions, they also illustrate the limits of the relevance of the victim’s conduct. Distributive justice considerations can clarify why in some contexts, such as provocation in homicide, the prior behavior of the victim is a reason to mitigate the punishment, while in other contexts, such as the prior conduct of a rape victim, the behavior of the victim should have no such relevance. See infra Part III.
doctrine of provocation in homicide will be redescribed as one instance where criminal law takes the victim’s provoking behavior as a relevant factor in determining the degree of culpability, and therefore the punishment, of the criminal. The classification of property offenses into burglary, robbery, and theft will be reinterpreted in a similar manner. The greater the precautions taken by a victim of property offenses (e.g., locking the property at home in the case of burglary or keeping it in one’s physical possession in the case of robbery), the more severe the sanctions that are imposed upon the criminal. To the extent that more severe sanctions increase deterrence, the potential victim benefits from greater protection by the law.

Exposing the operation of a criminal law principle of comparative fault in certain areas of criminal law raises questions concerning the overall aim of this Article. Is it aimed at expanding the scope of the principle of comparative fault to new terrains of criminal law or merely at exposing the existing underlying reality of criminal law?

A full reply to this principal question cannot be provided at this stage. Distributive justice considerations severely constrain the desirable scope of the criminal law principle of comparative fault. Those considerations illustrate that despite the theoretical importance of the principle of comparative fault, its actual implementation should be highly restricted. Thus, this Article need not necessarily be regarded as proposing radical reforms, or indeed any reforms in criminal law. It rather can be regarded as an attempt to explore the theoretical foundations of criminal law. Most importantly, this Article shows that the alleged absence of a criminal law principle of comparative fault in many areas of criminal law is not a trivial phenomenon, but rather a puzzle which requires clarification. This clarification may be achieved both by exposing—even in the face of denial—the actual existence of a criminal law principle of comparative fault, and by showing that where it is absent, its absence is justified by the distributive justice considerations elaborated in detail in this Article.

The reader may protest against the perceived heartlessness of this analysis. The victim is (morally) innocent, while the criminal is not, and it is felt that the criminal justice system should reflect this asymmetry. Putting the behavior of the criminal on a par with the victim’s behavior, one might argue, not only distorts criminal law, but also brutally violates entrenched moral sensitivities. I will show, however, that such moral indignation is ill-founded. Part V will illustrate that where distributive justice considerations are not determinative, taking into account the victim’s conduct in determining the sanctions imposed on the criminal is far more consonant with our underlying moral understandings than is a system which refuses to do so.

6. Provocation is a partial defense to criminal homicide. See infra text accompanying notes 77-82.
The traditional view describing criminal law as directing its injunctions exclusively at actual or potential criminals is misguided. According to my proposal, criminal law should address both criminals and victims in a way that acknowledges the important roles of efficiency and distributive justice in criminal law reasoning. Under this model, criminal law would be described as a mechanism aimed at providing incentives to perpetrators as well as victims of crime to prevent crime in an efficient manner and to distribute its costs fairly.

I
THE COMPLEX CHOICE OF AGENT IN LAW

In this Part, I shall explore and analyze a major asymmetry between private law norms and criminal law norms. While private law norms are often addressed to both parties to a civil transaction (e.g., tortfeasors and tort victims), criminal law norms seem to address only one actor in the criminal encounter, namely the criminal. Various alleged justifications for the peculiar structure of criminal law norms will be explored and rejected in this Part. First, the claim that the illegitimate nature of criminal behavior can explain the focus of criminal law norms being solely on the criminals will be rejected. Second and more importantly, this Part will reject the claim that two of the traditional rationales of punishment, deterrence and incapacitation, justify the exclusive focus of criminal law norms on criminals. These rationales are derived from the more fundamental goal of criminal law, namely the efficient provision of protection and its fair distribution among potential victims. Therefore, this Part concludes that the focus of criminal law norms on the criminal should be evaluated on the basis of efficiency and distributive justice considerations.

Legal norms are directed at individuals who are expected to conform their behavior to these norms, typically for the benefit of third parties. The individuals to whom the norms are addressed and whose behavior is to be modified may be termed “agents” and those who benefit from the modified behavior of the agents may be termed “beneficiaries.” Determination of who are agents and who beneficiaries should be made in light of the functions of the legal system.

One of the law’s primary functions is to induce agents to behave in a way which will bring about a socially desirable state of affairs. There are two strategies through which the policy maker can fulfill this function. One method is to change the preferences of individuals in ways that will bring about desirable behavior. The other strategy consists of changing the payoffs attached to behavior in a way that will lead the agent to behave in a

7. In the case of paternalistic norms—norms aimed at benefitting the same individuals who are required to conform to them—the agents and the beneficiaries are the same individuals.
socially desirable manner. Most law and economics scholars choose to regard law as a means of changing payoffs rather than changing preferences. Thus, under this view, the law's primary function is to attach payoffs to actions in order to induce agents to behave in a way that will bring about a socially desirable state of affairs.

The agent whose behavior is to be influenced may be identified as any one of a variety of individuals. A car accident, for instance, can be prevented by the pedestrian; by the driver; by the engineer who built the road; and by the police, who can invest more resources in the detection of traffic offenders. A breach of contract may be prevented, or its costs to the parties minimized, by the breaching party; by the victim of the breach, who can insure herself; and by the state, which can criminalize the breach of contract. A theft could be prevented by the thief; by the victim of theft, who could invest in better precautions; by the neighbor, who can call the police; and by the police themselves, who can invest more resources in patrolling the area.

Examples drawn from diverse fields of the law illustrate the complexity of the choices the legal system faces in determining who are agents. Practitioners as well as scholars of contract law, tort law, and other fields of

---

8. The distinction between influencing behavior by changing preferences and influencing behavior by changing payoffs may collapse under certain circumstances. It has often been observed that the payoffs involved in satisfying a preference influence the intensity or even the very existence of the preference. The adaptation of preferences to what seems possible and accessible (the sour grapes syndrome) is one example of such a phenomenon. The disutility involved in satisfying the preference makes the otherwise desirable object less attractive for the agent. See Jon Elster, Sour Grapes: Studies in the Subversion of Rationality 109-40 (1983); Cass R. Sunstein, Legal Interference With Private Preferences, 53 U. Chi. L. Rev. 1129, 1146-50 (1986). The counteradaptive preferences phenomenon ("the grass is always greener on the other side of the fence") is another example illustrating the difficulty of maintaining a strict distinction between changing behavior by changing preferences and changing behavior by changing the payoffs attached to behavior. In the case of counteradaptive preferences, the disutility involved in satisfying the preference for an otherwise undesirable object makes that object desirable for the agent. See Elster, supra, at 111-12; Jon Elster, Introduction, in Rational Choice 21 (Jon Elster ed., 1986). Both phenomena should be regarded as exceptional. Maintaining the distinction between preferences and payoffs has proved to be extremely valuable for the purposes of the economic analysis of law.

9. The emphasis on changing payoffs rather than preferences is based on powerful political reasons. Changing preferences is much more intrusive than changing payoffs. To change preferences is to change the individual herself, while to change payoffs is only to change the circumstances surrounding the person.

For an exception to this general approach, see Kenneth G. Dau-Schmidt, An Economic Analysis of the Criminal Law as a Preference-Shaping Policy, 1990 Duke L.J. 1 (arguing that criminal law aims to change behavior not only by changing payoffs, but also by changing preferences).

10. The forms that incentives take vary in different legal contexts, but they all promote a desirable state of affairs by changing the rewards of different actions. Legal sanctions are part of a broader enforcement system which includes other equally important means of regulating behavior. For instance, in tort law the efforts that plaintiffs invest in suing tortfeasors may influence tortfeasors' payoffs. Similarly, in criminal law the use of discretionary powers of judges or juries, or the use of discretionary powers of the executive in identifying and prosecuting criminals, all may affect the payoffs of socially undesirable behavior. All components of the enforcement system can be described as mechanisms for changing payoffs in order to influence behavior.
private law have developed complex doctrines identifying the agents to whom legal norms are directed under different circumstances. Contract law, for example, ordinarily directs its injunctions to the breaching party, but sometimes it addresses the aggrieved party, such as when it limits the rights of the plaintiff to collect for breach of contract when damages were avoidable.\footnote{By refusing to fully compensate the nonbreaching party for a loss that he could have avoided or mitigated, contract law induces the party to reduce damages to the extent they are within his control. For a discussion of the mitigation doctrine, see E. Allan Farnsworth, Contracts 896-907 (2d ed. 1990).} Tort law primarily addresses the potential tortfeasor, but also addresses the victim through the related doctrines of contributory negligence and comparative fault.\footnote{Contributory negligence bars recovery by a negligent plaintiff even when the defendant is at fault. Comparative fault, on the other hand, permits negligent plaintiffs to recover partially under certain circumstances. See infra text accompanying note 31.} Thus, both contract and tort law conceive of the parties to a contractual transaction or to a tort incident simultaneously as agents and beneficiaries.

Criminal law seems to be a striking exception to this paradigm. Scholars and practitioners of criminal law typically posit that the behavior of the victims of a crime is irrelevant for the purposes of criminal law (unless of course it constitutes an independent offense).\footnote{See, e.g., Dickerson v. State, 441 So. 2d 536, 538 (Miss. 1983) ("Contributory negligence is not a defense to manslaughter. All that the state must prove with respect to the victim is that he was prior to the incident a live human being. The homicide laws of this State protect all living beings within the jurisdiction, sinners as well as saints, drunks as well as deacons.").} Only the criminal is depicted as an agent whose behavior is to be modified by the criminal justice system.

The different attitudes of criminal law and private law with respect to identification of agents and beneficiaries can be clarified by distinguishing reciprocal from nonreciprocal norms. Nonreciprocal norms are directed to one class of individuals, to whom I have referred as agents (e.g., criminals), for the sake of another group of individuals, to whom I have referred as beneficiaries (e.g., victims of crime). Reciprocal norms, on the other hand, direct their messages to various groups of individuals who function simultaneously as agents (to whom the norms are addressed) and as beneficiaries (who benefit from the compliance of others).\footnote{Some criminal law scholars regard the absence of such a doctrine as a unique, essential feature of criminal law which distinguishes it from other areas of the law. See George P. Fletcher, Corrective Justice for Moderns, 106 Harv. L. Rev. 1658, 1673 (1993) (book review) ("The distinctive feature of criminal liability is that, in principle, the victim’s contributory fault is irrelevant to liability.").} This distinction raises the following puzzle: why does private law consist largely of reciprocal norms,
while criminal law norms generally seem to be nonreciprocal? Can we justify the exclusive use of nonreciprocal norms in criminal law?

By asking these questions, and by raising the possibility of broadening the scope of reciprocal norms in the criminal law, I follow Coase's path in tort law. Coase suggested that the traditional causal paradigm dominating tort law is misleading because it conceptualizes the tortfeasor's activity as the cause of the victim's harm. By contrast, Coase argued that both the tortfeasor's activity and the victim's activity should be regarded as causes of harm.15 This Article can be regarded as an extension of this Coasian insight from tort law to criminal law.

One explanation for the nonreciprocal nature of criminal law norms is based on an allegedly major difference between criminal law norms and private law norms. Private law aims to reconcile the legitimate conduct of two or more individuals; hence, the conduct of both parties may need modification. On the other hand, criminal law deals with a conflict between legitimate conduct (that of the victim) and illegitimate conduct (that of the criminal). It is only natural to focus our attention upon the criminal, whose conduct is illegitimate and hence needs to be modified.

This explanation is deficient in several respects. First, it is debatable whether conduct constituting a tort or a breach of contract is legitimate. Second, the objection seems to assume what needs to be proven: that it is legitimate for a potential victim to fail to take precautions against crime. Questioning the nonreciprocal nature of criminal law norms can be interpreted as raising doubts regarding the validity of this assumption. Third, adopting a scheme of reciprocal norms in the criminal law context would not necessarily imply the illegitimacy of the victim's negligent behavior. Instead, it might reflect the reluctance of society to protect legitimate conduct of the victim if protecting his conduct is too costly.16

A second explanation for the nonreciprocal nature of criminal law norms flows from the functions of the criminal justice system as traditionally defined. Textbooks of criminal law often state four theories concerning the major functions of punishment: general deterrence, specific deterrence, retribution, and reformation (or rehabilitation).17 Sometimes, incapacitation is also listed among the primary functions of criminal law. All of these theories focus exclusively on the actual or potential criminal. Deterrence (either general or specific) is understood as the effort to deter a potential

---


16. I return to this idea in Part V. It is common in various areas of private law to expose potential victims to risks stemming from the legitimate activities of those victims. For example, tort law refuses to compensate victims of tort whose legitimate activities are too costly, as where the victim exposes herself voluntarily to certain risks ("assumes the risk").

criminal from committing crimes in the future. The retributivist, on the other hand, is interested in inflicting suffering on the actual criminal, who deserves punishment. The advocate of rehabilitation theory hopes to reform the criminal and return him to society. Finally, the advocate of incapacitation theory aims at preventing the criminal from committing future crimes by restricting his opportunities to do so. Under all of these theories, criminal law is interested in the acts and the mental states of criminals (potential or actual) and regards victims' acts and mental state as irrelevant to the concerns of criminal law.

However, at least some of the traditional rationales, in particular deterrence and incapacitation, may be better explained as manifestations of a more basic aim, namely the provision of protection to potential victims. This interpretation of deterrence and incapacitation highlights questions that are not readily formulated in the traditional terminology. For instance, are deterrence and incapacitation of criminals the only or the most efficient ways to protect victims? Can protection of victims be better achieved by perceiving the victim (in addition to the criminal) as an agent? If potential victims can better be protected under such a scheme, how should it affect the legislature's or the judiciary's construction of the criminal law? Moreover, if indeed the ultimate function of criminal law is to protect potential victims, how should protection be distributed among potential victims? Should distributive justice considerations—in particular, distribution of protection among different classes of potential victims—influence criminal law? The intelligibility of these questions depends upon our willingness to subsume the traditional notions of deterrence and incapacitation within a more comprehensive aim: the protection of potential victims from crime.

The productive-distributive paradigm brings to the fore two fundamental issues not likely to be raised by someone using the traditional terminology. The first concerns the best (most efficient) means of protecting potential victims—means which do not necessarily rely upon deterrence or incapacitation of the criminal. The second concerns the proper distribution

18. For a good discussion of deterrence, see generally FRANKLIN E. ZIMRING & GORDON J. HAWKINS, DETERRENCE: THE LEGAL THREAT IN CRIME CONTROL (1973).
21. See supra note 4 for a definition of the provision of protection.
22. See supra note 3.
of protection among different classes of potential victims. The justifications for using nonreciprocal norms in criminal law should take into account these concerns. In the next Part, I shall explore the efficiency considerations and, in Part III, the just manner of distributing protection.

II
THE EFFICIENCY OF THE CRIMINAL LAW PRINCIPLE OF COMPARATIVE FAULT

Can we justify the nonreciprocal nature of criminal law norms on efficiency grounds? Can we demonstrate that neglecting victims' conduct in determining criminal sanctions is efficient? This Part explores these questions. It begins by examining the important role of victims in tort law as reflected in the comparative fault doctrine. Then, on the basis of analogy to tort law, it defines a criminal law principle of comparative fault and points out efficiency-based reasons for favoring the adoption of such a principle. Lastly, it examines and rejects various efficiency-based objections to the criminal law principle of comparative fault.

A. The Role of Victims in Tort Law

In examining the nonreciprocal nature of criminal law norms, it is useful to draw an analogy to tort law. The analogy is justified on two grounds. First, economists have explored extensively the efficiency considerations underlying tort law. Second, economists have pointed out important similarities between criminal law and tort law.

A tort is a harm generated by human action or inaction which can be prevented by taking (often costly) precautions. It is often the case that the tortfeasor or the victim, or both, will have to take costly precautions in order to minimize the total costs of human activity. Total costs are defined

23. I am grateful to Dick Craswell for detailed and insightful comments on this Part.

24. See, e.g., sources cited supra note 29.

as the sum of the costs of precautions and of expected accident losses.\textsuperscript{26} From an economic perspective, the aim of tort law is to provide optimal incentives such that both the injurer and the victim will take optimal precautions, that is, precautions which minimize the total cost of accidents.\textsuperscript{27}

In a world free of transaction costs, both parties would take optimal precautions irrespective of the initial allocation of legal responsibility since, if the legal system allocated responsibility in a nonoptimal manner, the parties would voluntarily shift the responsibility in a socially optimal manner.\textsuperscript{28} Very often, though, a voluntary shift in the allocation of responsibility is costly. Because of transaction costs, the allocation of legal responsibility may provide nonoptimal incentives and consequently may lead to suboptimal behavior.\textsuperscript{29}

The most important aspect of this framework for our purposes is the reciprocal nature of the norms of tort law, according to which both the alleged tortfeasor and the victim are required to act to minimize the total costs of accidents.\textsuperscript{30} Under certain circumstances, tort law shifts responsibility from the injurer to the victim. The principal doctrines that shift responsibility to the victim are contributory negligence and comparative fault.

Under a system of contributory negligence, the victim’s negligence bars his recovery completely. Dissatisfaction with the harshness of this

---

\textsuperscript{26} The expected accident losses are defined as the potential magnitude of the loss multiplied by the probability of suffering the loss. \textit{See Steven Shavell, Economic Analysis of Accident Law} 6 (1987).

Precautions should be interpreted here as broadly as possible. Thus, for my purposes precautions include any change in the activity of the injurer or the victim which is designed to reduce the expected costs of accident losses. These changes are divided into two sub-categories: changes in the level of activity and changes in the level of care. An example can illustrate the difference between the two sub-categories. Driving fewer miles can be interpreted as a change in the level of activity, while slowing for curves or paying attention to the presence of bicyclists can be interpreted as a change in the level of care. \textit{Id. at} 5.

\textsuperscript{27} \textit{See id. at} 26-32.

\textsuperscript{28} This is because if the law does not allocate responsibility in an optimal manner, both parties stand to gain from the surplus created by a reallocation of responsibility in an optimal manner. \textit{See generally} Coase, \textit{supra} note 15.

\textsuperscript{29} \textit{See Shavell, supra} note 26, at 5-32.

This framework may sound simple and straightforward, but there exists a serious controversy as to its proper application. The questions which arise in this debate include: What is a cost? How should the optimal behavior of the injurer and the victim be determined? Who should make determinations concerning the optimal behavior of the injurer and the victim? The debate among scholars who accept this basic framework began with the classical debate between Calabresi and Posner over the economic desirability of strict liability versus negligence systems. For Calabresi’s position, see Guido Calabresi, The Costs of Accidents: A Legal and Economic Analysis (1970). For Posner’s position, see William M. Landes & Richard A. Posner, The Economic Structure of Tort Law (1987); Richard A. Posner, A Theory of Negligence, 1 J. Legal Stud. 29 (1972); Richard A. Posner, Strict Liability: A Comment, 2 J. Legal Stud. 205 (1973). For a recent analysis of the issues, see Shavell, \textit{supra} note 26, passim.

\textsuperscript{30} Indeed, the very identification of an agent as a tortfeasor or as a victim depends upon the magnitude of the harm she causes. Coase argues that the legal system should classify as victim the person whose activity generates the larger net benefit. \textit{See Coase, supra} note 15.
result led to the development of the comparative fault system, under which the fault of the victim reduces his compensation proportionately, but does not bar recovery altogether. Both doctrines shift responsibility in order to provide incentives to potential tort victims to take reasonable precautions against the tort.\textsuperscript{31}

The similarity drawn by economists between criminal law and tort law may suggest that efficiency requires adoption of an analogous responsibility-shifting principle in criminal law.\textsuperscript{32} How would such a principle operate in the criminal law context? The criminal justice system could be redesigned so that it would acquit the alleged criminal,\textsuperscript{33} or at least mitigate his punishment, when the victim of the alleged crime failed to take precautions reducing the likelihood of the crime. I shall term such a scheme a criminal law principle of comparative fault and explore in the next Section efficiency-based reasons favoring the adoption of such a scheme.

Before exploring the rationale for adopting a principle of comparative fault in criminal law, it is important to qualify the analogy drawn between criminal law and tort law. The fundamental principles guiding policy in tort law can be useful in the context of criminal law. Thus, for instance, the

\textsuperscript{31} For discussions of the advantages and disadvantages of comparative and contributory fault systems, see Robert D. Cooter & Thomas S. Ulen, An Economic Case for Comparative Negligence, 61 N.Y.U. L. Rev. 1067 (1986); David Haddock & Christopher Curran, An Economic Theory of Comparative Negligence, 14 J. Legal Stud. 49 (1985); Daniel Orr, The Superiority of Comparative Negligence: Another Vote, 20 J. Legal Stud. 119 (1991); Samuel A. Rea, Jr., The Economics of Comparative Negligence, 7 Int'l Rev. L. & Econ. 149 (1987); Daniel L. Rubinfeld, The Efficiency of Comparative Negligence, 16 J. Legal Stud. 375 (1987).

\textsuperscript{32} One powerful objection to this suggestion would be to dispute the aptness of the analogy between ordinary torts and crimes. While tort law recognizes a principle of comparative fault, it excludes from its scope an important category of torts, namely, intentional torts. Arguably, the relevant analogy should be drawn between crimes and intentional torts rather than between crimes and ordinary torts because, unlike in the case of ordinary torts, the perpetrator of both crimes and intentional torts intends to cause harm.

The economic reasons for the exclusion of intentional torts from the scope of the principle of contributory fault have been explored in detail by economists. See, e.g., Landes & Posner, supra note 29, at 153-56. Under their analysis, the optimal level of activity of tortfeasors who commit intentional torts is zero. Given this optimal level, the optimal level of precautions by victims is zero. Therefore, the introduction of a contributory fault principle would induce victims of intentional torts to invest in precautions which are inefficient. \textit{Id}.

But this analysis, even if it is valid in the context of intentional torts, is based on certain premises which are inapplicable to criminal law. The criminal may be the "cheapest cost avoider" and hence it may be that the optimal level of crime is zero. However, it does not follow that it is optimal for the victim not to invest in any precautions against crime. Unlike tort law, enforcement of criminal law imposes costs upon the state. It is possible that it is better to deter the criminal from committing the crime through precautions taken by the victim rather than through expending resources of the law enforcement and criminal justice systems. Thus, precautionary measures taken by the victim should be regarded as an alternative mechanism for reducing criminal activity which, under certain circumstances, may be more efficient than the traditional mechanisms of the criminal justice system. For a more detailed elaboration of this argument, see \textit{infra} text accompanying notes 44-47.

\textsuperscript{33} I use the terms "alleged criminal" and "alleged crime" because a legal system which adopts a criminal law doctrine of comparative fault may decline to define as crimes activities directed against careless victims when those very activities would be defined as crimes when directed against cautious victims.
claim that one primary aim of tort law is to guide not merely the tortfeasor’s behavior, but also that of the victim, could spark an inquiry into whether criminal law does or should shape the behavior of both the criminal and the victim. Yet, use of the term “criminal law principle of comparative fault” may be misleading in three respects. First, it is important to bear in mind that despite the apparent implication of the term, I do not assume moral fault on the part of the victim. Second, the term comparative fault is used here in a much broader manner than it is used in tort law. Under my definition, it denotes a rule which may annul criminal responsibility altogether as well as a rule which mitigates a criminal’s punishment. The definition includes, therefore, what would be called contributory negligence in tort law as well as tort law’s doctrine of comparative fault. Lastly, the term “criminal law principle of comparative fault” may be misleading because of its affiliation with tort law. The determination of fault on the part of the victim, as well as the legal implications of fault within criminal law, differ from those in the tort law context. In particular, the proper scope of the principle, the conditions of its applicability, and the moral constraints upon its applicability are not the same in both legal contexts. The use of the term “comparative fault” in criminal law may be useful only if we bear in mind the important distinctions that its use in this new context implies.

B. Efficiency-Based Reasons Favoring the Adoption of a Criminal Law Principle of Comparative Fault

Efficiency considerations dictate that criminal law should strive to minimize the total costs of crime, defined as the sum of the expected costs of crime and the costs of precautionary measures against crime. The costs of precautionary measures are borne by the state as well as by potential victims. Efficiency requires the distribution of the costs of precautions between the state and potential victims in a way that minimizes the total costs of crime. Without appropriate incentives, potential victims will invest less in precautionary measures than is socially optimal. That is, potential victims will expose themselves to risks which are unjustified from the perspective of efficiency. I shall attempt to demonstrate that the best way to

34. See supra text accompanying note 15.
35. For a discussion of those matters, see infra Part V.
36. See ROBERT O. COOTER AND THOMAS S. ULEN, LAW AND ECONOMICS 536-39 (1988). This definition is quite similar to the one used in the context of tort law. See supra text accompanying note 26.
37. One of the most sophisticated foes of the economic analysis of criminal law has argued that the economic perspective sometimes requires criminalization of negligent behavior of potential victims of crime. The following paragraph illustrates the supposedly inevitable conclusion of economic analysis:

[When confronted with two inconsistent activities, we must decide which to discourage in order to avoid the cost of conflict. We must, in other words, decide who will be the criminal and who will be the victim. It may be, for example, that either raising the penalty for unauthorized use of a motor vehicle or imposing criminal penalties on owners who leave their
induce potential victims to take optimal precautions is to adopt a criminal law principle of comparative fault. This argument requires establishment of the following two propositions:

1. In the absence of a criminal law principle of comparative fault, potential victims’ incentives to take precautions will be suboptimal.
2. A criminal law principle of comparative fault can guarantee that the incentives for potential victims to take precautions will operate efficiently, or at least more efficiently than in the absence of such a principle.

I will establish each of these claims through a consideration of counterarguments. In contrast to the first claim, it could be argued that even in the absence of a criminal law principle of comparative fault, potential victims’ investment in precautions is optimal given that potential victims always bear the full expected costs of crimes directed against them. This argument suggests that there is a major difference between criminal law and tort law which undermines the attempt to import the principle of comparative fault from tort law into criminal law. In tort law, a victim who is likely to be compensated irrespective of her behavior will underinvest in precautions because she does not bear the full costs of the accident. Hence, in order to provide optimal incentives for potential victims of torts to take precautions, we deny compensation if the victim fails to take adequate precautions. This is not, so the argument goes, the case in criminal law. Irrespective of the enforcement policy, the victim of crime bears the full expected costs of the crime. Consequently, we do not need a criminal law principle of comparative fault in order to induce potential victims of crime to take optimal precautions. Moreover, adopting such a principle may provide stronger incentives than necessary and hence bring about a larger investment in precautions than is socially optimal.

cars unlocked will be equally effective in reducing the number of car thefts. Which course—or combination of courses—we choose reflects no more than how we prefer to distribute the cost of control.


The symmetry between the criminal and the victim according to the economic perspective, Seidman claims, is too counterintuitive to be true and therefore should lead us to reject the economic perspective on criminal law. Id. at 331-34.

This reasoning is misleading. Even tort law does not impose liability on a victim who fails to take reasonable precautions. Instead, it merely denies compensation, in whole or in part. The equivalent measure in criminal law would not be the attribution of criminal liability to the victim, but the exculpation of the alleged criminal or a reduction of his punishment. Such a measure is less counterintuitive than Seidman suggests and hence cannot provide a solid basis for rejecting the economic approach to criminal law.

Moreover, even if Seidman’s reasoning were correct, the normative conclusions he draws are too hasty. It was difficult for pre-Coase scholars to question the alleged asymmetry between tortfeasor and victim. It is even more difficult to digest the more radical suggested symmetry between criminal and victim. But the deconstruction of the concept of victim and criminal through Coasian insights should not be dismissed, even if “counterintuitive,” before it is granted a fair hearing.
This is a powerful objection to the first proposition, but its validity depends upon the "no externalities" premise, namely that precautions taken by potential victims have no impact on others. The "no externalities" premise, however, is false. Precautions taken by potential victims may provide benefits to others or impose costs on others. The potential victim is bound to ignore these costs or benefits in deciding how much to invest in precautions. Thus, the existence of externalities will lead potential victims to invest too much or too little in precautions against crime.

The precautions taken by potential victims have two types of positive externalities. First, increased precautions taken by some potential victims benefit other potential victims because those precautions increase the costs of crime for the criminal, thereby reducing the number and the severity of his crimes. While a successful crime directed against an individual raises terror and fear among other potential victims, the relative immunity of cautious victims contributes to a general societal feeling of security and stability. Hence, precautions taken by potential victims contribute to the well-being of other potential victims.

Second, the state itself benefits from victims' precautionary measures because such measures reduce the costs of the enforcement system. The reduction in the number and the severity of crimes as a result of better precautions taken by potential victims reduces the costs of identifying, prosecuting, and punishing criminals. The optimal level of victims' precautions should be defined by reference to all of the relevant costs. But potential victims will ignore the costs of the enforcement system in determining their investment in precautions. Once those externalities are identified, it is clear that victims' private incentives are suboptimal.

The "no externalities" premise, therefore, is subject to criticism. Precautions taken by victims provide two types of positive externalities. The first directly benefits other potential victims by making crime less prof-

---

38. This claim may raise the objection that precautions taken by one victim will simply lead to a crime directed against another victim. This phenomenon, known as displacement of crime, has been discussed extensively in the criminology literature. See infra text accompanying notes 48-52.

39. A successful crime directed against a careless victim will raise fear even among cautious victims who do not know that the victim was careless.

It is true that a crime does not have to be successful in order to sow fear in the community. Yet cautious people should not only be more immune to successful crimes, but also to attempted crimes, since criminals are less likely to target cautious individuals.

40. An analogous point has been made in tort law. Several writers have explored the impact of litigation costs (which are analogous to the costs of criminal law enforcement) on deterrence. See, e.g., Keith N. Hylton, The Influence of Litigation Costs on Deterrence Under Strict Liability and Under Negligence, 10 INT'L REV. L. & ECON. 161 (1990); A. Mitchell Polinsky & Daniel L. Rubinfeld, The Welfare Implications of Costly Litigation for the Level of Liability, 17 J. LEGAL STUD. 151 (1988).

41. I have discussed only positive externalities. Taking precautions may also have negative externalities. An excess of precautions (e.g., gated driveways, alarms, guard dogs) may, for example, lead to a feeling of insecurity and of being under siege. More importantly, precautions taken by one individual may shift crimes to other potential victims. See infra text accompanying notes 48-52.
itable; the second benefits the population at large by reducing the indirect costs of crime, in particular, the costs of the enforcement system.\(^42\)

Establishing the first proposition—that victims’ incentives to take precautions are suboptimal—is not sufficient for our purposes. In addition, we need to establish the second proposition—that a comparative fault principle would indeed increase victims’ incentives to invest in precautionary measures. In a legal system that adopts a criminal law principle of comparative fault, the sanctions imposed upon a criminal will depend upon the precautions taken by his or her victim. If the victim of a crime desires that the perpetrator receive a severe punishment, then the use of a principle of comparative fault would motivate the victim to take better precautions. Hence, in order to establish the second proposition, we need to show that victims of crime are in fact interested in the degree of punishment of the criminals who attacked them or their property.

Harsher sanctions for criminals might make their victims better off for two reasons. First, some actual victims no doubt harbor retaliatory sentiments and get a high degree of satisfaction from punishment of the guilty parties. Arguably, in a legal system that adopted a criminal law principle of comparative fault, potential victims would take better precautions in order to guarantee that criminals who acted against them would indeed be punished harshly. I suspect, however, that while a desire for retribution is common among actual victims, it hardly influences the behavior of potential victims. Retaliatory sentiments may be intense once a person suffers an actual loss. However, the prospect of suffering from a potential future crime and the wish to see the future criminal punished would rarely influence the behavior of potential victims.

A second, more satisfactory explanation for victims’ interest in punishing criminals relies on the relationship between the expected costs of crime to potential victims and criminals’ expected costs of committing crimes. Consistent with this explanation, a principle of comparative fault changes the incentives operating on criminals. This change, in turn, affects the incentives operating on victims, inducing them to act in a more efficient manner.

Under a scheme of comparative fault, criminals will be more reluctant to commit crimes against cautious victims than in a regime which neglects to treat victims as actors. If a criminal who is aware that his jurisdiction has adopted a comparative fault principle knows that a potential victim has taken a suboptimal level of precautions, that criminal will be more inclined to commit a crime against that careless victim because he knows that he will be subjected to less severe penalties. Similarly, the informed criminal

---

42. Why should the state not provide victims the necessary incentives to take optimal precautions by criminalizing their negligent behavior? In addition to the immoral nature of such a policy, the costs of criminalizing victims’ negligence may be too high. Identifying, prosecuting, and punishing negligent victims is costly.
will be less inclined to commit a crime against the cautious victim because the criminal knows that he will be subjected to higher penalties.

Criminals' "preference" for careless victims under a scheme of comparative fault will, in turn, influence the behavior of potential victims. Potential victims will be disposed to take better precautions given that criminals will be less likely to commit crimes directed at cautious victims. Thus, adopting a principle of comparative fault provides additional protection in the form of deterrent benefits to potential victims who have taken optimal care.

One qualification should be pointed out. This analysis is valid only if the precautions taken by victims can be detected by criminals. Non-observable precautions, those that cannot be detected by criminals in advance, cannot influence criminals' behavior and consequently cannot affect the degree to which cautious victims are exposed to crime.\(^4^3\)

Adopting a criminal law principle of comparative fault is only one means to induce victims to take better precautions. Adjustments to other components of the enforcement system can bring about similar effects. For example, a more lenient exercise of judicial discretion in punishing criminals who have acted against careless victims, or a lesser investment in detecting and prosecuting those criminals will bring about similar results. In each case, society invests more resources in protecting cautious victims and hence provides potential victims with stronger incentives to take precautions.

So far, two claims have been established. First, in the absence of a criminal law principle of comparative fault, the investment of potential victims in precautions is suboptimal. Second, adopting a criminal law principle of comparative fault, will increase the incentives of victims to invest in precautions. Both claims together imply that adopting a principle of comparative fault may guarantee more efficient investment in precautions by potential victims.

C. Efficiency-Based Objections to the Criminal Law Principle of Comparative Fault

In this Section, I shall examine three attempts to provide an efficiency-based justification for the absence of a criminal law principle of comparative fault.\(^4^4\) Posner provides two arguments which suggest that criminals are consistently the cheapest cost avoiders. Both arguments miss the mark.

\(^4^3\) This may seem like a very serious qualification. However, one should recall that if a criminal law principle of comparative fault is adopted, potential crime victims' incentives to notify criminals that they have taken precautions will increase. Under the current system, victims are already motivated to inform criminals of precautions they take, but adoption of a principle of comparative fault will increase such motivation. Hence, it is more likely that precautions which are currently invisible will become visible due to the stronger incentives of victims to inform criminals of the existence of precautions.

\(^4^4\) Obviously there are non-efficiency considerations which may explain the absence of a criminal law principle of comparative fault. See infra Part V (considering argument that criminal law
because the appropriate comparison is not between criminal and victim but rather between victim and state. A third possible justification for the absence of a criminal law principle of comparative fault, the theory of crime displacement, could be used to refute any deterrence-based argument, not just my own; moreover, it suffers from several weaknesses.

One argument has been (too) briefly raised by Posner:

Though in principle, . . . the victim rather than the aggressor might be the “cheaper cost avoider,” . . . it appears that the market-bypassing acts that have been made criminal are primarily those where the victim is never (or very rarely) the cheapest cost avoider. Other harmful acts are more likely to be governed by tort law, with its concepts of assumption of risk and contributory negligence that facilitate comparing the costs to potential injurer and to potential victim of avoiding injury.45

Posner believes that current law makes conduct criminal only after a cost-benefit calculation has been made: the person who will bear criminal liability has been identified by the legal system as the person who is in the best position to prevent the harm, while it is rarely socially optimal for the person who is labeled a victim to take precautions against crime. Therefore, the economic principles requiring us to allocate responsibility for an activity in an optimal manner are built into the very definition of criminal activity.

In his more recent writings, Posner develops a different efficiency argument.46 According to this argument, first developed in the context of intentional torts, the criminal is necessarily the cheapest cost avoider because his cost of avoiding the crime is negative. The criminal’s cost of avoidance is negative because commission of the crime requires him to expend resources; or, not committing the crime saves him resources. Therefore, the criminal’s cost of avoidance must be smaller than the costs to the victim of avoiding the crime, which are positive.

Posner’s second argument shares the same flaw that besets his earlier argument. Both arguments are based on an assumption, rooted in tort law, that the only relevant comparison is that of the alleged criminal and the victim. Comparing the costs of prevention by the criminal and by the victim will lead inevitably to the conclusion that the criminal is often, if not always, the person who is better suited to “avoid the risk.” The identification of the criminal as the “cheaper cost avoider” does indeed mean that it is socially desirable that criminals avoid carrying out crimes. But given the persistence of criminal activity, the salient question is who should bear the

principle of comparative fault is morally repugnant); but see infra Part IV (the premise that the principle is not present in criminal law is false).


46. See LANDES & POSNER, supra note 29, at 153-58.
costs of preventing such activity. It is not sufficient merely to point out the inefficiency of criminal behavior and the importance of preventing it. In addition, it is necessary to identify the person or the institution that is in the best position to take precautions. Comparing the victim with the criminal only confuses this process of identification, which requires us instead to compare the costs of precautionary measures to the victim and the costs of precautionary measures to the state.

The flaw in Posner's proposal is its failure to recognize that the imposition of costs on the criminal cannot be accomplished apart from the imposition of costs on the state. By focusing on the criminal, Posner's approach obscures the crucial issue: on which party (the potential victim or the state) the costs should fall.

The important question, from the perspective of efficiency, is not whether the criminal is in the best position to prevent the crime, but whether the costs of precautionary measures are better borne by the state or by the potential victim. The state can deter future crimes by imposing harsher sanctions or by increasing the probability of detection and conviction. Potential victims can prevent the crime by taking precautions that will make it difficult and more costly for the criminal to carry out the crime. A decision to punish the criminal, rather than being a decision to impose the costs of precautions on the criminal, is instead a decision by the state to relieve the victim of responsibility for those costs.47

An alternative efficiency-based explanation for the absence of a criminal law principle of comparative fault is rooted in the notion of crime displacement. Many scholars, in particular criminologists, have argued that when potential victims use better precautionary measures, the risks of crime are simply shifted to other potential victims. This phenomenon is known in the criminology literature as displacement of crime. Pessimistic criminologists believe that crime control activities simply redistribute crimes to more vulnerable locales and persons rather than prevent them.48 Under this view,

47. Investment in precautions by the state can take two forms. The state can either raise the probability of conviction or increase the severity of the sanctions imposed upon the criminal. Economists have discussed the impact of each one of those factors. See Gary S. Becker, Crime and Punishment: An Economic Approach, 76 J. Pol. Econ. 169 (1968), reprinted in Essays in the Economics of Crime and Punishment 1 (Gary S. Becker & William M. Landes eds., 1974); A. Mitchell Polinsky & Steven Shavell, The Optimal Tradeoff Between the Probability and Magnitude of Fines, 69 Am. Econ. Rev. 880 (1979).


Criminologists distinguish among (at least) five forms of displacement of crime: temporal, tactical, target, territorial, and functional. See Reppetto, supra, at 168-69. The criminal may change the time in which he commits the crime, the ways in which he commits the crime, the target of the crime, the place
an attempt by the state to induce victims to take precautions by adopting a
criminal law principle of comparative fault is inefficient given that victims' 
precautions merely redistribute costs of crime rather than reduce them.

Displacement of crime, however, can hardly justify the rejection of a
criminal law principle of comparative fault. First, displacement is limited 
to a subset of crimes. It is inapplicable, for example, to passion crimes. If a 
potential criminal's lover has left him, precautions taken by the lover are 
unlikely to lead the desperate person to look for an alternative victim. 
Second, as some studies show, displacement of crime often forces criminals 
to engage in less attractive and more risky activities, or to abandon criminal 
activity altogether. For example, better precautions taken by potential 
victims in an affluent neighborhood may force criminals to turn their 
attention to less affluent neighborhoods and subsequently may persuade them to 
leave criminal activity altogether and choose instead legal substitutes, or 
at least to choose to commit a less serious crime. Lastly, criminologists 
have observed that crime control measures lead to reduction of crimes 
beyond those they were meant to thwart. This "free rider" effect means 
that anti-crime steps taken by one victim may prevent crimes directed at 
other victims. The force of these counterarguments is perhaps limited, but 
they should at least raise serious doubts as to the validity of the displace-
ment hypothesis.

Efficiency seems to support the adoption of a criminal law principle of 
comparative fault. Attempts to explain the absence of such a principle on 
the basis of the claim that the very definition of crimes incorporates a cost-
benefit analysis, or on the basis of the claim that the costs of a criminal 
avoiding crime are negative, or lastly, on the basis of the crime displace-
ment hypothesis, fail. However, criminal law should not only aim at pro-
viding protection in an efficient manner. As the next Part points out, the 
law should also distribute protection fairly among different potential 
victims.

---

of the crime or the type of crime. Recently, criminologists have added a sixth form of displacement—
perpetrator displacement—where an attempt to restrict a perpetrator of a crime will lead another 
perpetrator to take his place. See Barr & Pease, supra, at 279.

49. See, e.g., Cornish & Clarke, supra note 48, at 934 (citing research that finds, for example, that 
 crackdowns on subway robbers force them to turn to street robbery).

50. Cornish and Clarke offer another example, where displacement of drunk driving will likely 
compel potential drunk drivers to seek legal means to transport themselves, rather than to turn to other 
crimes. See id. at 938.

51. The criminal’s choice of lesser crimes is known among criminologists as benign displacement. 
See Barr & Pease, supra note 48, at 284-89. Of course, some criminals will choose to commit more 
severe crimes ("malign displacement"). See id. at 289-91.

52. Under different names, this phenomenon has been noted by both criminologists, see Clark,
Introduction, in Situational Crime Prevention, supra note 48, at 3, 25; Terance D. Miethe, Citizen-
Based Crime Control Activity and Victimization Risks: An Examination of Displacement and Free-Rider 
Effects, 29 CRIMINOLOGY 419, 422 (1991), and economists, see Steven Shavell, Individual Precautions 
(finding a general deterrent effect when individuals jointly raise precautions).
III
THE FAIRNESS OF THE CRIMINAL LAW PRINCIPLE OF COMPARATIVE FAULT

Traditionally, criminal sanctions have been regarded either through the prism of utilitarianism (or the wealth-maximizing prism) or through the prism of retributivism. The literature rarely rationalizes punishment in terms of distributive justice.\(^{53}\) This Part will illustrate how notions of distributive justice can enrich our understanding of criminal law. According to my account, criminal law should be described as a mechanism for distributing protection in a fair manner to potential victims. The fair distribution of protection requires the selective introduction of a criminal law principle of comparative fault.

A. Protection as a Non-Public Good

It is hardly new to regard criminal law as a means of providing stability and security for society. Provision of protection, like the provision of any other good, should be subject to distributive justice considerations. Yet criminal law scholars have been reluctant to analyze the provision of protection through a distributive prism.

The reason for this reluctance seems to be based on the widely held belief that protection is a pure public good.\(^{54}\) A pure public good has two characteristics: jointness of supply and the impossibility or inefficiency of


Punishment may restore a fair distribution of resources, either by making the victim better off or by making the criminal worse off, or both. If punishment of criminals restores equality by making victims better off, in what way exactly are they better off? If victims are made better off simply because they are being vindicated, should the legal system reinforce the victims’ vindictive passions? For an examination of the legitimacy of vindictive passion, see Jeffrie G. Murphy, Getting Even: The Role of the Victim, in CRIME, CulpABILITY, AND Remedy 209 (Ellen F. Paul et al. eds., 1990).

If, on the other hand, punishment restores equality by making criminals worse off, should we restore equality by worsening the position of those who are better off, i.e., criminals, without thereby making those who are worse off, i.e., victims, better? There are moral and political grounds to argue that restoring equality should always be done by improving the position of those who are worse off rather than by worsening the position of those who are better off.

My suggestion, though it relies on principles of distributive justice, is different. Instead of interpreting punishment as a practice aimed at restoring a fair balance disturbed by criminal activity, I interpret punishment of criminals as the provision of protection to potential victims. The distribution of protection, like the distribution of any other good by the state, should conform with principles of distributive justice.

\(^{54}\) This belief poses a well known problem for the Hobbesian tradition. See Jean Hampton, Hobbes and the Social Contract Tradition 175 (1986).
excluding others from its consumption. The distribution of pure public goods—unlike the extent of their provision—by definition cannot be determined by the state, because individuals cannot be excluded from enjoyment of such goods. Consequently, one is tempted to conclude that distributive justice considerations are irrelevant to public goods analysis.

This conclusion certainly does not apply to protection, for the belief that protection is a pure, or almost pure, public good is mistaken. The state can control the distribution of protection. For example, the state can provide better police protection to certain neighborhoods; it can provide more resources for the identification and prosecution of criminals who have committed crimes against certain victims; or it can, over time, provide better or worse protection to some victims by imposing harsher or lighter sanctions upon criminals who commit crimes directed against those victims.

Once the misconception that protection is a pure public good is exposed, it is clear that the state has significant power and therefore great responsibility concerning the manner in which protection is distributed. Protection is a uniquely important good provided by the state to its citizens. Its distribution, like that of any non-public good supplied by the state, must be subject to distributive justice concerns.

The precise nature of a fair distribution of goods is a highly controversial subject. Once one admits that the provision of protection should be

55. A good definition of these terms can be found in Michael Taylor, The Possibility of Cooperation 5-6 (1987):

A good is said to exhibit perfect indivisibility or jointness of supply (with respect to a given set of individuals, or public) if, once produced, any given unit can be made available to every member of the public, or equivalently if any individual's consumption or use of the good does not reduce the amount available to others. A good is said to exhibit nonexcludability (with respect to some group) if it is impossible to prevent individual members of the group from consuming it or if such exclusion is "prohibitively costly"... Id. at 5-6 (footnote omitted). For another lucid characterization of public goods, see Dennis C. Mueller, Public Choice II, 11 (1989).

The feature of public goods which has attracted the most attention from economists and legal scholars is the need for centralized, nonmarket production of public goods in order to guarantee optimal provision. If there are no mechanisms to enforce individuals' participation in the production of public goods, these goods will not be provided because of the free rider phenomenon. Given the inability to exclude individuals from the benefits, free riders will enjoy public goods, but will refuse to participate in their provision. See Mueller, supra, at 11-13.

56. The ability of the state to provide protection is often regarded as the main justification for the existence of the state. The most forceful defense of this position can be found in Thomas Hobbes, Leviathan 153, 230 (Richard Tuck ed., 1991).

57. The phrasing of the issue in terms of distributing protection to victims, rather than sanctions to criminals, may seem biased. After all, the distribution of protection affects the interests of two groups: potential victims and potential criminals. Fairness in the distribution of protection to victims may lead to unfairness in the distribution of the costs of protection to the criminals, that is, unfairness in the distribution of sanctions. Nonetheless, the focus on distributing protection to the victims is justified because criminals who commit similar crimes do not have a right to equal sanctions; they merely have a right to sanctions which are not cruel or unreasonable.

subject to distributive justice concerns, the distribution of protection should hardly remain immune to the controversies surrounding the distribution of other goods. A full exploration of the meaning of a fair distribution of protection must therefore consider the conflicting principles governing the distribution of goods generally.

Before embarking upon this enterprise, the scope of the fairness requirement should be qualified. The resources of the state must be distributed in a fair manner, that is, in a manner which assures every individual an equal share of publicly financed protection. Some individuals, seeking greater protection, may invest privately in precautions. Such individuals should not be denied the right to purchase a higher degree of protection than is granted to them by a fair public distribution of protection by the state.

B. Two Models of Distributive Criminal Justice

The public distribution of protection can be governed by one of two models: the equal protection model or the equal costs model. Deciding which model to prefer has important ramifications for the applicability of the criminal law principle of comparative fault. As will be shown, when the equal protection model should govern, use of the criminal law principle of comparative fault is inappropriate, whereas when the equal costs model should govern, the use of the comparative fault principle can bring about fairness in the distribution of protection.

The equal protection model requires the enforcement system to furnish every individual the same degree of protection. The degree of protection of an individual can be defined as the expected harm from crime for that individual, that is, the probability of harm multiplied by its size. Under the equal protection principle, an enforcement system is fair if the expected harm from crime for each individual equals the expected harm from crime of every other individual in the society.59

The equal costs model, on the other hand, requires the state to invest equal resources in the protection of every individual. This may lead to a disparate degree of protection given the inevitable variance in individuals' vulnerability to crime.60 Under the equal costs model, an enforcement system is fair if the costs of protecting each potential victim are equal across all potential victims. The contrast between these two models can be articulated as follows: the equal protection model focuses upon the product or the output of the enforcement system, namely protection, while the equal

---

59. It may seem that the equal protection principle gives priority to the protection of the rich at the expense of the poor. Given that the poor are less wealthy, their expected cost of crime is bound to be lower. This assumes that we measure the expected costs of crime in terms of wealth rather than utility. The proper measure, however, is utility rather than wealth.

60. Vulnerability to crime encompasses two factors: likelihood of being the victim of a crime and the degree of harm in the event of victimization.
costs model focuses upon the inputs of the enforcement system, namely its costs.61

The remainder of this Part will develop criteria to determine the circumstances under which each one of these models should be used as well as explore the implications of these models for the criminal law principle of comparative fault.

1. The Equal Protection Model

According to the equal protection model, the expected harm of crime incurred by any individual should be equal to the expected harm of crime incurred by any other individual. The model requires special protection for individuals who are more vulnerable to crime in order to equate their expected harm to that of less vulnerable individuals.62 Protection can be supplied by imposing harsher sanctions on those who commit crimes directed against vulnerable victims or by devoting more resources to the identification and prosecution of criminals who commit those crimes.63

It is reasonable to apply the equal protection model when the special vulnerability of the potential crime victim is not voluntary.64 Society should protect victims who are particularly vulnerable to crime due to factors that are beyond their control.

This claim can be explored and refined with several illustrations. Women are more vulnerable to some crimes than men, since they are more likely to become victims of sexual offenses. Minorities are more vulnerable to racially based crimes than whites.65 The equal protection model should govern the distribution of protection to women and minorities because their vulnerability is involuntary. The model requires special protection for these groups in order to assure them, in the end, the same degree of protection enjoyed by less vulnerable groups. One way to grant vulnerable individuals special protection is to impose harsher sanctions on criminals who commit crimes directed against vulnerable individuals.66 Thus, under the equal protection model, the legal system should intervene in order to equalize the

61. In a world in which everyone is equally vulnerable to crime, the two models will coincide; but in a world like ours in which some are more vulnerable than others, the two models necessarily diverge.

62. Closer examination of the equal protection principle reveals ambiguities in the concept of equal protection. For instance, does equal protection mean that every member’s probability of falling prey to any crime is equalized? Or can a greater probability that victim A will suffer from one sort of crime be balanced by the greater probability that victim B will suffer from another sort of crime? Despite the importance of these ambiguities, resolving them is not crucial for my purposes.

63. See supra note 4 for an explanation of my use of the term “protection.”

64. It is beyond the scope of this Article to provide a comprehensive definition of voluntariness. The problem with the use of the term “voluntary” has been articulated as follows: “Whether we call a risk voluntary or involuntary depends mostly, it seems, on whether we like to call the cup half empty or half full.” Leo Katz, Bad Acts and Guilty Minds: Conundrums of the Criminal Law 43 (1987).

65. For an explanation of my use of the term vulnerable, see supra note 60.

66. This may justify enhancement of sentences for criminals motivated by racial animosity. There are seemingly more obvious forms of protection for vulnerable victims, such as increased police patrols. However, such methods of protection cannot adequately be targeted to the vulnerable group.
expected costs of crime to vulnerable groups and the costs incurred by less vulnerable groups.\textsuperscript{67}

One important qualification to the notion of voluntariness should be emphasized. The equal protection model is appropriate when the harm is \textit{predominantly} the byproduct of an involuntary vulnerability, although the involuntary vulnerability may be coupled with a voluntary act. The harm in rape or in racially-based attacks could be understood as the product of an involuntary state of affairs—the special vulnerability of women or minorities—and a voluntary behavior—the so-called "provocative behavior" of women, or the decision of a minority to stroll through a white neighborhood. I emphasize that I do not claim that the "provocative behavior" of women actually increases their chances of becoming victims of sex crimes. Moreover, a minority may suffer a racially motivated attack anywhere. My argument is simply that, even assuming a correlation between victim behavior and a sex crime or a racial crime, the use of the doctrine of comparative fault cannot be justified in these contexts. As the above examples illustrate, almost any crime can be conceived of as the conjunction of a voluntary and an involuntary state of affairs. Whether the voluntary or involuntary component should determine the principle for distributing protection is essentially a normative decision that is often controversial.

Sometimes the legal system grants special protection to vulnerable victims despite the voluntary nature of the victim’s vulnerability. Victims are protected if their vulnerability is attributed to a voluntary act which possesses an important social value. We insist on granting full protection to individuals who use their First Amendment rights in a way which provokes a hostile reaction, even if by doing so they voluntarily expose themselves to severe risks.\textsuperscript{68} The so-called "Heckler’s Veto" doctrine requires that police use available resources to protect a speaker who inflames a hostile audience because of the special importance of freedom of expression, despite the voluntary nature of the speech which generates the risks. Another example can be used to illustrate this point. Some owners of factories or stores decide, for ideological or commercial reasons, to place their businesses in a depressed, high crime area. Such a location is socially valuable because it

\textsuperscript{67} The equal protection principle can explain the moral indignation directed towards judges who take into account the "indecent" behavior of women in rape cases. The equal protection principle requires a special effort to guarantee that the ex ante costs of crime be equal between men and women. Why should the special vulnerability of women require them to avoid "provocative behavior" when men can behave "provocatively" without exposing themselves to grave risks? It is not only that men’s provocative behavior does not expose them to special risks; men’s behavior is not interpreted as provocative in the first place.

The irony is that the risks men impose upon women determine what behavior of women is interpreted as provocative. "Provocative behavior" is at least implicitly defined as that behavior which generates aggression by men towards women. Given that men do not bear those risks, men’s behavior cannot, by definition, be "provocative."

\textsuperscript{68} See, e.g., Terminiello v. Chicago, 337 U.S. 1 (1949); see also Laurence H. Tribe, \textit{American Constitutional Law} 853-55 (2d ed. 1988) (discussing \textit{Terminiello}).
provides economic opportunity to the disadvantaged. Hence it seems justified to devote more public resources to protection of those businesses despite the voluntary exposure of those businesses to the risks of crime.\textsuperscript{69}

This Section has explored two situations in which the equal protection principle should govern the distribution of protection. The first involves predominantly involuntary vulnerability; the second involves voluntary vulnerability which can be attributed to a socially valuable activity. The dichotomy between the two reasons for applying the equal protection principle is not as sharp as my analysis implies. For example, the justification for applying the equal protection principle in the context of rape victims can be based not only on the vulnerability of women to sexual crime, but also on the social importance of the activities which supposedly provoke their attackers. The freedom of women to dress as they wish and to behave in a manner that fits their identity is an important freedom which should be protected by the legal system. The protection granted by the Heckler's Veto doctrine to demonstrators can also be justified not only on the basis of the social value of freedom of speech, but also on the basis of the involuntary state of affairs which exposes those demonstrators to grave risk, namely the hostile reaction of others to their divergent beliefs. Although the demonstrators "contribute voluntarily" to the risk by publicly expressing their views, they generally do not voluntarily generate the hostility elicited by their views.

It is important to emphasize, however, that both justifications for applying the equal protection principle require making controversial normative judgments. These include judgments as to whether the vulnerability is primarily the byproduct of a voluntary action or an involuntary state of affairs; or judgments as to whether the voluntary action which generated the vulnerability has an important societal value. This Section provides only a general framework for applying the equal protection principle. The precise applicability of this framework to particular circumstances is bound to depend upon one's moral theory.\textsuperscript{70}

What are the implications of the equal protection model for the applicability of the criminal law principle of comparative fault? The equal protection model requires special protection to be granted to vulnerable victims in order to reduce their expected harm from crime to that of less vulnerable victims. In direct contrast, the criminal law principle of comparative fault requires granting lesser protection to careless victims. These victims are particularly exposed to crime and hence constitute an important class of vulnerable victims. Therefore, with respect to careless victims, use of the

\textsuperscript{69} Ed Rubin suggested this example to me.

\textsuperscript{70} Given the difficulty in applying the principles governing distribution of protection, it might be better to reject the principle of comparative fault. Such a conclusion may indeed result from a thorough investigation of the distributive implications of the current rule versus the expected distributive implications of the proposed one. This Part attempts to illustrate that such a conclusion does not follow.
equal protection model is inconsistent with employment of the criminal law principle of comparative fault. Indeed, if the equal protection principle had universal application, criminal law would rightly reject the criminal law principle of comparative fault. The next Subsection will illustrate that a different distributive principle should sometimes prevail and that this alternative distributive principle requires endorsement of the criminal law principle of comparative fault.

2. The Equal Costs Model

The equal costs model requires equal distribution of crime prevention resources to different potential victims. Under this model, it is unfair for some individuals to receive more resources for their protection even if those individuals are more vulnerable to crime without those additional resources. The equal costs model requires not an equality of the product or output of the enforcement system, namely protection, but equality in the resources devoted to the protection of different individuals.

One way of illustrating the persuasive force of this model is to think of protection costs as a collective fund which should be distributed equally among citizens. Every person should be given equal resources for her protection. Given the nature of protection, it may be more efficient to centralize the production of protection, but the distribution of resources should not differ from the one that would take place under a scheme that allocates to every individual equal resources for their protection. Providing more resources to the protection of some victims at the expense of other victims indicates that unjust priority is being given to the protection of the former over the protection of the latter.

Application of the criminal law principle of comparative fault offers an effective way of implementing the equal costs model. A negligent victim who contributes voluntarily to her vulnerability imposes larger ex ante costs on the enforcement system.\(^7\) Given the higher ex ante probability that a careless victim will become an actual victim of a crime, the expected cost of identifying, prosecuting, and punishing criminals who have committed crimes against careless victims is higher than the expected cost of identifying, prosecuting, and punishing criminals who have committed crimes against cautious victims. Unless the enforcement system adjusts its treatment of criminals by taking into consideration the behavior of victims, an unjustifiably large amount of resources will be devoted to the identification, prosecution, and punishment of criminals who have committed crimes against careless victims. Equalizing the costs of protection requires that fewer efforts be invested in the protection of careless victims relative to those invested in

---

71. It is also true that involuntarily vulnerable victims impose higher ex ante costs on the enforcement system. This should not deprive them of protection because, as I stated in the last Section, the model which governs their protection is the equal protection model rather than the equal costs model.
the protection of cautious victims. This can be done either by imposing lighter sanctions upon criminals who acted against careless victims or by reducing the resources devoted to the identification and prosecution of those criminals. The first option is tantamount to a recognition of a criminal law principle of comparative fault.

The equal costs model thus recommends the adoption of the criminal law principle of comparative fault. In the absence of such a principle, careless victims get priority over cautious ones. However, use of the equal costs model is appropriate only in certain contexts.

What is the proper context for implementing the equal costs model? When is it fair to "reward" a less vulnerable victim and "punish" a more vulnerable victim? It would seem, as discussed above, that the involuntary vulnerability of a victim should not lead to granting her lesser protection. Moreover, when the voluntary behavior generating the vulnerability has unique social value, it deserves full protection. In both cases, the equal protection model should determine the distribution of protection. But if the victim's exposure to the risk is both voluntary and of no special social value, the principle of equal costs should govern the distribution of protection and the equal protection principle should be disregarded.

3. The Scope of the Criminal Law Principle of Comparative Fault: Efficiency and Fairness Perspectives

Distributive justice considerations constrain the use of the criminal law principle of comparative fault. Under a distributive justice analysis, the criminal law principle of comparative fault should be used only when the

---

72. See supra notes 43-44 and accompanying text.
73. Sometimes the enforcement system takes an intermediate position which fits neither the equal protection principle nor the equal costs principle. The intermediate position may be justified on pragmatic grounds, but it violates fundamental principles of fairness.

What is the protection the enforcement system grants to a person who neglects to lock his door or take minimal precautions to protect his property? The protection a careless individual enjoys is undoubtedly lower than that required by the equal protection principle. Careless individuals are more exposed to crime than cautious individuals. Complying with the equal protection model would require the enforcement system to expend more resources in protecting careless individuals than in protecting cautious ones in order to guarantee equality in the expected costs of crime.

One could argue that the protection enjoyed by careless individuals is governed by the equal costs model, but this hypothesis cannot withstand scrutiny. The protection careless individuals enjoy is greater than that required by the equal costs model. Careless individuals impose larger ex ante costs on the enforcement system than cautious individuals. Given the higher probability of crime directed against careless individuals, the ex ante costs those individuals impose on the enforcement system are higher than the ex ante costs imposed on the enforcement system by cautious individuals. The equal costs model requires the enforcement system to exert less effort in tracing, prosecuting, or punishing individuals who harmed careless victims. But, it is often the case that the legal system refuses to take those factors into consideration and hence deviates from the equal costs model.

To sum up, our enforcement system seems to deviate both from the equal protection model and from the equal costs model. Careless individuals are more vulnerable to crime than cautious individuals. Hence, the requirements of the equal protection model are not satisfied. At the same time, careless individuals impose higher ex ante costs on the enforcement system than cautious individuals. Thus, the enforcement system does not satisfy the requirements of the equal cost model.
equal costs model supplies the proper distributive scheme. When it does not, the legal system should reject the comparative fault principle and instead embrace the principle that particularly vulnerable victims deserve greater protection than the average victim. As there are a limited number of criminal law contexts which demand use of the equal costs model, this Article's endorsement of a criminal law principle of comparative fault does not necessarily require a wholesale reform of the legal system.

It is important to note one primary difference between the scope of the principle as derived from the efficiency based analysis and its scope as derived from distributive justice considerations. Efficiency considerations require potential victims to take \textit{optimal} precautions, or those precautions which minimize the total costs of crime. Distributive justice considerations require potential victims to take \textit{reasonable} precautions, or those precautions which do not restrict their freedom unjustifiably or violate entrenched societal values. It may be efficient, and hence required by the efficiency considerations described in Part II, to require a woman to dress and act in a way which will reduce her chances of being raped (assuming a woman's attire can actually have such an effect), or to require vulnerable elderly persons to stay at home, but it is an unreasonable requirement because it unjustifiably restricts freedoms. Thus, the distributive justice perspective demonstrates that the criminal law principle of comparative fault, despite its theoretical attraction, is of limited scope. The general absence of comparative fault considerations from the criminal law is due not, as some scholars argue,\textsuperscript{74} to an absolute principle which forbids consideration of the victim's fault, but to distributive justice limitations.

Let me briefly summarize the discussion before turning to explore specific doctrines of criminal law. This Part examined criminal law from a distributive perspective, presenting two models which can govern the distribution of protection. Then, it analyzed the necessary conditions for the applicability of each model. Lastly, it showed that the equal costs model can be implemented by introducing a criminal law principle of comparative fault. Thus, it has been shown that both efficiency and fairness considerations sometimes support the introduction of a criminal law principle of comparative fault. Why has the legal system failed to recognize such a principle? The next Part will illustrate that the premise underlying this question is false. The criminal law principle of comparative fault is not merely a product of a fanciful, detached, scholarly imagination, lusting for originality. Judges and legislators already implicitly recognize a criminal law principle of comparative fault. It is the scholars who have yet to acknowledge its existence.

\textsuperscript{74} See, e.g., Fletcher, supra note 13, at 1672-78.
IV
REINTERPRETATION OF LEGAL REALITY: THE PRINCIPLE OF
COMPARATIVE FAULT AS PRACTICED IN CRIMINAL
LAW

Naturally, there is no a priori method for demonstrating that the criminal law has afforded limited recognition to the criminal law principle of comparative fault. Instead, one needs to review the doctrines of criminal law, explore its different rules, and extract those rules which can be interpreted as an implicit endorsement of a comparative fault principle. The rules explored in this Part are merely illustrative, not exhaustive, of the contexts in which the criminal law principle of comparative fault is implicitly endorsed by legislatures and courts.

Before embarking upon this enterprise, we need to distinguish the principle of comparative fault from other doctrines so that we do not mislabel those other doctrines and falsely identify them with the criminal law principle of comparative fault.

Two situations may be interpreted falsely as manifestations of the criminal law principle of comparative fault. In the first situation, the behavior of the victim may annul the responsibility of the perpetrator because it leads the perpetrator to lose his temper and consequently to act without criminal intent. The behavior of the victim in these contexts changes the mental state of the perpetrator in a way which annuls criminal responsibility. If an act of the victim leads the perpetrator to act without criminal intent, or with a different form of criminal intent, the acquittal of the perpetrator or the mitigation of his punishment should not be interpreted as a recognition of the criminal law principle of comparative fault.

Second, a victim who takes efficient precautions against crime may induce the criminal to commit a more serious offense in order to overcome those precautions. The criminal justice system’s imposition of harsher sanctions in such cases, though causally related to the victim’s precautions, is not an adoption of a criminal law principle of comparative fault. Instead, the more severe punishment is simply an indication of the more serious nature of the offense committed. For instance, if, as a result of precautions taken by a victim of theft, a thief uses force, convicting him of assault in addition to theft should not be interpreted as a manifestation of the criminal law principle of comparative fault.

Both situations share one common underlying feature: differences in sanctions imposed upon criminals are justified by differences in their culpability. In the first situation, the more lenient punishment is justified by the

75. The defense of provocation has traditionally been explained in this way. These arguments are non-consequentialist, i.e., they rely on the moral value of the behavior. I shall illustrate that the defense of provocation can also be explained on the basis of consequentialist arguments without refuting the traditional non-consequentialist arguments. Consequentialist arguments for a legal rule are those arguments which rely on the desirable or undesirable consequences of adopting such a rule.
less evil subjective state of mind of the criminal, whereas in the second situation, the harsher punishment is justified by the more serious objective nature of the act.

The next two Sections explore different rules of the criminal system and interpret them as manifestations of the criminal law principle of comparative fault. In addition to advocating my position, I hope to expose difficulties with alternative interpretations rooted in the two paradigms discussed above. However, at the same time, one has to recognize the possibility that the same legal doctrines can be explained by different considerations. The following discussion has a relatively modest purpose, namely to establish the claim that a criminal law principle of comparative fault can explain certain existing doctrines of criminal law without claiming that it is the only way by which those doctrines can be explained.

A. Violent Crimes

In this Section, I will explain how the doctrine of provocation and the no-retreat rule of the doctrine of self-defense can be described as instances of the criminal law principle of comparative fault.

1. Provocation

Provocation is often characterized as a unique doctrine because it is limited to the context of homicide. The surprising conclusion of this Article is that the doctrine is not peculiar to this context, but rather reflects an important feature of criminal law—the principle of comparative fault. Its alleged peculiarity is more a byproduct of the scholarly mind than a reflection of reality.

Criminal law distinguishes between murder and manslaughter. Murder is defined both at common law and by statute in many states as "unlawful

---

76. Scholars have struggled to explain the narrow scope of the doctrine. Some have suggested that the doctrine should be abolished because there are no reasons to distinguish homicide from other offenses. See, e.g., Stephen J. Morse, Undiminished Confusion in Diminished Capacity, 75 J. CRIM. L. & CRIMINOLOGY 1, 22-24 (1984) [hereinafter Morse, Undiminished Confusion]. Others have advocated the expansion of the doctrine of provocation to other offenses as well. See, e.g., Andrew Von Hirsch & Nils Jareborg, Provocation and Culpability, in Responsibility, Character, and the Emotions 241 (Ferdinand Schoeman ed., 1987). Lastly, some scholars have attempted to justify the narrow scope of the doctrine. One very popular explanation argues that fixed penalties for murder necessitate devising mechanisms to mitigate the harshness of the penalty. See Finbarr McAuley, Anticipating the Past: The Defense of Provocation in Irish Law, 50 MOD. L. REV. 133, 135 (1987).

Other explanations for the narrow scope of the doctrine of provocation can be found in the literature. See Martin Wasik, Partial Excuses in the Criminal Law, 45 MOD. L. REV. 516, 521 (1982) ("[I]t is a mistake to regard the fixed penalty and the availability of partial excuses as causally dependent on each other, as many writers have done. It is rather that both phenomena stem from a third consideration: the fact that murder is marked as a 'crime standing out from all others.'") (footnote omitted). For a discussion of the relation between the mandatory sentence of murder and the doctrine of provocation, see Celia Wells, The Death Penalty for Provocation?, 1978 CRIM. L. REV. 662, 668-72.
homicide with malice aforethought." One way to reduce a homicide charge from murder to voluntary manslaughter is to show that the intentional killing occurred in the heat of passion as a result of provocation. Provocation is an act of the victim (or someone else) which preceded the homicide and contributed to it. A successful claim of provocation is based on the co-existence of two dimensions: one subjective and one objective. The subjective dimension involves passion which is aroused by the provocation and which prevents cool reflection. The objective dimension requires the existence of an act which precedes the homicide and which is recognized by law as adequate to cause a reasonable person to lose his self control. In order to successfully use provocation as a defense, one needs to satisfy both the subjective test, that is, show an actual uncontrolled passion which caused the homicide, and the objective test, that is, show behavior conforming with the "reasonable person standard."

What can explain the law's recognition of the defense of provocation? Criminal law scholars have proposed two explanations. Some have argued that the doctrine is a partial excuse—a legal recognition of an excusable weakness of the actor—without endorsing the view that homicide carried out as a result of provocation is in any respect justified. Others have argued that the doctrine is a partial justification—the act (rather than the actor) is, in some respect, partially justified—and that the wrongdoing of the provoker makes the homicide less objectionable.

Both explanations have been criticized by scholars. Some have argued that provocation can be interpreted as a justification only if one assumes that the immoral conduct of provoking an attacker reduces the societal harm of the victim's death. Given that the victim's behavior never endangered the killer, the view of provocation as justified stands on dubious moral

---

77. See, e.g., CAL. PENAL CODE § 187(a) (Deering 1991); GA. CODE ANN. § 26-1101(a); IDAHO CODE § 18-4001 (Supp. 1993); IOWA CODE ANN. § 707.1 (West Supp. 1994).
79. See ROBINSON, supra note 78, § 102(b)(1).
80. See id. § 102(b)(3).
81. See id. § 102(b)(1) to (2).
84. See A.J. Ashworth, The Doctrine of Provocation, 35 CAMBRIDGE L.J. 292, 307; McAuley, supra note 76.
85. For a good articulation of the distinction between excuse and justification, see John Austin, A Plea for Excuses, 54 PROC. OF THE ARISTOTELIAN SOC'y 1, 3 (1956-57). Austin puts the question as follows: "[If the provoker] partly responsible because he roused a violent impulse or passion in me, so that it was not truly or merely me 'acting of my own accord' [excuse]? Or is it rather that, he having done me such injury, I was entitled to retaliate [justification]?"
ground. But regarding provocation as a justification does not commit one to the view that the value of one's life depends upon one's actions. Instead, one must be committed to the more plausible claim that "an individual is to some extent morally justified in making a punitive return against someone who intentionally causes him serious offence, and that this serves to differentiate someone who is provoked to lose his self-control and kill from the unprovoked killer."

It has also been argued that the subjective condition, or the requirement that the provoked person be in a special mental state, refutes the justification-based explanation of provocation. Arguably, justifications focus upon the act and not the actor and hence need not pay attention to the mental state of the perpetrator. I believe, however, that this argument is mistaken. Certain acts may be justified only if carried out for certain reasons.

The rationale based on excuse has its critics as well. Its main deficiency seems to be its inability to explain the criminal law's insistence on setting objective standards. The person who commits homicide out of unprovoked uncontrollable passion does not necessarily differ in the intensity of his emotions or in his mental state from the person who commits homicide as a result of provoked uncontrollable passion. Therefore, if provocation is indeed an excuse, the objective dimension of provocation cannot be explained, since the person who committed unprovoked homicide is not necessarily more culpable than the person who committed homicide as a result of a provocation satisfying the objective element.

But this critique of the excuse-based rationale rests on the premise that the availability of excuses cannot depend on the existence of objective conditions, namely, conditions unrelated to the mental state of the perpetrator.

85. See Dressler, Rethinking Heat of Passion, supra note 83, at 457 ("It is morally questionable to suggest that there is less societal harm in Victim's death merely because he acted immorally. One must remember that Victim's immoral conduct in no way jeopardized the life of the defendant or anyone else.") (footnote omitted).
86. Ashworth, supra note 84, at 307.
87. See Dressler, Provocation, supra note 83, at 479-80.
88. See Paul H. Robinson, A Theory of Justification: Societal Harm as a Prerequisite for Criminal Liability, 23 UCLA L. Rev. 266, 291 (1975) ("[T]he defense of justification should remain available in every situation for which no resulting harm can be demonstrated, regardless of any actor-oriented considerations such as prior fault, motive, belief, or knowledge.") (footnote omitted).
89. See George P. Fletcher, The Right Deed for the Wrong Reason: A Reply to Mr. Robinson, 23 UCLA L. Rev. 293, 295 (1975) ("[T]he theory of justification to make justificatory defenses available only to those whose intent is meritorious."). But see supra note 84 and accompanying text.
90. See Ashworth, supra note 84, at 307-14.
One objection to this premise is a moral one. Arguably, in order to claim an excuse, a person cannot rely solely on her failure to acquire knowledge or have a certain mental state, but should also establish that acquiring this knowledge or being in a certain mental state could not reasonably be expected from her. Consequently, a person who failed to know a relevant fact he should have known or to restrain his temper in a context where he could have been expected to do so should not be excused.

This is merely a brief and incomplete sketch of the main non-consequentialist accounts of the doctrine of provocation. It is beyond the aims of this Article to resolve the controversies raised by those arguments. For my purposes, it is enough to note that the controversy between the justification-based and the excuse-based rationales for the provocation doctrine has not been resolved. I now examine whether consequentialist arguments, in particular efficiency and fairness arguments, can provide insight into the logic of the doctrine of provocation, and perhaps yield compelling rationales for it.

91. See Jeremy Horder, Cognition, Emotion, and Criminal Culpability, 106 LAW Q. REV. 469, 480-81 (1990) (suggesting expansion of the permissible boundaries of culpability beyond instances in which defendants were aware of possible wrongdoing); Celia Wells, Swatting the Subjectivist Bug, 1982 CRIM. L. REV. 209 (arguing that objective standards have an important role in determining culpability).

The introduction of negligence as an appropriate basis for criminal liability leads to similar conclusions. For a defense of negligence as a proper basis for imposing criminal liability, see George P. Fletcher, The Theory of Criminal Negligence: A Comparative Analysis, 119 U. PA. L. REV. 401 (1971).

92. In the context of provocation, this view was advocated by Dressler, who states:

A person who becomes sufficiently enraged to kill because the decedent acted in a nonwrongful manner arguably does not deserve to be excused. At the least, the nonwrongfulness of the decedent's actions is highly pertinent in determining whether the actor's loss of self-control was excusable. Thus, the individual who becomes angry and responds violently when another justifiably strikes him in self-defense and the person who unjustifiably creates the situation in which the provocation gives birth are blameworthy and should not be excused.

Dressler, Provocation, supra note 83, at 475 (footnote omitted).

A similar position is taken by Fletcher. See George P. Fletcher, Rethinking Criminal Law 247 (1978) ("Obviously, there are some impulses such as anger and even mercy . . . that we do expect people to control. If they fail to control these impulses and they kill another intentionally, they are liable for unmitigated homicide or murder. The basic moral question in the law of homicide is distinguishing between those impulses to kill as to which we as a society demand self-control, and those as to which we relax our inhibitions.").

93. For an explanation of the distinction between consequentialist and non-consequentialist arguments, see supra note 75.

94. Many have denied the value of consequentialist arguments in this context. See, e.g., Dressler, Rethinking Heat of Passion, supra note 83, at 434 n.136 (contending that utilitarian principles do not help to explain provocation doctrine).

95. There have been previous attempts to justify the doctrine of provocation on efficiency grounds. According to Posner, a person who commits a homicide as a result of provocation is more likely to be discovered and successfully prosecuted than a person who commits an unprovoked homicide (who presumably can plan an escape). Moreover, the person retaliating for provocation is less likely to succeed in committing homicide and hence the ex ante risks he imposes on society are lower. Therefore, in order to achieve an optimal deterrent effect criminal law needs to impose harsher sanctions on the murderer than on the person who commits manslaughter. See Posner, Economic Analysis, supra note 25, at 219.
Before exploring these arguments, it is necessary to clarify and defend an important premise. Many crimes, in particular passion crimes, are not committed with a precise cost-benefit calculation. This, however, does not mean that criminal sanctions do not influence the perpetrators of passion crimes. The legislature can affect the behavior of potential killers or of potential provokers in an indirect fashion. Imposing severe sanctions for crimes committed out of passion is likely to lead to a more restrained society as a whole and hence to a society where the ex ante costs of provocation are lower. Members of a society which severely sanctions provoked homicide may develop both more restrained reactions to provocations and a more intense disposition to provoke given that provocations are less likely to result in homicide. Without unrealistically assuming a rational calculation of costs and benefits in crimes of passion, the legal system can partly determine the dispositions of potential killers as well as the dispositions of potential provokers.9

To understand the rationale for the doctrine of provocation, assume a legal system that refuses to recognize provocation as a partial defense. Individuals who are provoked to commit homicide would be treated in the same manner as people who commit unprovoked homicide. Given the premise established above, such a system would presumably succeed better than our current system in encouraging those individuals who have been provoked to resist the temptation to commit homicide. The restrained attitude of provoked individuals would have an impact on the dispositions of provokers: a potential provoker would be more likely to provoke given that her provocation would be less likely to result in homicide. Eliminating the defense of provocation may therefore lead to an increase rather than a decrease in cases of homicide because it may increase the total number of provocations.

Assume, now, a legal system which regards provocation as a complete defense, that is, a legal system in which the existence of provocation will lead to a complete acquittal of a person who commits homicide. In such a society, the rate of provocation is likely to be very low; but any given provocation is more likely to result in homicide. Eliminating the defense of provocation may therefore lead to an increase rather than a decrease in cases of homicide because it may increase the total number of provocations.

Posner's arguments, however, focus on the subjective mental state of the perpetrator rather than on the objective circumstances: a person who was subjectively provoked is less likely to succeed in committing murder and more likely to be caught. The question whether the provocation was sufficient to provoke a reasonable person is irrelevant. Hence, Posner's analysis fails to explain the objective component of the doctrine of provocation.

96. Support for this claim is not limited to advocates of law and economics. In his analysis of provocation, Ashworth argues as follows:

It is wise to be sceptical [sic] when deterrent arguments are applied to impulsive crimes, especially when the only direct deterrent effect would be limited to the difference between the penalty for murder and the probable sentence for manslaughter upon provocation. On the other hand, there might be an indirect effect stemming from the role of law in supporting or reinforcing a code of social behaviour which deprecates loss of temper.

Ashworth, supra note 84, at 310-11 (emphasis added) (footnote omitted).
defense of provocation may therefore increase the total number of homicides.

A legal system in which provocation functions as a partial defense provides incentives both for the potential victim to avoid provocation (by reducing the punishment levied on persons who commit homicide as a result of provocation) and for the person who considers committing homicide to avoid carrying it out (by imposing criminal sanctions upon him). Recognizing provocation as a partial defense promotes efficiency by providing appropriate incentives for the potential provoker as well as the potential killer.

Before turning to fairness considerations underlying the provocation doctrine, let me briefly address two possible objections to the efficiency-based rationale for the doctrine. First, it might be argued that the rationale cannot explain the subjective component of provocation, namely the requirement that the defendant actually lose control over her actions. If criminal law aims to deter provokers, it should be indifferent to the mental state of the defendant. This objection ignores the importance of balancing the need to provide incentives to the provoker with the need to provide incentives to the provoked. Where the provoked person has not lost her control, it may be more efficient to provide incentives to the provoked person to restrain herself than to deter provokers from provoking by mitigating the punishment of the provoked person. The subjective component of provocation is not therefore necessarily incompatible with the efficiency-based argument.

Second, the efficiency-based argument may be seen as incompatible with the recent recognition of misdirected retaliation, in which the provoked person acts against an innocent third party. There has been some controversy as to whether a person who misdirects retaliation should benefit from the defense of provocation. But recognition of the misdirected retaliation doctrine does not undermine the efficiency-based justification of provocation. Depriving a defendant of the defense of provocation simply because the victim was not the provoker imposes a risk upon the provoked person—a risk that he may, in the heat of passion, retaliate against the wrong person. This risk may deter potential killers and hence, in turn, may provide incentives to increase the frequency of provocations. Recognizing misdirected retaliation as a basis for mitigating the sentence cannot, therefore, refute the efficiency-based justification.


98. Arguably, the weakness of the efficiency-based justification of provocation is its inability to provide precise tests to measure the cumulative behavioral impact on provokers and provoked individuals.

Measuring the cumulative effects of providing incentives to provokers or provoked person can only be done, if at all, with empirical psychological tools which are beyond the scope of this Article. The
Fairness considerations based on the equal costs model also support the doctrine. The provoker imposes extra costs on society. His voluntary conduct (the provocation) lacks significant social value and therefore should not be rewarded by society. Therefore, a fair distribution of protection should be determined by the equal costs model, which requires expending equal resources for the protection of provokers and nonprovokers. Given the higher probability that provokers will become actual victims, society should expend fewer resources to punish people who commit homicide as a result of provocation.99

Provocation is but one instance in which criminal law grants certain privileges to the perpetrator of a crime directed towards a victim who is partly to blame for the crime directed against him. The scheme provided in this Article can also help to expose the structural similarity between the defense of provocation and other doctrines of criminal law granting privileges of this type. The next Subsection will explore whether the consequentialist logic underlying provocation can also explain the boundaries of the doctrine of self-defense, which exculpates individuals who commit crimes in order to protect themselves from an apparent threat of injury or death.

2. The "No Retreat" Rule

Drawing the boundaries of self-defense is difficult because violent encounters often involve a process of escalation to which both parties contribute. The scope of self-defense should be delimited to provide appropriate incentives to the initial aggressor as well as to her victim. The problem is, however, that any change in the boundaries of self-defense has different and often contrasting impacts on the behavior of the parties to a violent encounter. Depriving one person (e.g., the initial aggressor) of the privilege to claim self-defense provides him an incentive to withdraw from the violent encounter; but, precisely for this reason, it provides the other party an incentive to pursue violence. The scope and the qualifications of self-defense as developed in common law and in statutory law convey a subtle recognition of the balance to be drawn between the incentives of the initial aggressor and those of his victim.

Let me illustrate this claim by examining the no retreat rule. Generally speaking, force may not be used in self-defense unless it is reasonably necessary for the protection of the victim. A major exception to this necessity requirement is the no retreat rule. Under this rule, the defendant is permit-

---

99. It is also plausible to argue that the provocation doctrine is endorsed through the use of other mechanisms such as the enforcement policy of the police forces and the use of discretionary powers of judges and juries in criminal cases. See supra notes 43-45 and accompanying text.
ted to use deadly force to combat a deadly attack directed against him even if the defendant could have avoided using the deadly force by retreating.100

Adoption of a no retreat rule creates an asymmetry between the initial aggressor and the initial victim. While the initial aggressor has a duty to retreat if he wants to be exculpated from criminal responsibility, his victim has no such duty. A no retreat rule increases the risk taken by initial aggressors by removing the incentives of their victims to retreat. At the same time, it may provide an incentive to the victims of an initial aggressor to resort to violence.101

Admittedly, the tendency of modern codes is to narrow the scope of the initial victim’s privileges. However, even the Model Penal Code maintains a modified version of the no retreat rule. While rejecting the traditional no retreat rule, the Code specifies an important qualification.102 Under this qualification, the no retreat rule still applies when the defendant was attacked in his dwelling or in his place of work, so long as the defendant was not the initial aggressor.103

The no retreat rule, even in its more limited modern manifestation, can be interpreted as an instance of a broader principle—the no fault rule—which imposes limitations on the use of self-defense. Under the no fault rule, a claim of self-defense is available only if the defendant is “without fault in bringing the conflict on.”104 Hence, it is denied to “slayers who incite the fatal attack, encourage the fatal quarrel or otherwise promote the necessitous occasion for taking life.”105 In other words, in delineating the boundaries of self-defense, the legal system takes into account—as required by the principle of comparative fault—the fault of the victim.

Admittedly, this explanation of the no retreat rule contributes little to the resolution of the dispute concerning just where the boundaries of self-defense should lie. It, however, highlights the underlying considerations behind the dispute and illustrates that determining the scope of self-defense is an instance of a much broader phenomenon of criminal law—the phenomenon which is the subject matter of this Article.

The treatment of violent encounters is complex because of the need to provide incentives to both parties to act in a way which will minimize harm. Granting privileges to the initial victim of an aggression and denying those privileges to the initial aggressor are means by which the legal system can

100. For a discussion of the no retreat rule and its justifications, see JOSHUA DRESSLER, UNDERSTANDING CRIMINAL LAW 196-97 (1987).
101. This argument against the no retreat rule has influenced modern penal codes. See MODEL PENAL CODE AND COMMENTARIES § 3.04 commentary at 52-54 (Official Draft and Revised Comments 1985).
102. Id. § 3.04(2)(b).
103. Id. § 3.04(2)(b)(ii)(A). The provisions in subsection (A) maintain a limited asymmetry between the initial aggressor and the initial victim and grant the latter privileges based on the relative fault of the former. See id. § 3.04 commentary at 55-57.
105. Id. at 1231 (footnote omitted).
deter initial aggression. However, the more protection given to the initial victim, the greater is the risk that the initial victim will have no incentives to break off the violence by withdrawal or by other means. Similarly, the more protection given to a provoker, the greater the frequency of provocations; while the less protection given to the provoker, the greater the frequency with which provocations lead to violence.

Underlying efficiency and fairness considerations support application of the principle of comparative fault not just in the context of violent crimes against the person, but in the context of property crimes as well. Leaving one's door unlocked, for example, is structurally similar to provocation in the sense that in both cases the victim can be regarded as partially responsible for the crime. Hence, these cases may require a similar response from the legal system. The next Section will expose an anomaly in the way that property offenses are classified and will argue that the classification of property offenses is intelligible only if interpreted as a subtle recognition of the comparative fault principle.

B. The Enigma of the Classification of Property Offenses

The legal system distinguishes among various property crimes, including theft, burglary and robbery. I shall argue that the distinction between these offenses embodies a subtle recognition of the principle of comparative fault. The greater the precautions taken by a victim of a property offense, the more severe the sanctions that are imposed upon the criminal and consequently the greater the protection that is granted to the potential victim by the legal system. By varying the sanctions imposed upon the criminal according to the efforts made by the victim to prevent the crime, the legal system provides incentives for the victim to take better precautions and distributes fairly the costs of protection.

The distinction between simple theft and robbery or between simple theft and burglary has been traditionally explained on the grounds that robbery and burglary are more serious crimes than theft. The more severe sanctions for robbery or burglary are simply reflections of the more serious nature of the offense. Is this traditional explanation satisfactory?

Robbery is defined in the Model Penal Code article 222 as follows:

A person is guilty of robbery if, in the course of committing a theft, he:

- recklessly inflicts serious bodily injury upon another; or
- threatens another with or purposely puts him in fear of immediate serious bodily injury; or
- commits or threatens immediately to commit any felony of the first or second degree.  

106. I am grateful to Yoram Schachar for first bringing this idea to my attention.
Robbery can also be described as larceny committed under aggravating circumstances. A larceny is aggravated (and hence becomes a robbery) when (1) the property is taken from the person or presence of another, and (2) the taking is accomplished by the use of force or threatened force. Robbery often (although not always) involves, therefore, battery plus larceny or assault plus larceny.

It is worthwhile exploring why the legal system creates this combined offense rather than simply prosecuting the robber for the different components of robbery, namely for assault or battery and theft. The puzzle becomes especially acute given that the combined offense is treated much more harshly than the aggregation of its components. Why should the robber not be punished in the same manner as a person who commits both a theft and assault but does not combine them in the manner the robber does?

One possible answer relies on the fact that some robbers might, under current law, escape conviction for assault or battery. Arguably, the special attention of criminal law to robbery is a way to punish those criminals who currently can be convicted of robbery, but who cannot be convicted of assault or battery. This explanation fails for two reasons. First, the broad definitions of assault and battery in criminal law guarantee that most of the criminals whose activities currently fall into the definition of robbery could be prosecuted for assault, battery, or other offenses. Second, those robbers who could not independently be prosecuted for assault or battery are the least dangerous robbers. If a robber's threat to the victim is not itself punishable under criminal law, it does not seem reasonable that the robber should be punished more severely than an ordinary thief.

The commentary to the Model Penal Code suggests another explanation for the special severity of the sanction for robbery:


110. Commentators who have pointed out this phenomenon have not found it particularly surprising. See Note, A Rationale, supra note 108, at 102.

111. The argument was discussed in the commentary to the Model Penal Code as follows:

Robbery appears, then, to consist of a combination of theft and actual or threatened injury, each element constituting, at least in our day and under this Code, a separate crime. It might be thought sufficient, therefore, to prosecute for these crimes, cumulating punishment where appropriate. Closer analysis reveals that the premise is not accurate, and that the conclusion does not follow. Many threats are not criminal, apart from special circumstances. For example, a threat (as distinguished from actual attempt) to punch someone in retaliation for a slight is generally not criminal. Only a minority of the states provide [sic] misdemeanor penalties even for coercive threats, i.e., those designed to secure some non-pecuniary concession from the person threatened.


Even if all threats were subject to minor penalties, e.g., as "disorderly conduct," the combination of penalties for a petty theft and petty threat or minor violence by no means corresponds to the undesirability and danger of the offense. The violent petty thief operating in the streets and alleys of our big cities, the "mugger," is one of the main sources of insecurity and concern of the population. There is a special element of terror in this kind of depredation. The ordinary citizen... abhors robbers... just as he abhors burglars who penetrate the security of his home or shop. In proportion as the ordinary man fears and detests such behavior, the offender exhibits himself as seriously deviated from community norms, requiring more extensive incapacitation and restraining.\(^{113}\)

The commentary argues that the fear and terror caused by physical confrontation with the criminal require more serious treatment of the robber than the mere thief. However, the same physical confrontation takes place in any instance of battery or assault. The commentary fails, therefore, to illustrate the special quality of the combination of theft and battery or theft and assault which justifies a harsher treatment of the robber.

The enigma of the classification of property offenses is not limited to the case of robbery. A similarly inexplicable phenomenon can be identified in the common law distinction between theft (or other crimes) and burglary.

The Model Penal Code defines burglary as follows:

A person is guilty of burglary if he enters a building or occupied structure, or separately secured or occupied portion thereof, with purpose to commit a crime therein, unless the premises are at the time open to the public or the actor is licensed or privileged to enter.\(^{114}\)

Burglary is therefore a conjunction of an independent offense—criminal trespass\(^{115}\)—and an intention to commit theft (or other crime). Unlike the case of robbery, burglary does not even consist of two independent offenses, but consists merely of one punishable offense and an independently nonpunishable intention to commit a crime. As in the case of robbery, it is not merely the existence of a combined offense which is puzzling. It is also inexplicable why the combined sanctions imposed upon criminal trespass and an intention to commit theft or another crime are relatively light in comparison to the severe sanctions imposed upon the burglar.\(^{116}\)

\(^{113}\) MODEL PENAL CODE § 222.1, at 69 (Tentative Draft No. 11, 1960).

\(^{114}\) MODEL PENAL CODE AND COMMENTARIES § 221.1 (Official Draft and Revised Comments 1980).

\(^{115}\) Criminal trespass is entering or remaining on land or in a structure or vehicle by one who knows he is not authorized or privileged to do so and either defies the owner's order by entering or remaining on the property, or ignores a posting or other indication that he is intruding. BLACK'S LAW DICTIONARY 1347 (5th ed. 1979).

\(^{116}\) This has not escaped the notice of the Model Penal Code commentators. See MODEL PENAL CODE § 221, at 57-59 (Tentative Draft No. 11, 1960).
The commentary to the Model Penal Code concedes that the very existence of burglary as a distinct offense can be justified only in historical terms. The drafters add that they would have been content to eliminate the offense altogether had it not been so entrenched in common law systems.\textsuperscript{117} In support of their position advocating the elimination of burglary as a separate offense, they mention the absence of an equivalent offense in civil law countries. Civil law countries instead penalize crimes involving intrusion into another's home by adding a minor term of imprisonment for criminal trespass to the sentence for the other crimes committed or attempted, or sometimes regard the commission of certain crimes in inhabited dwellings as an aggravating circumstance.\textsuperscript{118}

Alternatively, one could argue that the severe treatment of burglary is accounted for by the sanctity of the home.\textsuperscript{119} Intrusion into one's "castle" is a serious threat to one's privacy and comfort and the legal system should acknowledge the special severity of this invasion.

However, this explanation fails to address the comparatively forgiving attitude of the Model Penal Code (as well as many other criminal codes) towards criminal trespass.\textsuperscript{120} It is unclear why the same habitation values violated so blatantly by the burglar are not also violated to the same degree by the criminal trespasser.\textsuperscript{121} Moreover, an account based on habitation values cannot explain the expansion of the definition of burglary under modern statutes. Burglary can now be committed not only in dwellings, but in almost any building and sometimes even in a vehicle. For example, the New York Penal Code defines burglary as any unlawful entry into a building;\textsuperscript{122} moreover, it broadly defines building to include, "in addition to its ordinary meaning . . . any structure, vehicle or watercraft used for overnight lodging of persons, or used by persons for carrying on business therein, or used as an elementary or secondary school, or an inclosed motor truck, or an inclosed motor truck trailer."\textsuperscript{123} This expansion of burglary, not atypical

\begin{itemize}
  \item[117.] See id. at 57.
  \item[118.] See Note, Statutory Burglary—The Magic of Four Walls and a Roof, 100 U. PA. L. REV. 411, 424-27 (1951) [hereinafter Note, Statutory Burglary].
  \item[119.] See Note, A Rationale of the Law of Burglary, 51 COLUM. L. REV. 1009, 1020-21 (1951) [hereinafter Note, Burglary]. This explanation finds support in traditional common law. Burglary has traditionally been classified as a crime against the habitation. See, e.g., Stowell v. People, 90 P.2d 520, 521 (Colo. 1939); State v. Surles, 52 S.E.2d 880, 882 (N.C. 1949).
  \item[120.] Criminal trespass in the Model Penal Code is a petty misdemeanor. Even if committed at night—when the violation of habitation values might be greatest—the MPC classifies the offense only as a misdemeanor. See Model Penal Code and Commentaries § 221.2(1) (Official Draft and Revised Comments 1980).
  \item[121.] See Note, Burglary, supra note 119, at 1021 ("[A] completed trespass accompanied by intent to commit a crime does not differ from a common trespass in what has occurred . . . .").
  \item[122.] N.Y. Penal Law § 140.20 (McKinney 1989).
  \item[123.] Id. § 140.00.
\end{itemize}
of modern statutory codes, is inconsistent with the argument that severe penalties for burglary protect habitation values.\textsuperscript{124} Can we explain the harsh treatment of burglars relative to that of trespassers in terms of the additional crimes burglars intend to commit? Creating a separate offense seems at best cumbersome. The severe sanctions imposed for burglary can be explained only by maintaining that the statutory penalties specifically attaching to the conduct of the burglar would be inadequate in the aggregate.\textsuperscript{125} A simpler remedy would be to revise the sanctions for these crimes.\textsuperscript{126}

Yet, what if the harm inherent in the commission of some crimes varies depending on the context in which it is committed? Committing some crimes within a building may be intrinsically more dangerous and harmful than committing those very same crimes outdoors. Even so, the natural way to reflect the severe nature of those crimes when committed in dwellings would be to regard the commission of those crimes within a building as an aggravating circumstance rather than impose a harsher sanction upon all crimes committed within a building.\textsuperscript{127}

Lastly, it could be argued that two or more crimes directed toward one victim might ruin the life of the victim while committing several crimes toward different victims is less likely to have disastrous consequences. "Distributing" the costs resulting from a criminal activity among several

\textsuperscript{124} This incompatibility has not gone unnoticed in the scholarly literature. See, e.g., Note, Statutory Burglary, supra note 118, at 433.

\textsuperscript{125} The burglar need not commit any crime in addition to the trespass, but often he can be convicted of an attempt to commit the crime he intended to commit.

\textsuperscript{126} See Note, Burglary, supra note 119, at 1021 ("If greater deterrent measures are advisable for particular criminal conduct other than burglary, they should be provided through the medium of the statute which declares such conduct to be criminal.").

\textsuperscript{127} See id. at 1021-22. It could be argued that all crimes committed in the circumstances which are currently characterized as burglary present special risks because they generate a high probability of incidental crimes unrelated to the original criminal intent of the burglar. The victim is likely to commit violence against the invader if the invader is a burglar, i.e., if the invader intends to commit an additional crime; while if the invader is merely a trespasser, the victim is unlikely to use violence against him. See id. at 1022-23. However, despite its appeal, this explanation is empirically dubious. It assumes that victims can make fine distinctions between those who invade their homes in order to commit crimes and those who enter their homes for other purposes.

Another explanation, suggested to me by Alex Stein, is that the severe sanctions imposed upon the burglar or the robber reflect the risks imposed upon the society by the conjunction of multiple criminal acts in a concerted course of conduct. Arguably, the willingness to commit different crimes (e.g., assault or battery and theft or criminal trespass and theft) in a coordinated manner demonstrates a more evil nature than the willingness to commit those very same crimes sporadically, with no clear general purpose. Discriminating between the robber and the thief or the burglar and the thief is a means of distinguishing between those who commit crimes which are related to each other in an organic way and those who commit those very same crimes with no overriding plan.

While it is true that the coordinated commission of several offenses may often reflect a more evil criminal nature than a sporadic commission of those very same offenses, this argument is inapplicable in our context. The battery and theft are so naturally related in cases of robbery that no particularly sophisticated or evil mind is required to combine them. The same applies to the relation between criminal trespass and theft. The combination of theft and assault or theft and criminal trespass is not indicative of a mind more evil than that of a criminal who commits these offenses separately.
persons seems less harmful than imposing all of these costs upon one person. Hence, while the commission of theft and assault or theft and trespass directed against one victim may not reflect the more evil nature of the criminal, it is simply objectively more harmful and hence deserves harsher treatment.

This is indeed an ingenious attempt to salvage the intelligibility of robbery and burglary. It is too ingenious, for why limit the principle to burglary and robbery? Why not apply it in the context of other offenses? If this explanation truly reflected the logic of criminal law, then the very nature of criminal law would be different from what it is today.

The enigma of classification may be summarized as follows: the traditional explanation for the classification of property offenses is based on the relatively severe nature of robbery and burglary relative to theft. Arguably, robbery and burglary are more serious offenses because they involve not only an attack upon property, but also an attack upon the body of the owner or the possessor (robbery) or a serious invasion of one's privacy and comfort (burglary). Hence, the legal system needs to express the serious nature of these crimes through the sanctions it imposes.

But the traditional explanation fails to clarify the peculiar severity of the penalties for the complex offenses of robbery and burglary relative to the comparatively light penalties for the components of these offenses. If, indeed, these complex offenses are more serious than theft alone only because they involve an intrusion upon the sanctity of human body or home, the criminal system could reflect their heinous nature by convicting and punishing the criminal separately for intruding upon each one of those values.

Given the failure of the traditional account in explaining the special treatment of robbery and burglary, can the special treatment be justified on other grounds? Can we explain what otherwise may be regarded as an exaggerated reaction of the legal system to those forms of criminal activity? I believe that the otherwise enigmatic classification of property offenses can be understood and justified on the basis of the criminal law principle of comparative fault.

The case of theft is a case in which the victim took no special precautions to protect his property. In contrast, robbery is a case where the victim held the object close to her body and was forced to give it up; and burglary is a case where the object was taken from the victim's home. In both robbery and burglary the victim held the object in what may be called her "territory." The most basic precaution required of the victim (if she wants to benefit from special protection of her property) is to hold that property in what is socially and legally understood to be her territory. By imposing harsher sanctions on the robber or the burglar, the legal system rewards the victim for taking that basic precaution.
The particular severity of the penalty imposed for robbery and burglary does not reflect, then, the harm caused by the theft plus the harm caused by additional offenses attempted or committed by the criminal. Instead, the severe sanctions can be justified as a means of "rewarding" cautious potential victims. Such a "reward" is needed in order to promote both efficiency and distributive justice. By providing the potential victim with additional incentives to take care of her property, the reward helps spur efficient behavior. Equalizing the costs of protecting cautious and careless victims serves distributive justice.

Admittedly, this explanation is based on a generalization about the degree of caution exercised by victims of the crimes at issue. One can think of examples which would constitute theft under criminal law without any negligence on the part of the victim. Furthermore, one could conceive of instances of robbery or burglary which involve careless behavior by victims. Criminal law as well as law in general is bound, however, to use generalizations. These generalizations may be too crude to capture the subtleties of reality. It is certainly possible that the traditional classification of property offenses is indeed too crude and in need of refinement. My argument is limited to demonstrating that efficiency and fairness are plausible rationales for this classification; perhaps more explicit recognition of their roles may ultimately lead to refinement of the classification of property offenses.

128. The term "rewarding" in this context should be taken metaphorically. It is not that we provide the cautious victim better protection because of her better behavior, but simply that we want to influence the behavior of potential victims as well as to distribute fairly the costs of protection. It does not necessarily involve any judgment as to her moral character.

129. The examples used are examples drawn from substantive criminal law. However, as has been emphasized (see supra notes 43-45 and accompanying text), the doctrine of comparative fault can be implemented through law enforcement mechanisms, such as police investigation and judicial sentencing, which can be made sensitive to the conduct of the victim.

The Uniform Sentencing Guidelines provide a concrete illustration of the influence of the notion of comparative fault on federal sentencing.

Take, for instance, section 5K2.10 of the 1994 edition of the United States Sentencing Commission's Federal Sentencing Guidelines Manual, entitled "Victim's Conduct." The section states, "If the victim's wrongful conduct contributed significantly to provoking the offense behavior, the court may reduce the sentence below the guideline range to reflect the nature and circumstances of the offense." U.S. SENTENCING COMM'N, FEDERAL SENTENCING GUIDELINES MANUAL § 5K2.10 (West 1994) [hereinafter GUIDELINES MANUAL]. The section also specifies a list of considerations which have to be taken into account, including: the size and strength of the victim and other relevant physical characteristics in comparison with those of the defendant, the persistence of the victim's conduct, and the efforts of the defendant to prevent confrontation. The section specifically excludes criminal sexual abuse from its scope as well as nonviolent offenses. But this latter exclusion is not absolute. The Guidelines Manual states that there may be unusual circumstances in which substantial victim misconduct would warrant a reduced penalty for a nonviolent offense.

In contrast, section 3A.1.1 of the Guidelines Manual, entitled "Vulnerable Victim," can be interpreted as a recognition of the equal protection principle. The section recommends that the judge increase the sanction "[i]f the defendant knew or should have known that a victim of the offense was unusually vulnerable due to age, physical or mental condition, or that a victim was otherwise particularly susceptible to the criminal conduct." GUIDELINES MANUAL, supra, § 3A.1.1.
The principle of comparative fault is implemented in the American
criminal law system, though under a disguise. But why is there a need for
such a disguise? Why is tort law willing to boldly affirm the centrality of a
principle of comparative fault, while criminal law refuses to whisper even a
silent confession? The legal system may be reluctant to confess the preva-
lence of the principle in criminal law because of the moral indignation that
the principle would arouse in the typical observer. But the mere indignation
such a principle may arouse is not in itself a reason to dismiss it. The
sources of this indignation must first be explored and found to be valid.
The next Part will explore some of the reasons we viscerally recoil from the
use of a principle of comparative fault in the criminal law context.

V

Comparative Fault and Retributive Sentiments

Why does a criminal law principle of comparative fault cause moral
indignation? Have we neglected some important moral considerations
which move our sentiments, but for which the productive-distributive para-
digm fails to account?

The use of a principle of comparative fault in criminal law may be
viewed as immoral because it indicates a differential moral evaluation of a
criminal who committed a crime directed towards a careless victim and a
criminal who committed a crime directed towards a cautious victim. Given
that the caution or carelessness of the victim seems irrelevant to the culpa-
bility of the criminal, the principle apparently discriminates unjustifiably
between criminals who are equally culpable because they committed similar
acts while in similar mental states. Because this objection is based on the

1 I am grateful to Sarah Beale for drawing my attention to the relevance of the Sentencing
Guidelines.

130. This objection is based on a confusion between the state’s right to punish and the state’s duty
to punish. The state may indeed have a right to punish equally individuals who have committed equal
wrongs under similar circumstances, but it does not necessarily have a duty to punish them equally.
When efficiency or distributive justice concerns require differential treatment, the state is entitled or
even obliged to take those concerns into account.

A different moral objection would be that my analysis obscures the role of individual rights in
criminal law. This objection comes in two forms. First, the analysis proceeds from a flawed perspective
because it does not focus upon rights. Second, the use of a criminal law principle of comparative fault
violates preexisting rights.

The first version of the objection simply takes a different perspective of criminal law (a rights-
based perspective) and argues that this perspective should be dominant or exclusive. I believe, however,
that an exclusive focus upon rights cannot provide a sufficient foundation for a system of criminal law.
Rights should be regarded as constraints, and within those constraints other considerations may play an
important role.

The second version is incomplete without specifying the rights which supposedly are violated by
the doctrine of comparative fault. Is it, as argued by some, the right of criminals who acted against
careless victims to be punished? See Duff, supra note 19, at 262-66. Is it the right of criminals who
acted against cautious victims to be treated in the same manner as criminals who acted against careless
ones? Or is it the right of cautious victims to the protection of criminal law? The latter hypothesis
seems to be the most promising. However, in the absence of a comparative fault doctrine, victims who
belief that the aim of criminal sanctions is to punish culpable individuals, I shall term it the retributivist objection, and explore whether the principle of comparative fault can be interpreted in a way which accommodates this objection.

The core of the retributivist objection is the assumption that the severity of criminal sanctions necessarily reflects a moral judgment concerning the degree of culpability of criminals. This premise is not always accurate. Culpability is indeed a necessary condition for imposing punishment, but the severity of punishment may be influenced by other factors. Two interpretations of the criminal law principle of comparative fault can explain the differential severity of sanctions imposed upon criminals without relying on a moral judgment concerning the criminals' culpability. First, the principle of comparative fault can be understood to express a moral judgment concerning victims—either a condemnation of careless victims or praise of cautious ones. Alternatively, one can adopt the principle of comparative fault without making any moral judgments about either criminals or victims.

According to the first interpretation, the comparative fault principle can be justified on the basis of the culpability of careless victims. The careless victim is irresponsible. He imposes unjustified costs on society and consequently he deserves lesser protection. Some forms of provocative behavior by victims fall under this category and may explain statements about victims such as "he deserved it." A different, although analogous, description would avoid the condemnation of careless victims and instead would interpret the criminal law principle of comparative fault as a reward granted to cautious victims.

The difference between interpreting the criminal law principle of comparative fault as a condemnation of the careless victim and interpreting it as a reward for the cautious victim is not merely an academic one. It may have important implications for criminal law theory. If indeed the criminal law principle of comparative fault is based on the culpability of the careless victim, the best way to express societal condemnation of his negligence is to treat the carelessness of the victim as a mitigating factor in sentencing. If, on the other hand, the criminal law principle of comparative fault aims at rewarding cautious victims, the best way to effectuate the principle is by treating the care exercised by the victim as an aggravating circumstance in sentencing. The former description indicates that we provide lesser protec-

voluntarily expose themselves to crimes drain more resources from the legal system than victims who do not do so. If careless potential victims have a right to equal rather than special protection, the introduction of a criminal law comparative fault doctrine is necessary to respect the rights of cautious victims to protection from the state.

131. More precisely, it is based on a more demanding moral premise—namely, that people who are equally culpable should be punished equally.
tion to careless victims; the latter indicates that we provide better protection to cautious victims.\(^{132}\)

Typically, though, victims' carelessness is not conceived of as culpable behavior; neither is their caution deemed praiseworthy. Hence, we cannot always rely on victims' perceived virtues or vices to justify use of the criminal law principle of comparative fault. Can our retributive inclinations accommodate the principle when victims' carelessness is not culpable and their caution is not praiseworthy?

The criminal law principle of comparative fault can alternatively be described as a statement that society cannot afford to protect legitimate, nonculpable behavior of victims when that behavior is too costly. Under this interpretation, the victim's culpability is neither condoned nor deplored. The principle simply reflects the fact that societal resources are limited and thus should be directed to the most urgent societal needs. The protection of careless victims is a particularly costly enterprise and consequently we may have to sacrifice some of the protection granted to careless victims. This sacrifice need not imply that the victims' careless conduct is condemnable nor that the criminals' behavior is less culpable.

Both interpretations of the criminal law principle of comparative fault seek a reconciliation between the principle and our retributive sentiments. The first interpretation does so by interpreting the principle itself as reflecting a moral judgment of victims. Under this description, the principle of comparative fault is regarded as a means of punishing culpable victims or as a means of rewarding careful victims. The second interpretation does so by acknowledging the value of retribution, while denying that the use of the principle involves any moral judgment as to the degree of evil embodied in the crime, or in the victim's carelessness. Efficiency and fairness need not therefore be incompatible with our retributive sentiments.

VI

How to Reconcile Heart and Mind

I have suggested that criminal sanctions should be analyzed as a means of providing and distributing protection to potential victims. From this new perspective, labelled the productive-distributive paradigm, it becomes apparent that efficiency and distributive justice considerations strongly support adoption of a criminal law principle of comparative fault. Moreover, the principle of comparative fault can describe various doctrines of criminal law which have so far resisted satisfactory explication.

It would be unfair to end this discussion without making a confession to the reader. It has been difficult for the author to convince himself of the justness of the criminal law principle of comparative fault. The heart

---

\(^{132}\) The examples in Part IV illustrate that in practice contemporary criminal law uses both techniques. Provocation is interpreted as a mitigating circumstance in homicide, while robbery and burglary can be regarded as theft committed in aggravating circumstance.
resisted the mind's judgments. The gradual conversion of the heart occurred only after I recalled again and again two observations which have not yet been discussed in the paper.

First, the instances which were discussed earlier as paradigmatic of the comparative fault principle, namely complex violent crimes and property offenses, seem to be normal, almost natural, features of criminal law. In thinking about a criminal law principle of comparative fault, we need not think primarily of the controversial rape cases. We should instead focus upon the numerous situations in which such a principle is implemented without an eyebrow being raised and develop selective criteria for expanding the principle in accordance with our moral principles.

Secondly, adopting the criminal law principle of comparative fault is an official recognition of the fact that the victim of a crime is an autonomous agent responsible for his actions. Criminal law is often regarded as affirming the status of the criminal as an agent. However, it has failed to accord such recognition to the victim. By directing its injunctions not only towards criminals, but also towards potential victims, criminal law recognizes and respects the potential victims as agents who would otherwise remain passive beneficiaries of criminal law norms. This suggests that in addition to the consequentialist arguments analyzed in this Article there may be some further Kantian reasons to support its conclusions, but I save those for another day.

133. See Morris, supra note 53, at 46-49.