

Optimal Imprisonment with General Enforcement of Law

Massimo D'Antoni*

University of Siena

Tim Friehe

University of Marburg

Avraham Tabbach

Tel-Aviv University

October 15, 2022

Abstract

We study the optimal use of imprisonment when enforcement efforts are general (i.e., when effort increases the detection probability for a range of acts). In contrast to the conventional wisdom that optimal imprisonment rises with the act's harmfulness and is equal to the maximum level only for the most harmful acts, we show that – when the distribution of criminal benefits exhibits a standard monotone hazard rate – optimal imprisonment can only be zero or maximal. Thus, having general as opposed to specific enforcement effort does not alter the fact that only extreme sanction levels should optimally be employed.

Keywords: Imprisonment; General Enforcement; Maximal Sanctions

JEL classification: D31; D62; H23; K14; K42

*Corresponding author

1. Introduction

Becker (1968) famously showed that it is optimal to deter crime by combining the lowest possible probability (to economize on enforcement costs) with the severest possible penalty (to maintain adequate deterrence). This proposition remains the basic tenet in the optimal law enforcement literature, although several papers have identified specific circumstances in which this combination of instrument levels is not optimal (Garoupa, 1997).

In an important contribution, Shavell (1991) argues that acknowledging the relevance of *general* instead of *specific* law enforcement can align prescriptions from theory with the real-world observation of non-maximal sanctions. According to Shavell (1991), specific enforcement concerns apprehending and penalizing individuals who have committed a single kind of criminal act as identified by its level of social harm (e.g., enforcers who write tickets for overtime parking). In contrast, general enforcement enables apprehending and penalizing individuals who have committed any of a range of acts (e.g., a police officer on patrol may assist in the conviction of perpetrators of minor or major crimes).¹ Shavell (1991) reports that the maximum imprisonment term is socially optimal when enforcement is specific. For each act in isolation, the Becker proposition applies, that is, any expected sanction should comprise the severest sanction possible in order to economize on the costs of enforcement. In contrast, he finds that, if enforcement efforts are general, optimal imprisonment rises with the act's harmfulness and is equal to the maximum term only for the most harmful acts. The intuition is that the higher level of harm increases the marginal benefit of a longer imprisonment term without influencing the marginal costs. These results have become a corner stone in the analysis of optimal law enforcement.²

This paper challenges this conventional wisdom. We show that optimal imprisonment is either zero or maximal when the the distribution of criminal benefits exhibits a monotone hazard rate (MHR) that is not only usually assumed to hold in many fields of applied microeconomics (including law and economics) but also fulfilled for a wide range of actual distribution functions.³ The intuition for the extreme-imprisonment result is that welfare evolves in a u-shaped manner with the imprisonment term when MHR applies, signifying that the interior solution *minimizes* welfare.

¹Mookherjee and Png (1992) similarly distinguish between monitoring (efforts not tailored to the severity of the act) and investigation (efforts undertaken when the severity of the act is known) in their study focused on monetary sanctions.

²See, for example, Garoupa (1997, pp. 272-73) and Polinsky and Shavell (2007, pp. 431-32).

³The most commonly used parametric distributions – such as the normal and the beta and gamma distributions with bell-shaped density – exhibit MHR on their support. The same holds for a distribution on a finite support such as the uniform distribution or triangular distribution with any mode.

With the MHR assumption, the optimal pattern of enforcement and imprisonment is as follows: (i) no enforcement and any imprisonment term at all harms level, or (ii) positive enforcement and no (maximum) imprisonment for acts that cause harm below (above) a critical harm level. As a result, both specific and general enforcement of law are associated with the use of only extreme sanctions.

Whereas the result by [Shavell \(1991\)](#) that imprisonment increases in the harmfulness of the act remains conventional wisdom to this day, [Kaplow \(1990\)](#) already mentioned that interior solutions may require possibly strong assumptions. Our paper adds to [Kaplow \(1990\)](#), who analyzes only a single act (i.e., specific enforcement), by identifying the precise characteristics of the distribution functions that induce boundary solutions, by highlighting that these characteristics are taken for granted in many strands of the literature, and by describing the optimal regime when enforcement is general.⁴

2. Model

We employ the model used by [Shavell \(1991\)](#) in which risk-neutral individuals choose whether to commit harmful acts by comparing the expected sanction with the criminal benefit. Individuals differ regarding this benefit b and the act's harmfulness denoted h . The benefit is distributed on the support $B \subseteq [0, \infty)$ according to the cumulative distribution function $F(b)$ and density $F'(b) = f(b)$. We also refer to the hazard rate function

$$H(b) = \frac{f(b)}{1 - F(b)},$$

and assume that it is non-decreasing. This is the widely used monotone hazard rate assumption (see, for example, [Tirole, 1994](#), p. 156). The harm is distributed on the support $[0, \infty)$ according to $G(h)$ as the cumulative distribution function and $g(h)$ as density. We assume that the distribution of benefits is the same for different harm levels.

If an individual commits an act, she will be detected with probability p and receive an imprisonment term z . We focus on general enforcement, so p cannot depend on h . Let y denote the costs of enforcement and assume that $p'(y) > 0$ and $p''(y) \leq 0$. The imprisonment term can depend on the level of harm (i.e., the kind of act) so we write

⁴In [D'Antoni et al. \(2022\)](#), we describe how the unobservability of wealth changes the optimal combination of fines and imprisonment when enforcement is specific. In the present contribution, wealth levels are irrelevant as we focus on imprisonment as the only kind of sanction in order to contribute to the understanding of optimal imprisonment when enforcement is general (i.e., when there is a range of acts to consider)

$z(h)$. Imprisonment cannot exceed an upper limit \bar{z} , which may reflect physical limits or moral constraints and is the same for all individuals.

Imprisonment is costly for the convicted offender and for society. The cost per-unit amount to σ and comprise the individual's costs and the additional social costs, $\sigma > 1$.

3. Analysis

An individual will commit an act if and only if her benefit exceeds the expected sanction, that is, if $b > p(y)z(h)$. Understanding this decision criterion, the social planner chooses enforcement efforts y and imprisonment $z(h)$ to maximize social welfare W defined to be the benefits individuals obtain from committing acts, less the harm done, less the social costs of imposing imprisonment, less the cost of enforcement. Formally,

$$W = \int_0^\infty \int_{p(y)z(h)}^\infty (b - h - p(y)z(h)\sigma)f(b)dbg(h)dh - y. \quad (1)$$

When enforcement efforts are positive, the marginal welfare effect from a longer imprisonment term results as

$$\frac{\partial W}{\partial z} = p(h + pz(\sigma - 1))f(pz) - p\sigma(1 - F(pz)), \quad (2)$$

where the first term reflects the marginal benefit from deterring the marginal offender, and the second term reflects the marginal cost from increasing imprisonment on those who will continue to offend even at the higher sanction level. A higher imprisonment term increases the term $h + pz(\sigma - 1)$ in the marginal benefits and lowers the marginal costs of imprisonment as it reduces the population of undeterred individuals.

[Shavell \(1991\)](#) considers an interior solution for the level of imprisonment (see equation (A17) in [Shavell, 1991](#), p. 1105). Assuming that the marginal welfare effect in (2) is equal to zero, we can also state the necessary requirement as

$$H(pz) - \frac{\sigma}{h + pz(\sigma - 1)} = 0, \quad (3)$$

using the hazard rate H . When we derive this term with respect to z for the second-order condition, we obtain (after dividing by p)

$$H'(pz) + \frac{(\sigma - 1)\sigma}{(h + p(\sigma - 1)z)^2}, \quad (4)$$

which is *positive* as long as the hazard rate function is non-decreasing (i.e., $H'(b) \geq 0$). A non-decreasing monotone hazard rate signifies that the ratio between the marginal probability of deterrence (captured by the criminal benefit density function $f(b)$) to the crime rate (captured by the survival function $1 - F(b)$) is non-decreasing. A very intuitive understanding results from noticing that the hazard rate is equal to the absolute value of the elasticity of the crime rate with respect to the expected sanction \mathcal{S} divided by the expected sanction \mathcal{S} , that is,

$$H(\mathcal{S}) = \left| \frac{d(1 - F(\mathcal{S}))}{d\mathcal{S}} \frac{1}{1 - F(\mathcal{S})} \right|. \quad (5)$$

Assuming $H'(b) \geq 0$, our result rules out that an interior solution represents a global maximum, and points to boundary solutions. After establishing that imprisonment is either zero or maximum for any given level of harm under the MHR assumption, we now analyze the general pattern of imprisonment.

Proposition 1. *Suppose that general enforcement is employed, that imprisonment is used as the only kind of sanction, and that the distribution of benefits fulfills the monotone hazard rate condition. Then, optimal law enforcement features either (i) zero enforcement effort and any imprisonment term or (ii) positive enforcement and maximum (no) imprisonment for acts imposing $h \geq h^*$ ($h < h^*$).*

Proof. The fact that only extreme sanctions will be employed has been shown above. Next, define $\hat{y}(h_c)$ as the solution of

$$\int_{h_c}^{\infty} p'(\hat{y})\bar{z} [(h + p(\hat{y})\bar{z}(\sigma - 1))f(p(\hat{y})\bar{z}) - (1 - F(p(\hat{y})\bar{z}))\sigma] g(h)dh = 1, \quad (6)$$

where the sanction is maximal if $h \geq h_c$ and zero otherwise. Note that the marginal effect for a given level of harm is similar to the one for imprisonment except for the marginal enforcement cost term but that, for example, with p' diminishing sufficiently fast, an interior solution can be ensured.⁵ With the integrand positive at all levels of $h \geq h_c$, it holds that $\hat{y}' < 0$ and that $\hat{y} = 0$ for h_c sufficiently large. Using this definition of the enforcement effort in our definition of welfare

$$W(h_c) = \int_0^{h_c} \int_0^{\infty} (b - h)f(b)dbg(h)dh + \int_{h_c}^{\infty} \int_{p(\hat{y})\bar{z}}^{\infty} (b - h - p(\hat{y})\bar{z}\sigma)f(b)dbg(h)dh - \hat{y}, \quad (7)$$

⁵See [Kaplow \(1989\)](#) for a discussion.

we find that welfare changes with the value of h_c as follows

$$W'(h_c) = g(h_c) \left(\int_0^{p(\hat{y})\bar{z}} (b - h_c) f(b) db + \int_{p(\hat{y})\bar{z}}^{\infty} p(\hat{y})\bar{z}\sigma f(b) db \right), \quad (8)$$

where the indirect effects via the enforcement effort \hat{y} cancelled out as a result of the envelope theorem. Scenario (i) obtains when this marginal effect is positive until h_c is so high that $\hat{y} = 0$. Scenario (ii) results when the two marginal effects balance out at a level of $\hat{y} > 0$, which requires that acts with $b \in (0, p(\hat{y})\bar{z}]$ create sufficiently high net harm. At $h_c = 0$, the marginal effect is clearly positive, such that a scenario in which all acts are punished by the maximum sanction (i.e., where $h^* = 0$) is not possible. \square

4. Discussion and Conclusion

According to the conventional wisdom, optimal imprisonment increases with the harmfulness of the act when enforcement effort is general, whereas this is not true when enforcement effort is specific. For the case of general enforcement, this paper shows that optimal imprisonment is either zero or maximal when the distribution of benefits fulfills a widely used and intuitive characteristic that is also fulfilled by many regularly used distribution functions. Our result implies that general as opposed to specific enforcement cannot generally improve the alignment of theoretically prescribed and practically observed sanction regimes.

Although our conclusions do not contradict widely accepted results, they put them into perspective by pointing out that their validity relies on hypotheses that are not usually made explicit and could be easily violated. As clarified by our analysis, reliance on internal solutions requires a discussion of the implied shape of the hazard rate of the criminal benefits.

Under some circumstances, a decreasing hazard rate leading to an interior solution may be justified. One possibility is a fat-tailed distribution of benefits, for example because its upper tail follows a Pareto distribution.⁶ The Pareto distribution is used in finance, to deal with catastrophic events, and in the analysis of income and wealth distributions, as it is consistent with empirical data for that context.⁷ It is not clear, however, whether it could be a realistic representation of the distribution of criminal benefits.

⁶This is the assumption made by [Mungan \(2017\)](#) in a model of optimal sanctioning in which social welfare does not include criminals' benefits.

⁷Notably, the assumption of a Pareto distribution of income plays a role in reconciling the optimal taxation analysis of the income tax with the circumstance that the marginal tax rate is not declining at high levels of income ([Diamond, 1998](#)).

A second possibility is that individuals may commit a crime by mistake. If there's a probability that individuals commit a harmful act even when they do not want to (and, ex post, it is not possible to discriminate intentional and unintentional acts), then the frequency of violation will stop declining as sanctions increase even before full deterrence is reached. This amounts to a declining hazard rate.

Finally, some individuals may not engage in a rational cost-benefit comparison and thus be unresponsive to incentives provided by sanctions. This implies that some individuals remain undeterred at all levels of punishment. This kind of "irrational" behavior by some individuals may be realistic in some cases of criminal behavior (e.g., drug addicts), although it deviates from the standard economic framework. In these instances, an interior solution to imprisonment in the general enforcement model is possible.

References

- Becker, G.S., 1968. Crime and punishment: an economic approach. *Journal of Political Economy* 76, 169–217.
- Diamond, P.A., 1998. Optimal income taxation: an example with a U-shaped pattern of optimal marginal tax rates. *American Economic Review* 88, 83–95.
- Garoupa, N., 1997. The theory of optimal law enforcement. *Journal of Economic Surveys* 11, 267–95.
- Kaplow, L., 1989. A note on the optimal use of nonmonetary sanctions. *Harvard Program in Law and Economics Discussion Paper* 61.
- Kaplow, L., 1990. A note on the optimal use of nonmonetary sanctions. *Journal of Public Economics* 42 (2), 245–247.
- Mookherjee, D., Png, I., 1992. Monitoring vis-a-vis investigation in enforcement of law. *American Economic Review* 82 (3), 556–65.
- Mungan, M.C., 2017. Over-incarceration and disenfranchisement. *Public Choice* 172 (3–4), 377–395.
- Polinsky, A.M., Shavell, S., 2007. The theory of public enforcement of law. *Handbook of Law and Economics* 1, 403–54.
- Shavell, S., 1991. Specific versus general enforcement of law. *Journal of Political Economy* 99 (5), 1088–1108.
- Tirole, J., 1994. *The Theory of Industrial Organization*, MIT Press, Cambridge, Mass.