האוניברסיטה העברית בירושלים

THE HEBREW UNIVERSITY OF JERUSALEM

THE CASE FOR DISCRIMINATORY SENTENCING: WHY IDENTICAL CRIMES MAY DESERVE DIFFERENT SANCTIONS

By

ALON HAREL and EYAL WINTER

Discussion Paper # 569 Mar 2011

מרכז לחקר הרציונליות

CENTER FOR THE STUDY OF RATIONALITY

The Case for Discriminatory Sentencing: Why Identical Crimes May Deserve Different Sanctions

Alon Harel*and Eyal Winter‡

March 3, 2011

Abstract

The traditional premise of criminal law is that criminals who are convicted of similar crimes under similar circumstances ought to be subject to identical sentences. This article provides an efficiency-based rationale for discriminatory sentencing, i.e., establishes circumstances under which identical crimes ought to be subject to differential sentencing. We also establish the relevance of this finding to the practices of sentencing and, in particular, to the Sentencing Guidelines. Most significantly, we establish that the model can explain why celebrities, leaders, or recidivists ought to be subject to harsher sanctions than others. Discriminatory sentencing is optimal when criminals confer positive externalities on each other. If a criminal A who imposes (non-reciprocal) large positive externalities on criminal B is punished sufficiently harshly, B would expect A not to commit the crime and, consequently, he would expect not to benefit from the positive externalities conferred on him by A. Given that B's expected benefits

are lower, the sanctions sufficient to deter B are also lower than the ones imposed on A. The result can be easily extended to the case of reciprocal externalities. Assume that a criminal A imposes positive externalities on B and B imposes identical positive externalities on A. If A is subject to a sufficiently harsh sanction and B knows this, B would expect A not to perform the crime and therefore would expect not to benefit from the positive externalities otherwise conferred on B. Consequently, a more lenient sanction than the sanction imposed on A would be sufficient to deter B.

1   Introduction

The traditional premise of criminal law is that criminals who are convicted of similar crimes under similar circumstances ought to be subjected to identical sentences. This view has been the basis for the endorsement of the sentencing guidelines. The sentencing guidelines were designed to promote consistency in sentencing and consistency is understood as requiring the imposition of similar sentences for similar crimes.[1] The purpose of this article is to provide an efficiency-based rationale for discriminatory sentencing, i.e., to explain when and why similar crimes committed under similar circumstances ought to be subjected to differential sentencing.

The justification for discriminatory sentencing is established in two steps. First we establish the justifiability of discriminatory sanctions in cases in

---

[1] See United States Sentencing Commission Guidelines Manual Part 1A1.1 (Nov. 2009) (noting that one reason Congress propagated the federal guidelines was to promote "reasonable uniformity in sentencing by narrowing the wide disparity in sentences imposed for criminal offenses committed by similar offenders"). In an earlier paper one of us has questioned this fundamental principle and raised the possibility that deterrence considerations may require that criminals committing similar crimes under similar circumstances face a lottery over sentences rather than a fixed sentence. See Alon Harel & Uzi Segal, "Criminal Law and Behavioral Law and Economics: Observations on the Neglected Role of Uncertainty in Deterring Crime," *American Law and Economics Review*, Oxford University Press, vol. 1 (1–2), pages 276–312 (1999). But, even under the premises of this proposal, while the actual sanctions imposed on different criminals would be different, the expected sanctions faced by criminals who committed similar acts would be identical.

which criminals confer positive non-reciprocal externalities on other criminals. Non-reciprocal externalities are defined as externalities that differ in their size such that the externality conferred by a criminal A (or a group of criminals) on another criminal B (or another group of criminals) is higher than the one conferred by criminal B on criminal A. We later apply the analysis to the more partiuclar cases of reciprocal externalities – cases in which criminals confer identical externalities on each other reciprocally.

The economic rationale justifying discriminatory sentencing in cases of non-reciprocal externalities is an intuitive one. Assume two criminals A and B and assume that both have an identical benefit from committing a crime alone. Yet if both of them commit the crime, A confers a large positive externality on B, e.g., because A is a religious leader and committing the crime reduces the stigma associated with it. In contrast B confers little or no benefit on A. If a criminal A, who imposes (non-reciprocal) positive externalities on criminal B, is expected to be punished particularly harshly, B would expect A not to commit the crime and, consequently, he would expect not to benefit from the positive externalities that can potentially be conferred by B. Given that B's expected benefits are lower, the sanctions sufficient to deter B are also lower than the ones imposed on A.

The result applies also in the particular case of reciprocal externalities. Assume that a criminal A imposes positive externalities on B and B imposes identical positive externalities on A. This often happens when the investigation of each crime imposes a burden on the police and therefore reduces the probability of detecting other criminals. If A is subjected to a sufficiently harsh sanction and B knows this, B would expect A not to perform the crime and therefore would expect not to benefit from the positive externalities otherwise conferred upon B. A more lenient sanction than the sanction imposed on A would therefore be sufficient to deter B. Part II of this paper develops a formal model establishing more rigorously the optimal sanctions in cases in which criminals confer positive (reciprocal or non-reciprocal) externalities on each other.

Discriminatory sentencing violates deep-seated moral intuitions many of us share. Yet Part III establishes that existing legal doctrine could be interpreted in a way that is consistent with the conclusions of this paper. We examine two types of positive externalities: "stigma-based externalities," namely the reduction of the stigma associated with crime as a result of the criminal activity of other criminals and "educational externalities," namely the information provided to some criminals by the performance of crimes by

other criminals. We explore some of the doctrinal intricacies of criminal law and establish that under some circumstances criminals who confer benefits on other criminals (either because they reduce the stigma associated with the crime or provide relevant information to other criminals) are indeed punished more harshly.

Translating the proposal from the language of incentives to a retributive framework would require maintaining that a person who commits a crime (and thereby confers positive externalities) commits a more wrongful act than a person who confers no such externalities. We conclude our analysis by examining how our suggestions could be translated into the retributive framework. We briefly explore this issue in Part IV of this paper.

2   A Formal Model and Results

In this section we introduce a simple model of externalities among criminals. We analyze this model and solve it for the optimal legal (sanction) scheme in two main environments starting with the case of externalities that are non-reciprocal, i.e., where one party A induces a higher externality on party B than party B induces on A. We then move to the symmetric case in which all agents are identical and show that in spite of the full symmetry the optimal scheme is still discriminatory. This intriguing result arises from the fact that we wish to screen out the bad equilibria of mis-coordination in which crime takes place.

2.1   The Model

A certain crime is potentially committed by a group N of n agents. The police and the court are jointly trying to eliminate the crime. We model the police as a non-strategic player. Specifically we assume that the police's available resources allow it to catch a criminal with a constant probability q and that the distribution is statistically independent across potential criminals. We assume that the n agents are ranked by seniority/experience, with player 1 being the most senior. Specifically, each agent j who is commiting the crime induces a positive externality $w_j > 0$ on every other agent who commits the crime. Seniority implies that $w_j > w_{j+1}$. We denote the vector of externalities by $w = (w_1, ...w_n)$. The positive externality induced by agent j can be interpreted either as an "educational externality" or as a "stigma-based

externality." *Educational externality* denotes the benefits resulting from the information and instruction given to the criminal by the criminal activity of other criminals. Such an education does not presuppose explicit instruction; it may be the by-product of a successful or unsuccessful crime performed by one which conveys information to the other. *Stigma-based externality* refers to the reduction of the negative stigma resulting from the criminal activity of others. We denote by B the material benefit that a criminal can acquire from the crime. A *legal scheme* is a vector of sanctions $s = (s_1, ... s_n)$ where s_j is the sanction inflicted on agent j if caught. The legal scheme s gives rise to a normal form game among the agents.. In this game each agent has two strategies $\{0, 1\}$, where 1 stands for committing the crime and 0 stands for refraining from it. For a vector of externalities w and a set of agents T we shall denote $w(T) = \sum_{i \in T} w_i$. For a strategy profile in which the set of agents who commit the crime is T each member j of T enjoys a total externality of $w(T\backslash\{i\})$ and the payoff to the criminal j in T is given by $B + w(T\backslash\{j\}) - qs_j$. Furthermore, this payoff is zero for each player who chooses to refain from the crime.

We shall say that $s = (s_1, ... s_n)$ is an optimal deterrence scheme if s deters all agents from committing the crime in any equilibrium of the game and furthermore there exists no other scheme s' such that full deterrence is the unique equilibrium of the game with a lower total sanction, i.e.,$\sum_{j \in N} s_j > \sum_{j \in N} s'_j$.[2]

Our model assumes that the positive externalities add up in an additive manner to the crime benefit. This assumption is not essential for our results. Our results will all hold if instead we assume that the benefits for a given criminal from committing the crime is some increasing function $B(w)$ of the total externality conferred by all the other criminals. In this case the total payoff for a criminal is given by $B(T\backslash\{i\}) - qs_i$.This payoff function applies when we think of educational externalities (that increase the benefits from the crime). Alternatively, in the context of stigma-based externalities (that reduce the negative stigma suffered by a convicted criminal) we can define a stigma function $\theta(w)$ that is decreasing as a function of the total externalities (the more frequent is the crime and the more senior are the people who commit the crime, the lower is the level of the stigma from which an additional criminal suffers. In this case the total payoff function of agent

---

[2]We assume that a criminal who is indifferent between his two options chooses to refrain.

i who commits the crime together with the set of agents $T\backslash\{i\}$ is given by $B = q(s_i + \theta(w(T\backslash\{i\})))$. As mentioned earlier our results apply to these models is the educational context as well as the stigma context.

The model we describe here involves an application of the literature in Contract Theory and in particular the part involving externalities[3] Most related to our model is Berenstein and Winter (2011)[4], which considers an abstract model of contracting with asymmetric externalities. However our results here require separate analysis as the special structure of externalities in this paper allows for more informative results regarding the structure of the optimal contracts, which cannot be derived in the abstract framework.

2.2   Non-reciprocal Externalities

We start with a two-person example and discuss the basic idea behind the structure of the optimal scheme of sanctions in the non-reciprocal environment.

Example 1

Suppose two potential criminals share the same benefits from committing the crime B and the same probability q of being caught, but differ in terms of their externalities/seniority. Suppose that the positive externality that agent 1 induces on agent 2 is 10, while agent 2 induces only 5 units of positive externality on agent 1. To be able to deter both agents from committing the crime at any possible equilibrium one of the agents must be threatened by a sanction severe enough to deter him from committing the crime even if he believes that the other agent commits the crime (from which the first player gains). This is because if both agents are deterred by a milder sanction there will be an equilibrium in which both commit the crime (while correctly believing that the other one commits the crime). We therefore have two options: impose the harsher sanction on agent 1 or impose it on agent 2. Under the first option the punishment that will deter agent 1 even if he

---

[3]See for example: Ilya Segal, "Coordination and Discrimination in Contracting with Externalities: Divide and Conquer?," 113 *Journal of Economic Theory* 147–181 (2003) and Eyal Winter, "Incentives and Discrimination," 94 *American Economic Review* 764–773 (2004).

[4]Shai Bernstein & Eyal Winter, "Contracting with Heterogeneous Externalities," The Center for the Study of Rationality, The Hebrew University of Jerusalem (2010).

believes that agent 2 commits the crime must satisfy $B + 5 - qs = 0$, or $s = \frac{B+5}{q}$. Under this option agent 1 is deterred and we can deter agent 2 by imposing a milder sanction satisfying $B - qs = 0$ or $s = \frac{B}{q}$. If we decide to impose the harsher sanction on agent 2 we will have to satisfy the equation $B + 10 - qs = 0$ or $s = \frac{B+10}{q}$. Now that agent 2 is deterred (regardless of agent 1's decision), it would be sufficient to impose a sanction of $s = \frac{B}{q}$ on agent 1. We see that the total sanction is smaller under the first option (by $\frac{5}{q}$ units). The basic intuition here is that by punishing the more senior agent more severely (such that he will be deterred no matter what the other agent does), we are eliminating a larger chunk of the total positive externality and can guarantee complete deterrence with a lower total sanction.

This intuition extends to the general case, as Proposition 1 demonstrates. Instead of only two options to design the mechanism in the two-person game, in the n-person game there are $n!$ options, which is the number of all possible orders of the n players. In every such order a sanction for a player will be just enough to deter him from committing the crime assuming all his predecessors in the order are out of the criminal circle and all of his successors are committing the crime. We shall use the term "the order of incentives" to refer to such orders and the term "the optimal order of incentives" to refer to the order that minimizes the total sanction. As we shall see, the optimal order of incentives will be unique in the non-reciprocal case (yielding higher sanctions to more senior criminals), while in the reciprocal case the order of incentives is arbitrary, still yielding discriminatory sanctions for all optimal schemes.

Proposition 1: If w_j are all distinct, then there exists a unique optimal deterrence scheme $s = (s_1, ...s_n)$ and this scheme inflicts a higher punishment on more senior criminals, i.e., $s_j > s_{j+1}$ for $1 < j < n$. Furthermore, $s_j = \frac{B+w(\{j+1,j+2,...,n\})}{q}$.

Proof: We first note that one of the agents must receive a sanction severe enough to deter him from comitting the crime even if he believes that all the other agents commit the crime. If all the agents receive a sanction that is less than this deterrence level, then there will be an equilibrium in which all agents commit the crime. Let i_1 be an agent who receives this sanction. We next note that there must be at least one agent other than i_1 who receives a sanction that is sufficient to deter him, assuming that

7

he believes that all but agent i_1 commit the crime. If all agents get sanctions less severe than that, then there will be an equilibrium in which all the agents but i_1 commit the crime. Let i_2 be a player receiving this level of sanction. We proceed by induction. Suppose that we designated agents $i_1, i_2, ...i_j$ in the above-described manner; then there must be an agent, call it i_{j+1}, whose level of sanction will be sufficient to deter him, assuming that he believes that all the agents in $N\backslash\{i_1, i_2, ...i_j\}$ commit the crime. Otherwise there will be an equilibrium in which all the agents in $N\backslash\{i_1, i_2, ...i_j\}$ commit the crime. This process yields an order $i_1, i_2, ...i_n$ and a scheme $s = (s_{i_1}, ...s_{i_n})$ satisfying $B + w(\{i_1, i_2, ...i_j\}) = qs_{i_j}$, which implies that $s_{i_j} = \frac{B+w(N\backslash\{i_1, i_2, ...i_j\})}{q}$. Furthermore s induces the game to be dominant solvable; i.e., the unique no-crime equilibrium is attained by iterative elimination of dominated strategies. We shall use the term "divide and conquer" to refer to schemes of the type described above. Showing that the optimal scheme must be divide and conquer implies that there are precisely $n!$ candidates for there to be an optimal deterrence scheme, each characterized by a different order of players. This means that the task of finding the optimal scheme boils down to finding the optimal order for which the above-described procedure has to be implemented. We shall now argue that there is a unique optimal order and this order is the identity order; i.e., 1 is first, 2 is second, ... and n is last. Suppose by way of contradiction that the optimal order is different from the identity order. Let π denote this order and let π^* denote the identity order. Since $\pi \neq \pi^*$, there must be two adjacent agents in π, $\pi(j)$ and $\pi(j + 1)$ such that $w_{\pi(j+1)} > w_{\pi(j)}$. We shall argue that the order in which the positions of these two players is swapped, denoted by π', corresponds to a deterring scheme with a smaller total sanction, which contradicts the optimality of π. We first note that for all agents $\pi(k)$ with $k < j$ we have $s_{\pi(k)} = s_{\pi'(k)}$. Furthermore, for all agents $\pi(k)$ with $k > j + 1$ we have $s_{\pi(k)} = s_{\pi'(k)}$. It is therefore enough to compare $s_{\pi(j)} + s_{\pi(j+1)}$ with $s_{\pi'(j)} + s_{\pi'(j+1)}$. The LHS equals $\frac{B+w(N\backslash\{\pi(1),...\pi(j-1)\})}{q} + \frac{B+w(N\backslash\{\pi(1),...\pi(j)\})}{q}$ and the RHS equals $\frac{B+w(N\backslash\{\pi(1),...\pi(j-1)\})}{q} + \frac{B+w(N\backslash\{\pi(1),...\pi(j-1),\pi(j+1)\})}{q}$ so the LHS - RHS equals $\frac{w_{\pi(j+1)}-w_{\pi(j)}}{q}$. But $w_{\pi(j+1)} > w_{\pi(j)}$, which means that $s_{\pi(j)} + s_{\pi(j+1)} > s_{\pi'(j)} + s_{\pi'(j+1)}$ and establishes the desired inequality. QED.

The intuition behind Proposition 1 is evident. Agents who confer higher externalities on others should be deterred "earlier" to eliminate their influence on others and thereby allow the court to inflict lower sanctions on the

rest. We should point out, however, that the reference to "early" is only virtual. There is no sequentiality in the order of moves of criminals; they all decide simultaneously, based on the sanction scheme that is assumed to be known. The word "early" refers to the agent's position in the order of incentives. "Early" agents receive a more severe punishment so that "later" agents can infer that they are out of the criminal circle and therefore later agents know that they will not benefit from the positive externalities induced by early agents.

We now turn to examine the treatment of reciprocal externalities.

2.3 Reciprocal Externalities

In our reciprocal world every two agents induce the same level of externalitiy on each other. Still, the general structure of the scheme will have to remain unchanged and some agents have to receive a higher sanction so as to be deterred regardless of the actions of others. Thus, even in the fully symmetric case the optimal scheme must be discriminatory, as Proposition 2 indicates.

Proposition 2: Assume that all agents are symmetric, i.e., $w_i = w_j$ for all i and j. Then there exist exactly $n!$ optimal deterring schemes and all these schemes are fully discriminatory, i.e., $s_i \neq s_j$ for all i and j.

Proof: We have shown in the proof of Proposition 1 that the optimal scheme must have the divide and conquer property. If all agents are symmetric, then the total sanction is the same for all divide and conquer schemes. Therefore any such scheme must be an optimal deterrence scheme. Furthermore, for each order π player $\pi(j)$ receives a sanction $s_{\pi(j)} = \frac{B + w(N \setminus \{\pi(1), ... \pi(j-1)\})}{q}$ in the corresponding mechanism. Since w_i are all positive, any two players will receive a different legal sanction. QED.

2.4 Partial Deterrence

We next consider a model in which the mechanism does not necessarily induce full deterrence such as eliminates crime entirely, but rather weights the total cost of deterrence (through sanctions) to the total damage caused by the crime. Indeed, if the cost of deterrence is negligible relative to the social damage caused by the crime, then full deterrence is optimal and we are back at our benchmark model. But if this is not the case, then it may well be

9

optimal to exempt some (or even all) agents from responsibility and impose zero punishment on these individuals. We shall assume that apart from their externalities all agents are symmetric; i.e, they all cause the same social damage D when committing the crime, they all receive the same benefit B, and they all have the same probability of being caught q.

To demonstrate the consequences of this model we shall go back and examine Example 1. As we have argued, the total sanction that deters the two agents is $\frac{2B+5}{q}$. Suppose now that we allow agent 1 to commit the crime without any threat of punishment. Agent 1 will rationally commit the crime regardless of his beliefs about the actions of other agent. In doing so he'll make the crime more attractive to agent 2. The optimal sanction deterring agent 2 will now have to satisfy $B - sq + 10 = 0$ or $s = \frac{B+10}{q}$. If instead we exempt agent 2 from punishment, then agent 2 will have a dominant strategy of committing the crime and will increase by 5 the benefits of agent 1 from committing the crime. The optimal sanction for agent 1 will therefore have to satisfy $B - sq + 5 = 0$ and $s = \frac{B+5}{q}$. It is therefore better to exempt agent 2 than it is to exempt agent 1 as agent 2's externality on agent 1 is milder than the externality that agent 1 induces on agent 2. The discussion above can be generalized now to the n-person case.

Proposition 3: Consider a court that attempt to minimize the total damage of the crime net of the deterrence cost. Assume a constant damage D for each criminal, a constant benefit B, and a constant probability of being caught q. Then the optimal scheme (with mixed externalities) has the following properties:

(1) Depending on D there will be a critical agent $j(D)$ such that all $j < j(D)$ will receive zero punishment $s_j = 0$ and all $j \geq D(j)$ will receive a positive punishment.

(2) For agents $j \geq D(j)$ the optimal sanctions are given by $s_j = \frac{B+w(\{1,2,3,...j-1\})}{q}$.

Note that assertion (1) in Proposition 3 says that as we decrease the social damage of a crime, we shall have to exempt more and more people from punishment, but those who are exempted first are always those who induce milder externalities on others (i.e., less senior criminals).

Proof of Proposition 3: Suppose by way of contradiction that there are two agents j and $j+1$ such that $s_j > 0$ and $s_{j+1} = 0$; i.e., agent j is deterred and agent $j + 1$ commits the crime. Suppose that we have the two agents

10

swap roles so that agent $j+1$ is deterred and agent j is exempted . This will not affect the total damage from the crime as the number of active criminals will remain unchanged. Since $w_j < w_{j+1}$, having the two agents swap roles will have two effects that reduce the total cost of deterrence. First, the cost of deterring $j+1$ is lower than the cost of deterring j in the original scheme, and, furthermore, the cost of deterring the other agents who need to be deterred weakly declines. This is because allowing agent j to commit the crime attracts other agents less than when agent $j+1$ is exempted.

Assertion 2 of Proposition 3 is a direct consequence of Proposition 2. Q.E.D.

One assumption made in this paper requires justification. Given that the only equilibrium resulting from the sanction proposed by us is one in which no criminal commits the offence, why should we take care to impose the minimal possible sanctions? Arguably, if no person commits a crime, no person is subjected to a criminal sanction. It follows therefore that it is costless to impose a large but equal sanction on all criminals (as long as this sanction is sufficiently large to deter all criminals). We propose two answers. First, it is false to assume that if we deter all criminals we never impose the sanction. False convictions would exist in such a system and, in such cases, we ought to minimize the size of the criminal sanctions. Second, harsh sanctions have costs even when they are never imposed as they contribute to the brutalization of society. Criminal sanctions are a mirror of the society and its humane or inhumane convictions. Imposing a sanction that is unnecessary for the purpose of deterring crime is cruel, and hence costly, even if, in reality, these sanctions are never imposed.

2.5   Legal Doctrine and Discriminatory Sentencing

Section II developed a model to analyze the precise effects of externalities conferred by one criminal on another on the optimal sanction. It also distinguished between reciprocal and non-reciprocal externalities. There are two results that are of relevance to law: 1) if criminals confer non-reciprocal externalities on each other, the optimal sanction on the person who confers a higher benefit ought to be higher. 2) If criminals confer reciprocal externalities, the optimal sanction is discriminatory. In such a case it is optimal to impose arbitrarily a harsher sanction on one criminal and a lighter sanction on the other. This section establishes that existing legal doctrines could be

rationalized in terms of efficiency. We show below that criminals who confer higher positive externalities than those that are conferred on them are (sometimes) punished more harshly. Needless to say, we do not maintain that the economic rationale is necessarily the rationale underlying the doctrines we explore here.

The legal literature has identified numerous factors influencing the propensity to commit crimes. These include the social sanctions (stigma), the social rewards resulting from a crime, the expected gains resulting from the crime, the costs of learning the criminal craft, etc.[5] The legal sanction ought to be designed so that it is sensitive to the overall effects of these incentives. These incentives are often affected by the presence/absence of other criminals performing crimes. The criminal activity of other criminals may affect the probability of detection;[6] the stigma attached to a crime;[7] the social rewards resulting from the crime;[8] it may influence the profits resulting from

---

[5]For a summary of factors influencing criminals to commit illegal activities, see Isaac Erlich, "Crime, Punishment, and the Market for Offenses," 10 *Journal of Economic Perspectives* 43, 46 (Winter 1996, No.1)

[6]Oren Bar-Gill & Alon Harel, "Crime Rates and Expected Sanctions: The Economics of Deterrence Revisited," 30 *Journal of Legal Studies* 485, 488, 489–491 (June 2001). This article argues that sometimes the greater burdens on the police (resulting from other crimes) reduce the probability of detection while at other times the fact that other crimes are committed increases the probability of detection.

[7]In some cases the more frequent the crime is, the lesser the stigma attached to it. As Bentham said: "where robberies are frequent, and unpunished, robberies are committed without shame." See Jeremy Bentham, *An Introduction to the Principles of Morals and Legislation*, ch. XII secc. I. Par. X (1789). This view was developed by Dan Khahan. See Dan Kahan, "Social Influence, Social Meaning and Deterrence," 83 *Virginia Law Review* 349, 357 (1997). For an attempt to explain why stigma is negatively correlated with the rate of crime, see Alon Harel & Alon Klement, "The Economics of Stigma: Why More Detection of Crime May Result in Less Stigmatization," 36 *Journal of Legal Studies* 355–377 (2007). In other cases the increase in the rate of crime may reinforce the social sanction. An increase in the rate of crime may raise public attention and strengthen the intensity of the societal condemnation of the crime. Thus, for instance, it was argued that "recognition of the prevalence of date rape in society was one of the factors that increased the stigma associated with it." See Bar-Gill & Harel supra note 6 at 493–494.

[8]Some crimes are rewarded by the community. For instance it was noted that so-called "honor killings" are often perceived as heroic. See, e.g., Valerie Plant, "Honor Killings and the Asylum Gender Gap," 15 *Journal of Transnational Law and Policy* 109, 117 (2005). Similarly, war crimes may be perceived as heroic by the society or the nation of the perpetrators of such crimes. Yet, the more people are engaged in such crimes, the less heroic these crimes may seem and the fewer the resulting social rewards conferred upon their perpetrators.

the crime;[9] or it may provide useful information to other criminals. In accordance with the analysis in Section II we divide the discussion into two parts: non-reciprocal externalities and reciprocal ones. Most of the discussion in this section is devoted to non-reciprocal externalities of two main types: stigma-related externalities and educational externalities.

2.6   Non-reciprocal Externalities

To examine the nature of non-reciprocal externalities let us distinguish between senior and regular criminals. The former have greater positive impact on the well-being of other criminals while the latter have little or no positive impact on the well-being of other criminals. There are two types of positive externalities that are characteristically non-reciprocal: stigma effects and "educational externalities."

2.6.1   Non-reciprocal Stigma Effects: High-esteemed/Low-esteemed Criminals

It has been noted that the effect of a crime on the social stigma attached to the crime depends on the identity of the person who commits the crime, in particular on her social status.[10] Criminals who have high status in their communities, e.g., spiritual or religious leaders or simply established members of the community, may have much greater effects on the stigma attached to a particular crime than individuals who do not enjoy such status. Consequently, such high esteemed criminals reduce the costs of committing the crime and confer great benefits on other criminals. The most paradigmatic cases are ones where the very involvement in the crime is designed to reduce the stigma attached to the crime and challenge the legal prohibitions. The willingness of a respectable and charismatic priest to conduct marriages of underage children was presumably designed to convey his conviction that

---

[9]For instance a recent research maintains that "a large volume of stolen guns competes with guns sold by traffickers and depresses black market prices, reducing both the profit incentive for traffickers and the need for their services." See Gary Kleck & Shun-Yang Kevin Wang, "The Myth of Big-time Trafficking and the Overinterpretation of Gun Tracing Data," 56 *UCLA Law Review* 1233, 1282 (2009).

[10]For some empirical support for this hypothesis, see Brian Mullen, Carolyn Copper & James E. Driskell, "Jaywalking as a Function of Model Behavior," 16 *Personality and Social Psychology Bulletin* 320–330 (1990).

such marriages are legitimate. The harsh charges against the priest (including charging him with being an accomplice to rape) were presumably based on the concern that others would follow him and conduct such marriages due to the effects of his acts on the resulting stigma.[11] At other times, while the crime is not designed to reduce the negative stigma attached to the crime, it does so inadvertently simply because of the social status or eminence of the criminal. As my leader/priest/teacher/mentor has been committing the crime, it follows that the crime cannot be that bad.[12]

Who is esteemed/despised within a particular community may differ of course from one community to the next. It is not necessarily the case that individuals with high socio-economic status are esteemed in a community. Yet, it seems likely that in our society the socio-economic status of an individual is significant in evaluating the potential stigma effects that the individual's involvement in a crime have. We can assume that the higher the socio-economic status of the individual is, the greater the impact she has on the

---

[11] We here refer to the prosecution, conviction, and now recent reversal of Fundamentalist Latter Day Saints leader Warren Jeffs, for his role in the propagation and affirmation of underage polygamous marriages in his isolated religious community. See John Dougherty, "Polygamist Leader Sentenced to Ten Years in Prison," New York Times (Nov. 21, 2008); Dan Frosch, "Polygamist's Rape Convictions are Overturned in Utah," New York Times (July 28, 2010). Celebrities have been known to receive harsher punishments than the general populace because of their special status and influence. A very current anecdote supporting this conjecture involves American actress Lindsay Lohan, frequently in the media for her wild antics. Lohan recently violated the terms of a court order by testing positive for drug use. Though the underlying crime was a misdemeanor, the lower court judge sent her to jail and denied her bail until a hearing on the merits scheduled a month later. See Richard Winton & Andrew Blankstein, "Lindsay Lohan Judge Found Way around Early Jail Release for Actor," Los Angeles Times (Sept. 24, 2010), http://articles.latimes.com/2010/sep/24/local/la-me-0925-lohan-early-m

[12] The issue has been explicitly raised with respect to celebrities, in partiuclar athletes. Thus when high school basketplayer Dajuan Wagner was sentenced for assault, the judge noted that: "to the young people of New Jersey and Camden he's a star" and imposed a harsh sentence. See Michael M. O'Hear, Symposuim: Blue-Collar Crimes/White-Collar Criminals: Sentencing Elite Athletes Who Commit Violent Crimes 12 Marq. Sports L. Rev. 427, 438 (2001). Other cases also indicate that courts impose indeed particularly harsh sanctions on athletes. Another famous example are the harsh sanctions imposed on Kansas City Royals basketball players. See, Laurie Nicole Robinson, Comment: Professional Atheletes-Held to to a Higher Standard and Above the Law: A Comment on High-Profile Criminal Defendants and the Need for States to Establish High-Profile Courts 73 Ind L.J. 1313, 1327 (1998). See also Jared Chamberlain et al, Celebrities in the Courtroom: Legal Responses, Psychological Theory and Empirical Research 3 Vand.J.Ent & Tech. L. 551, 556-558 (2006).

stigma attached to the crime. High-status criminals may affect the stigma attached to the crime for a different reason. To the extent that crimes committed by high-status individuals (in particular celebrities) are more likely to be noticed by the media, such crimes become public knowledge and they therefore generate the feeling that the relevant legal prohibitions are not fully observed. The mere fact that crimes committed by eminent individuals are more likely to become public knowledge implies that they are more likely to erode compliance. It would be interesting therefore to explore to what extent the socio-economic status of the criminal affects the sentence imposed upon the crime.

The Sentencing Guidelines prohibit discrimination on the basis of race, sex, national origin, creed, and socio-economic status of offenders.[13] In specifying the duties of the Sentencing Commission, 28 U.S.C. Section 994(d) proclaims that: "The Commission shall assure that the guidelines and policy statements are entirely neutral as to the race, sex, national origin, creed, and socio-economic status of offenders." The Sentencing Guidelines indeed prohibit discrimination on the basis of socio-economic background.

Yet a more thorough investigation raises doubts whether the bar on discriminatory sentencing is indeed realized. There are many reasons why socio-economic status may be relevant to sentencing despite the prohibition in the Sentencing Guidelines. For our purposes we wish only to examine the conjecture that so-called "identical sentences" required by the Sentencing Guidelines may have differential effects on the welfare of high-esteemed and low-esteemed individuals.[14]

---

[13] It does not however prohibit discrimination on the grounds of celebrity status. Some have suggested that courts may take into account the "moral leadership" of the defendant into account. See United States v. Gunderson, 211 F. 3d 1088, 1089 (8th Cir. 2000).

[14] Surprisingly, even Immanuel Kant who is not known for his great concern for utilitarian considerations was aware of the differential effects of criminal sentencing on the welfare of criminals belonging to different classes. Kant believed that upon committing a serious crime and being given a choice between death and convicted labor:

"The man of honor would choose death and the scoundrel convicted labor. This comes with the nature of the human mind; for the man of honor is acquainted with something he values even more highly than life, namely honor, while the scoundrel considers it better to live in shame than not to live at all. Since the man of honor is undeniably less deserving of punishment than the other, both would be punished quite proportionately if all alike were sentenced to death. The man of honor would be punished mildly in terms of his sensibilities and the scoundrel severely in terms of his."

See Immanuel Kant, The Metaphysics of Morals (ed, Mary Gregor) 6:334 (1996).

It has been noticed that the sentences imposed on high-esteemed criminals are harsher (in terms of welfare) than those imposed on low-esteemed criminals. The Sentencing Guidelines and the courts' consistent decisions expressly prohibit differential sanctions based on socio-economic considerations. Yet, "identical sentences" (in terms of the time of incarceration) impose greater costs on high-esteemed criminals than on low-esteemed criminals. The effects of incarceration on individuals of high socio-economic background are much harsher than the effects of incarceration on individuals of low socio-economic background.[15] Some legal scholars, most notably Adam Kolber, have argued passionately on this ground that if criminal sanctions are retributivist, then it follows that high-esteemed individuals should get shorter prison terms. In their view, imposing identical prison terms on the poor and the rich discriminates against the rich.[16]

Kolber may be right in his conjecture that the rich indeed suffer more from incarceration than the poor and that therefore justice requires that the rich get more lenient sentences than the poor. But, to the extent that crimes committed by high-esteemed individuals weaken the stigma attached to the crime, it may be justifiable to impose harsher sanctions on such individuals. The Sentencing Guidelines bar on discrimination between high-esteemed and low-esteemed individuals (in terms of the length of prison terms) may ironically amount to imposing harsher sanctions (in utility terms) on the former

---

[15]With respect to fines it is evident that the same fine may have differential effects on poor and rich criminals. It was suggested already by Bentham that "pecuniary punishments should always be regulated by the fortune of the offender." See Jeremy Bentham, *The Theory of Litigation* 217. See also Richard Posner, "An Optimal Theory for White Collar Criminals," 17 *American Criminal Law Review* 409, 415 (1980). Many legal systems therefore distinguish between poor and rich criminals and impose differential fines. See, e.g., Derek A. Westen, "Fines, Imprisonment and the Poor: Thirty Dollars or Thirty Days," 57 *California Law Review* 778, 813 (1967). With respect to imprisonment, opposite claims have been made. Richard Posner believes that "the disutility of imprisonment rises with income." See Posner ibid. at 415. See also Adam J. Kolber, "The Comparative Nature of Punishment," 89 *Boston University Law Review* 1565, 1567–1568 (2009); Adam J. Kolber, "The Subjective Experience of Punishment," 109 *Columbia Law Review* 182, 235 (2009). For a powerful critique of the relevance of the subjective experiences of criminals in prisons on the appropiate punishment, see Kenneth Simmons, "Retributivists Need Not and Should Not Endorse the Subjectivist Account of Punishment," *Columbia Law Review Sidebar* (March 12, 2009). For a more comprehensive attack on the subjectivist school of punishment, see Dan Merkel, "Bentham on Stilts: The Bare Relevance of Subjectivity to Retributive Justice," 98 *California Law Review* 907–988 (2010).

[16]See Kolber, supra The Subjective Experience of Punishment, supra note 15.

than on the latter and thus address the positive externalities that the former provide to the latter.

2.6.2 Non-reciprocal Educational Benefits

The more experienced a criminal is, the greater the benefits his activity confers on other criminals. An experienced criminal therefore confers benefits on other criminals. The simplest way that this is done is when an experienced criminal looks for a non-experienced partner. The non-experienced partner's benefits are not merely the direct monetary benefits resulting from the crime but also from the benefits resulting from the experience of the more senior criminal. The more experienced criminals are also more likely to develop new techniques to commit crimes and to avoid being detected. To the extent that these techniques are adopted by other less-experienced criminals, the experienced criminals provide non-reciprocal benefits to the non-experienced criminals. Such benefits need not presuppose any communication between the criminals; techniques can be disseminated and criminal experience accumulated without direct instruction.

The best proxy for experience is the degree of recidivism of the criminal. Recidivism is considered to be a standard consideration in sentencing.[17] The practice of imposing harsher sanctions on recidivists is entrenched in the legal system.[18] One of the interesting implications of this view is that harsher

---

[17]See the Guidelines Manual, supra note 1 at 5A1.1 for the sentencing table. Previous criminal history is strongly related to the severity of punishment. See also commentary to 4A1.1 ("A defendant with a record of prior criminal behavior is more culpable than a first offender and thus deserving of greater punishment. General deterrence of criminal conduct dictates that a clear message be sent to society that repeated criminal behavior will aggravate the need for punishment with each recurrence.") Upward modifications are also available for defendants who were an "organizer, leader, manager or supervisor in any criminal activity" (3B1.1), abuse of position of trust or use of a special skill (3B1.1)

[18]Interestingly, law and economics theorists have questioned the practice of imposing harsher sanctions on recidivists. To the extent that we wish criminals to internalize the costs they impose on others, there is no reason to impose harsher sanctions on recidivists. Recidivists and non-recidivists deserve the same sanction. For some attempts to rationalize the application of escalating penalties from an economics perspective, see C. Y. Cyrus Chu, Sheng-cheng Hu & Ting-yuan Huang, "Punishing Repeat Offenders More Severely," *International Review of Law & Economics* 20, 127, 134–136 (2000) (cited in note 92) (arguing that penalties should be lesser for first-time offenders and greater for repeat offenders because of the increased risk of erroneously convicting first-time offenders); Ehud Guttel & Alon Harel, "Probability Matching: The Behavioral Law and Economics

sanctions ought to be imposed on repeat offenders who commit similar crimes rather than those who commit different crimes as only the former are likely to produce educational benefits.

The Sentencing Guidelines also dictate that criminals who possess special skills and use the skills either in the commission of the offense or in the concealment of the offense are potentially subject to harsher sanctions.

The "special skills" provision has raised a lot of controversy.[19] To properly

---

of Repeated Behavior," 72 *University of Chicago Law Review* 1126–1129 (2005) (using behavioral heuristics to justify the imposition of harsher sanctions on recidivists); A. Mitchell Polinsky & Steven Shavell, "On Offense History and the Theory of Deterrence," 18 *International Review of Law & Economics* 305 (1998) (arguing that escalating penalties may be optimal insofar as they deter both initial and repeat offenders); A. Mitchell Polinsky & Daniel L. Rubinfeld, "A Model of Optimal Fines for Repeat Offenders," 46 *Journal of Public Economics* 291, 303 (1991) (arguing that escalating penalties may be efficient if the probability of detection is lower for repeat offenders because of their increased experience in committing crimes). But see Moshe Burnovski & Zvi Safra, "Deterrence Effects of Sequential Punishment Policies: Should Repeat Offenders Be More Severely Punished?," 14 *International Review of Law and Economics* 341 (1994) (arguing that if total punishment is held constant, then increasing the punishment for an initial violation and decreasing the punishment for subsequent violations increases total deterrence). None of these attempts, however, provides a satisfactory explanation. See Winand Emons, "A Note on the Optimal Punishment for Repeat Offenders," 23 *International Review of Law and Economics,* 254–255 (2003).

[19]A broad swath of special skills have been considered for upward modification: most involve the criminal's possession of important information and skill sets relevant to the commission of the crime. For instance, accountants accused of tax fraud have been found to be in possession of a special skill. See, e.g., U.S. v. Rice, 52 F.3d 843 (N.M. 1995). Use of technical knowledge regarding computers may be special skills; see, e.g., U.S. v. Peterson, 98 F.3d 502 (Cal. 1996), but see U.S. v. Godman, 223 F.3d. 320 (Tenn. 2000) (holding it is not). Ability to drive an 18-wheel big-rig has been consistently held to be a special skill. See, e.g., U.S. v. Smith, 332 F.3d 455 (Ill. 2003) and U.S. v. Tolar, 268 F.3d 530 (Ill. 2001). The knowledge required to manufacture methamphetamine, LSD, or other drugs is a special skill for sentencing purposes. U.S. v. Spencer, 4. F.3d 115 (Vt. 1993); U.S. v. Pickard, 298 F.Supp.2d 1140 (D.Kans. 2003). Attorneys forging judicial orders have been found to employ special skills. U.S. v. Reich, 479 F.3d 179 (N.Y. 2007). Using knowledge of electronic skills for placing a PIN-recording device inside an ATM has been deemed a special skill. U.S. v. Lavin, 27.F.3d 1994 (N.Y. 1994). However, special knowledge regarding the design and functioning of ATMs was held not to be a special skill for sentencing purposes, although this knowledge was employed by a criminal to steal from the ATM. See U.S. v. Harper 33 F.3d. 1143 (Cal. 1994). But c.f. U.S. v. Aubin, 961 F.2d.980 (Mass. 1992), holding that such knowledge is a special skill. Knowledge of automobiles and car parts has been deemed a special skill for purposes of sentencing professional car thieves. See U.S. v. Exarhos, 135 F.3d 723 (Fla. 1998). Captaining a

evaluate the "special skills" provision and rationalize it we need to uncover its underlying logic. The most natural economic rationale for the "special skills" provision may be to increase the sanctions on criminals who are less likely to be detected or more likely to succeed due to their skills. As criminals with "special skills" are more likely to succeed in committing the crime and avoid detection, it seems rational to impose harsher sanctions on them. Yet our paper suggests an additional rationale, namely, the potential contribution of a criminal with special skills to his peers and the positive externalities she confers on them. In evaluating whether a qualification of a criminal ought to be classified as a "special skill" under this provision, one ought to examine the potential of the criminal to transmit this qualification to others, i.e. confer benefits upon them.

Upward modifications are also available for defendants who serve some form of leadership role in the commission of crimes. For criminal activity involving more than five people, defendants who served as "leaders or organizers" of the activity can be subject to four levels of upward modification. Defendants who serve as "managers or supervisors" but not "leaders or organizers" are subject to three levels of upward modification.

The leadership provision was often used by courts to increase the sentences for criminals who impart knowledge to their co-conspirators. A defendant who created fraudulent checks and provided them to other conspirators with advice and instruction on how to use them properly was given an upward adjustment.[20] The Seventh Circuit has ruled that a defendant who provided instruction to his co-conspirators so that they could better conceal the nature of the stolen goods in order to "fence" them on E-Bay was prop-

---

vessel on the high seas requires use of a special skill. See U.S. v. Calderon, 127 F.3d 1314 (Fla. 1997), but only if it involves the use of a special skill set not available to the general public. The ability to manufacture and reproduce keys is a special skill. See U.S. v. Muzingo, 999 F.2d. 391 (Iowa 1993). Language is not considered a special skill, at least by some jurisdictions. See U.S. v. Elefant, 999 F.2d. 674 (N.Y. 1993). Skill as an architect used to build an illegal smuggling tunnel has been considered a special skill. See U.S. v. Corona-Verbera, 509 F.3d. 1105. Airplane piloting (U.S. v. Mettler, 938 F.2d 764 (Wis. 1991)), radio operation (U.S. v. Malzoga, 2 F.3d 1107 (Fla. 1993)), tax-preparation (U.S. v. Noah, 130 F.3d 490, (R.I. 1997)), securities trading ability, (U.S. v. Keiser, 578 F.3d 897 (N.D. 2009)) are further examples of special skills.

[20] U.S. v. M.A. Yah, 500 F.3d 698 (8th Cir. 2007) (co-defendant testified that Yah "taught him how to 'scam'") (affirming court's discretion for the upward modification of sentence, but overturning conviction on other grounds).

erly given an upward adjustment for a leadership role.[21] An organizer of an insurance fraud scheme involving staged car accidents properly received an upward adjustment, in part because he instructed his co-conspirators as to what they should say to insurance companies when reporting the accident.[22] A defendant in a credit card fraud scheme was held to be in a leadership position as he had taught his co-defendant exactly how to commit the fraud.[23] Courts have found leadership roles on the basis of teaching, either in whole or in part, in cases regarding methamphetamine production.[24] Lastly, courts used this provision in cases regarding the manufacturer of counterfeit financial instruments.[25]

Our conclusion is therefore that the Sentencing Guidelines often increase the sanctions for criminals who provide "educational benefits" to other criminals. This result can be rationalized on the grounds of the positive externalities conferred by these criminals. We turn now to examine the treatment of reciprocal externalities.

2.7 Reciprocal Externalities

There are numerous reciprocal externalities that criminals confer or impose on each other. Such externalities include the effects of crimes on the probability of detection, the stigma effects that may (at times) be reciprocal, the mutual effects of crimes on the profitability of the crime (competition among criminals), and mutual learning. We also showed in Section II that efficiency requires the imposition of differential sanctions even in cases of reciprocal externalities. Can legal doctrine be interpreted to accomplish such a result?

---

[21]U.S. v. Watz, 450 F.3d 720 (7th Cir. 2006) (explaining that because the defendant had directed his co-conspirators to steal specific items, and had taught co-conspirators how to create fake invoices to hide the source of these items, the defendant was not merely a colleague, but played a leadership role).

[22]U.S. v. Colello, 16 F.3d 193 (7th Cir 1993).

[23]U.S. v. Abdiodun, 442 F.Supp. 2d 88 (S.D. NY 2006) ("In other words, he had agreed to teach [his co-defendant] how to commit the fraud and, accordingly, was a supervisor or manager and should be sentenced accordingly") (conviction later overturned on other grounds, though upward modification for leadership affirmed in U.S. v. Abdiodun, 536 F. 3d 162 (2d cir. 2008)).

[24]U.S. v. Scott, 287 Fedd. Appx. 801 (11th Cir. 2008); U.S. v. West, 232 Fed. Appx. 333 (4th Cir. 2007); U.S. v. Grey, 109 Fed. Appx. 74 (6th Cir 2004)).

[25]U.S. v. Williams, 88 Fed. Appx. 194 (9th Cir. 2004) (cash); U.S. v. McDaniels, 398 F.3d 540 (6th Cir. 2005) (altered checks).

It is hardly surprising that there are no explicit cases in which the criminal law system or the Sentencing Guidelines discriminate among criminals who commit similar crimes in similar circumstances. The imposition of differential sanctions in cases involving non-reciprocal externalities seems intelligible at times, e.g., the differential treatment of recidivists and non-recidivists. Even when such distinctions do not seem plausible (as in cases involving criminals of different social classes), we have shown that the concept of equality is sufficiently flexible such that in effect high-status criminals are subject to harsher sanctions than low-status criminals (due to the differential welfare effects of incarceration on these two groups). In contrast, differentiating between criminals in the case of reciprocal externalities seems to require designating "arbitrarily" some individuals and imposing greater or lesser sanctions on them. This practice seems to violate foundational retributive sentiments.

We believe that such a conclusion is too hasty. Admittedly, the legal system is reluctant to use "arbitrary" designation. Yet what is needed in cases of reciprocal externalities is not "arbitrary" designation but any designation whatsoever. Prior designation of some criminals for harsher sanctions will achieve the purpose of promoting efficiency even if the classificatory methods used in sentencing are not "arbitrary" but are grounded in retributive considerations (such as the distinctions drawn between recidivists and non-recidivists). In cases of reciprocal externalities, any classificatory criteria designating some for harsher sanctions and others for more lenient sanctions could bring about efficient outcomes. The fact that retributive considerations support the distinction between recidivists and non-recidivists makes the discriminatory sentencing scheme politically feasible, but it does not undermine its efficiency as a tool to address reciprocal externalities.

2.8   Retributivism and Disctiminatory Sentencing

The analysis in this paper may strike the criminal lawyer as counterintuitive and as deviating too radically from the retributivist intuitions underlying criminal law. More specifically, it seems to offend the strong retributivist sentiments underlying criminal law. Unde the traditional view, the size of the criminal sanction ought to reflect the culpability and the wrongfulness of the perpetrator of the crime.[26] If two individuals who performed a criminal

---

[26]For the wrongfulness-culpability paradigm, see George P. Fletcher, *Rethinking Criminal Law*, 454–459 (1978); Robert Nozick, *Philosophical Explanations*, 363–397 (1981).

act are equally culpable and the two acts are equally wrongful they ought to be subject to equal sanctions. Hence it follows that discriminatory sentencing conflicts with retributive concerns.

This may be true; this paper is not designed to support the impositon of discriminatory sentencing but only to explore the circumtances under which it is beneficial from an efficiency-based perspective. Furthermore, we believe that at least with respect to the case of non-reciprocal externalities our model could be made consistent with the wrongfulness-culpability paradigm. The wrongfulness-culpability paradigm is a formal framework and its content depends on identifying what parameters affect the culpability of the perpetrator and the wrongfulness of the act. In evaluating the wrongfulness of an act, the benefits conferred by performing this act on other criminals could be taken into account. To the extent that a criminal act by one person confers great benefits on another criminal while the crime of the second criminal confers little or no benefits on the first criminal, the act of the first criminal could be judged to be more wrongful than the act of the second criminal. If this is so, discriminatory sentencing could be made compatible with traditional retributive principles.