



University of Connecticut

Department of Economics Working Paper Series

Deterred or Detained? A Unified Model of Criminal Punishment

Thomas J. Miceli
University of Connecticut

Working Paper 2009-16

June 2009

341 Mansfield Road, Unit 1063
Storrs, CT 06269-1063
Phone: (860) 486-3022
Fax: (860) 486-4463
<http://www.econ.uconn.edu/>

This working paper is indexed on RePEc, <http://repec.org/>

Abstract

The standard economic model of crime since Becker (1968) is primarily concerned with deterrence. Actual punishment policies, however, appear to rely on imprisonment to a greater extent than is prescribed by that model. One reason may be the incapacitation function of prison. The model developed in this paper seeks to incorporate incapacitation into the standard model. A key finding of the hybrid model is that when prison is the only form of punishment and the probability of apprehension is fixed, incapacitation can result in a longer or a shorter optimal prison term compared to the deterrence-only model. It is longer if there is underdeterrence, and shorter if there is overdeterrence. When fines are also available and are not constrained by offenders' wealth, the optimal prison term is zero. Since the fine achieves first-best deterrence, only efficient crimes are committed, and hence, there is no gain from incapacitation. Other aspects of the standard model are also studied within the context of the hybrid model.

Journal of Economic Literature Classification: K14, K42

Keywords: Deterrence, imprisonment, incapacitation, law enforcement

I appreciate the comments of Steve Shavell on an earlier draft.

Deterred or Detained?

A Unified Model of Criminal Punishment

1. Introduction: The Compatibility of Deterrence and Incapacitation

The standard economic model of crime as first developed by Becker (1968), and brought to its current level of maturity by Polinsky and Shavell (2000, 2007), is primarily concerned with deterrence. Central to the BPS model (as I shall call it), and crucial for deterrence to be possible, is the *rational offender assumption*, which maintains that would-be criminals decide whether or not to commit a crime by comparing the gains from the illegal act to the expected punishment. Although some may doubt the validity of this assumption, there is ample empirical evidence to support it.¹

On the basis of the rational offender assumption, the enforcement authority is able to choose the optimal law enforcement policy, consisting of the probability of apprehension and severity of punishment (the fine and/or prison term), so as to maximize a social welfare function that consists of the net harm from crime plus the cost of enforcement. At the optimum, some amount of crime is tolerated, both because enforcement is costly, and because the benefits to offenders are generally counted as part of social welfare, which means that some crimes are “efficient”.² In other words, optimal deterrence in the BPS model does not generally entail complete deterrence.

One of the clearest implications of this model as regards the optimal structure of punishment is that fines should be used to the maximal extent possible (i.e., up to the offender’s wealth) before prison is even contemplated. The obvious reason is that, while

¹ For a survey of the empirical literature, see Levitt and Miles (2007).

² This assumption is by no means uncontroversial. See the discussion in Lewin and Trumbull (1990). The implications of relaxing this assumption are examined in Section 5 below.

finer and prison are equally capable of deterring crime, fines are costless to impose while prison is costly (Polinsky and Shavell, 1984). The extensive use of prison in actual penalty schemes therefore appears to represent a departure of legal practice from the prescriptions of the BPS model.

A different rationale for criminal punishment that is perhaps more consistent with the actual use of prison is incapacitation, which involves the removal of dangerous offenders from society. For example, this is undoubtedly the principal motivation behind the recent spate of “three-strikes” laws, which prescribe escalating penalties, culminating with life imprisonment, for repeat offenders (Shepherd, 2002). In a sense, though, one can think of incapacitation and deterrence as complementary rationales for punishment. That is, when the threat of punishment is not expected to prevent certain offenders from committing undesirable or dangerous acts, for whatever reason, incapacitation may be necessary to prevent them from inflicting such harm on society. In this context, Shavell (1987) has shown that it is optimal (cost minimizing) to imprison those offenders who would impose more harm on society than the cost of detaining them. The incapacitation model is silent, however, about why the offenders were not deterred in the first place by the threatened imprisonment. At the same time, the BPS model generally ignores the possibility that some offenders are undeterrable; indeed there is no reason to punish these offenders at all from a deterrence perspective.

The current state of the literature therefore involves a somewhat troubling disconnect between the two theories of law enforcement, which naturally raises the question of whether there is a way to reconcile the deterrence and incapacitation motives

within the context of the BPS model.³ A first step in this direction is to make the BPS model dynamic so as to allow for the time dimension inherent in the incapacitation motive. Several recent efforts have been made along these lines by way of investigating the pervasiveness of escalating penalty schemes,⁴ but none has extended the analysis to incapacitation. These studies have nevertheless revealed an important insight that sheds light on the compatibility of deterrence and incapacitation. Specifically, they have shown that escalating penalties can never be optimal in a pure deterrence model if deterrence is perfect (Polinsky and Shavell, 2000, p. 67). The reason is that, when deterrence is perfect, only efficient crimes are committed, so it would not be socially advantageous to deter offenders from committing them repeatedly. By the same logic, incapacitation of offenders who are expected to commit efficient crimes would be undesirable. Based on this reasoning, it would seem that incapacitation can only be justified in a rational offender model if there is some underdeterrence at the optimum; that is, if there is a social gain from imprisoning undeterred offenders in order to prevent them from committing further, *inefficient* crimes.

As it happens, inadequate deterrence is a possible outcome of the BPS model when enforcement is costly. For example, when prison is used as one form of punishment (the only case that is relevant for incapacitation), social welfare is maximized at the point where the marginal benefit of deterring an additional crime equals the marginal cost of raising the expected sanction. Polinsky and Shavell (2000, pp. 50-51) show that this may entail either overdeterrence or underdeterrence. The model developed in the next section exploits this fact to develop a fully rational, dynamic model of law

³ See Ehrlich (1981) for an early effort at this.

⁴ See, for example, Polinsky and Rubinfeld (1991), Polinsky and Shavell (1998), Emons (2004), and Miceli and Bucci (2005).

enforcement in which both deterrence and incapacitation can arise naturally as motivations for imprisoning convicted offenders (assuming that prison is desirable).

The analysis is organized as follows. Section 2 lays out the basic model. Section 3 then considers the simplest case, where the probability of apprehension is fixed and prison is the only form of punishment. In that setting, we show that incorporating the possibility of incapacitation can either result in a longer or a shorter prison term compared to the pure deterrence model. Intuitively, if there is underdeterrence at the optimum, it is desirable to increase the prison term for incapacitation purposes so as to prevent convicted offenders from committing further inefficient crimes. However, if there is overdeterrence, then the optimal prison term should actually be shortened so as to permit offenders to resume committing efficient crimes.

When the model is extended to allow the use of fines along with prison, we show that if the fine is not limited by the offender's wealth, then, as in the BPS model, it is never optimal to use prison. The reason here is that, because the fine can achieve the optimal (first-best) level of deterrence, only efficient crimes are committed, so there is no social gain from incapacitating offenders. However, if the fine is constrained by the offender's wealth, prison may be optimal based on the same considerations as in the prison-only model.

Section 4 further extends the model to allow the probability of apprehension to be endogenous. The conclusions of this version of the model essentially coincide with those of the BPS model. Specifically, when prison is the only form of punishment, it should be maximal, but when fines are also available, the fine should again be used to the maximal

extent possible (limited only by the offender's wealth) before prison is considered. In this case, it is no longer generally optimal to make the prison term maximal.

Section 5 briefly considers a model in which the offender's utility is not counted as part of social welfare. This case is of interest because crimes for which incapacitation is appropriate tend to be violent and hence are not likely to be of any social value. In the prison-only model with a fixed probability of apprehension (the only case we consider), we show that the optimum involves either a finite prison term (if the harm from crimes is less than the cost of imprisonment) or an infinite term (if the harm from crimes is greater than the cost of imprisonment). Finally, Section 6 summarizes the results of the model and offers concluding comments.

2. The Model

The dynamic law enforcement model to be developed in this section is based on a formulation first proposed by Davis (1988). The model involves fully rational offenders who have infinite life spans and potentially commit crimes throughout their lives when not imprisoned. As in the BPS model, offenders decide whether or not to become criminals by comparing the gain from illegal activity to the expected punishment. The gain, g , is determined when, at time zero, a potential offender takes a random draw from the density function $z(g)$. An offender's realized g then defines his "type" throughout the remainder of his life; that is, whenever he encounters a future criminal opportunity.⁵ Time is continuous, and so g is defined to be the gain per instant of time that the offender is free and committing crimes. The instantaneous gain from not committing crimes is normalized to be zero.

⁵ The results would not be affected if he took another random draw at each new opportunity.

The offender makes his initial crime decision immediately upon realizing g (i.e., at time zero). If he chooses to become a criminal, he enjoys the gain up until the time he is apprehended, which is a random variable denoted by t . The density function for t is given by

$$u(t) = pe^{-pt}, \quad (1)$$

where p is the instantaneous probability of apprehension.⁶ The corresponding distribution function for t is thus $U(t)=e^{-pt}$, and the expected time until apprehension is $1/p$.

Since we are interested in incapacitation, we assume punishment involves imprisonment, possibly combined with a fine. Thus, at the time of apprehension, the offender pays the fine (if any) and is imprisoned for a length of time s . Then, when he is released at time $t+s$, he immediately confronts another criminal opportunity, and the process begins again. We consider a time-invariant enforcement policy.⁷ Thus, given the offender's type, g , he will make the same decision whenever confronted with a criminal opportunity. In other words, offenders who find it optimal to commit a crime at time zero will also find it optimal to do so whenever they are free, while those who are deterred at time zero will never commit crimes. Figure 1 depicts the time line of events as just described.

[Figure 1 here]

⁶ Loury (1979) and Mortensen (1982) use a similar approach to model the discovery of technological innovations.

⁷ Thus, for simplicity, we ignore enforcement policies that condition punishment on an offender's history. See the references in note 4 *supra*.

3. Fixed Probability of Apprehension

Following Polinsky and Shavell (2000, 2007), we initially consider optimal enforcement for the case where the probability of apprehension, p , is fixed. We first consider punishment by prison only, and then prison combined with a fine. In Section 4 we turn to the case where p is also endogenous.

3.1. Prison only

We begin by deriving the expected lifetime gain for an individual who chooses to become a criminal at time zero. As noted, the offender enjoys a gross gain of g at each instant until he is apprehended at time t . He is then imprisoned until time $t+s$, at which time he is released and confronts another criminal opportunity. Letting δ represent the disutility of prison to the offender per unit of time, we compute the following net gain for his initial episode of crime and punishment as a function of g , s , and t :

$$\int_0^t g e^{-r\tau} d\tau - \int_t^{t+s} \delta e^{-r\tau} d\tau = \frac{g}{r} (1 - e^{-rt}) - \frac{\delta}{r} (e^{-rt} - e^{-r(t+s)}), \quad (2)$$

where r is the discount rate. Since the apprehension date is a random variable, we need to compute the expected value of this expression by weighting it by the density function in (1) and integrating over all t :

$$\begin{aligned} G(g, s, p) &= \int_0^\infty \left[\frac{g}{r} (1 - e^{-rt}) - \frac{\delta}{r} (e^{-rt} - e^{-r(t+s)}) \right] p e^{-pt} dt \\ &= \frac{1}{p+r} \left[g - \frac{p\delta}{r} (1 - e^{-rs}) \right]. \end{aligned} \quad (3)$$

Expression (3) represents the net expected gain to the offender for the initial episode of crime and punishment. Given the assumption that the offender continues to

behave in this way over an infinite lifetime (assuming that this gain is positive, as discussed below), we compute his expected lifetime gain by the recursive equation

$$\Gamma(g, s, p) = G(g, s, p) + \beta(s)\Gamma(g, s, p), \quad (4)$$

where the expected discount factor, $\beta(s)$, is calculated as follows:

$$\beta(s) = \int_0^{\infty} e^{-r(t+s)} p e^{-pt} dt = \frac{p e^{-rs}}{p + r}. \quad (5)$$

Substituting (5) into (4) and solving for $\Gamma(g, s, p)$ yields the expression for the offender's expected net gain from a lifetime of crime:

$$\Gamma(g, s, p) = \frac{1}{p(1 - e^{-rs}) + r} \left[g - \frac{p\delta}{r} (1 - e^{-rs}) \right]. \quad (6)$$

Given that the return from legal activity is normalized to zero, a potential offender will commit a crime at time zero, and at all subsequent points in time that he is free, if this expression is positive. Thus, those individuals who draw a value of g such that

$$g \geq \frac{p\delta}{r} (1 - e^{-rs}) \equiv \hat{g}(s) \quad (7)$$

will become career criminals, while those who draw a lower g never commit crimes.

Note that $\partial \hat{g}(s) / \partial s = p\delta e^{-rs} > 0$, which shows that increasing the prison term will deter some offenders from becoming criminals. This reflects the deterrence function of prison.

Now consider the social cost of crime. This consists of the harm to victims plus the costs of enforcement. The harm is captured by h , which is defined to be the harm suffered by victims per unit of time that the offender is free (i.e., from time zero up to time t for the initial offense). The cost of enforcement consists of two components, the fixed cost of apprehension, denoted $k(p)$, and the variable cost of punishment, which consists of the cost to society of imprisoning the offender once he is apprehended. This

latter cost is reflected by c , the per-unit cost society incurs whenever the offender is in prison (e.g., from time t up to time $t+s$ for the initial prison term). The harm plus imprisonment cost for the offender's first episode of crime, written as a function of the initial apprehension date, t , is thus given by

$$\int_0^t h e^{-r\tau} d\tau + \int_t^{t+s} c e^{-r\tau} d\tau = \frac{h}{r} (1 - e^{-rt}) + \frac{c}{r} (e^{-rt} - e^{-r(t+s)}). \quad (8)$$

Computing the expected value of this expression as above, we obtain

$$\begin{aligned} C(s,p) &= \int_0^\infty \left[\frac{h}{r} (1 - e^{-rt}) - \frac{c}{r} (e^{-rt} - e^{-r(t+s)}) \right] p e^{-pt} dt \\ &= \frac{1}{p+r} \left[h + \frac{pc}{r} (1 - e^{-rs}) \right]. \end{aligned} \quad (9)$$

Finally, we convert this to the present value of expected costs over the offender's infinite lifetime as follows:

$$\Psi(s,p) = C(s,p) + \beta(s)\Psi(s,p). \quad (10)$$

Substituting for $\beta(s)$ from (5) and solving yields the present value of harm plus variable enforcement costs:

$$\Psi(s,p) = \frac{1}{p(1 - e^{-rs}) + r} \left[h + \frac{pc}{r} (1 - e^{-rs}) \right]. \quad (11)$$

Adding the fixed apprehension costs to this expression yields the total expected social costs of crime, $\Psi(s,p) + k(p)$.

Before examining the general model, we demonstrate, by way of comparison, that both the pure deterrence and pure incapacitation models emerge from the above formulation as special cases.

3.1.1. Pure Deterrence Model. The pure deterrence model emerges by assuming that the offender's crime decision is one-time. That is, the would-be offender decides whether or not to commit a crime at time zero, and then refrains from committing any future crimes, whether or not he is caught and punished. Thus, imprisonment offers no incapacitation benefits.

In this context, only the initial episode of crime and punishment is relevant for any given offender. Thus, the net benefit of crime to the offender is given by the expression in (3), while the social cost of crime is given by the expression in (9). Social welfare consists of the benefit minus the cost, summed over those offenders who choose to commit the crime, or those with $g \geq \hat{g}(s)$ (where $\hat{g}(s)$ is define by (7)), minus fixed enforcement costs. The resulting expression is

$$\begin{aligned} W_d &= \int_{\hat{g}(s)}^{\infty} [G(g, p, s) - C(p, s)] z(g) dg - k(p) \\ &= \int_{\hat{g}(s)}^{\infty} \frac{1}{p+r} \left[g - h - \frac{p(c+\delta)}{r} (1 - e^{-rs}) \right] z(g) dg - k(p). \end{aligned} \quad (12)$$

The optimal prison term is found by maximizing (12) with respect to s . The resulting first order condition, assuming an interior solution, is

$$\left[h + \frac{pc}{r} (1 - e^{-rs}) \right] z(\hat{g}) \delta = \int_{\hat{g}}^{\infty} (c + \delta) z(g) dg. \quad (13)$$

The left-hand side of (13) represents the marginal deterrence benefits of a longer prison term, as reflected by the savings in harm to victims plus the present value of savings in punishment costs. The right-hand side is the marginal cost of increasing s , consisting of the cost of a longer prison term in the form of higher punishment costs to society and the

offender. This condition is essentially identical to the condition for the optimal prison term in the standard BPS model.⁸

3.1.2. Pure Incapacitation Model. We next show that the optimal incapacitation policy as derived by Shavell (1987) emerges from the above formulation by choosing the prison term to minimize the present value of harm plus imprisonment costs, holding the crime rate fixed. In this case, the infinite-time version of the model is relevant. Thus, taking the derivative of the cost expression in (11) with respect to s yields

$$\frac{\partial \Psi}{\partial s} = \frac{pre^{-rs}(c-h)}{[p(1-e^{-rs})+r]^2}, \quad (14)$$

which is ambiguous in sign. It is positive if $c > h$ and negative if $c < h$. In the former case, the cost of imprisonment exceeds the harm imposed by criminals, so it is efficient to release dangerous offenders immediately upon apprehension rather than to hold them in prison. However, when the cost of imprisonment is less than the harm from crime, it is efficient to keep dangerous offenders in prison for life rather than to release them. The optimum thus entails a corner solution, given that both c and h are constant. Shavell extends the model, however, to allow the possibility that an offender's danger to society can change over time (for example, it may decline with the offender's age), thus possibly yielding a finite prison term.⁹

3.1.3. Combining Deterrence and Incapacitation: The Hybrid Model. We now turn to the general, or hybrid, model, which incorporates both deterrence and incapacitation. Social welfare in this model consists of the present value of gains to

⁸ See, for example, condition (6) in Polinsky and Shavell (2007, p. 410). The only difference is that here, the punishment costs on the left-hand side are in present value terms. (Also, Polinsky and Shavell assume $p=1$ in the fixed-probability version of their model.)

⁹ The rehabilitative effects of imprisonment may also reduce the dangerousness of an offender over the course of his incarceration, thus leading to eventual release.

offenders in (6), minus the present value of social costs in (11), summed over the range of actual offenders, minus fixed enforcement costs. The resulting expression is

$$\begin{aligned}
W_h &= \int_{\hat{g}(s)}^{\infty} [\Gamma(g, s, p) - \Psi(s, p)] z(g) dg - k(p) \\
&= \int_{\hat{g}(s)}^{\infty} \frac{1}{p(1 - e^{-rs}) + r} \left[g - h - \frac{p(c + \delta)}{r} (1 - e^{-rs}) \right] z(g) dg - k(p). \quad (15)
\end{aligned}$$

The optimal prison term is found by maximizing W_h with respect to s . The first-order condition (assuming an interior solution) is given by

$$\left[h + \frac{pc}{r} (1 - e^{-rs}) \right] z(\hat{g}) \delta = \frac{1}{p(1 - e^{-rs}) + r} \int_{\hat{g}}^{\infty} (c + \delta + g - h) z(g) dg. \quad (16)$$

The left-hand side of (16) is identical to the left-hand side of (12), and again represents the marginal deterrence benefits of a longer prison term. The marginal cost of increasing s on right-hand side of (16), however, differs from the right-hand side of (13). It consists of the usual cost of a longer prison term in the form of higher punishment costs to society and the offender, captured by the $c + \delta$ term in the integral, but there is an additional component, captured by the $g - h$ term, which reflects the incapacitation effects of imprisonment. Specifically, $g - h$ represents the foregone net social benefits (the opportunity costs) of those crimes that the offender is not able to commit because he is kept in prison longer, conditional on the fact that, given $g \geq \hat{g}$, he would find it optimal to continue committing crimes if set free. If $g - h$ is negative at the optimum, this term lowers the marginal cost of prison, which will tend to increase the optimal prison term compared to a pure deterrence model. However, if $g - h$ is positive at the optimum, this term raises the marginal cost of prison, which will tend to reduce the optimal prison term compared to the pure deterrence model.

The intuitive explanation for these results is as follows. If prison prevents offenders from committing *inefficient* crimes (those for which $g < h$), then incapacitation is socially desirable, and the optimal prison term will be longer to reflect this benefit. But if prison prevents offenders from committing *efficient* crimes (those for which $g > h$), then incapacitation is socially undesirable and will actually work in the direction of reducing the optimal prison term. It is not possible to tell from (16) whether, for those crimes that are committed, $g-h$ is positive or negative at the optimum. All we can conclude is that the right-hand side as a whole must be positive (given that the left-hand side is positive). It follows that, at the optimum, it must be true that

$$c + \delta > h - E[g \mid g \geq \hat{g}(s)], \quad (17)$$

or the marginal cost of imprisonment, $c + \delta$, must be greater than the net social harm imposed by those crimes that are actually committed. This condition necessarily holds if crimes are on efficient on average (i.e., if the right-hand side of (17) is negative), but it can also hold if crimes are inefficient on average (i.e., if the right-hand side of (17) is positive). Which of these is true will depend on the nature of the distribution of offender gains. Ultimately, therefore, we can only conclude that, in the prison-only model with p fixed, *incapacitation may either raise or lower the optimal prison term compared to the pure deterrence model.*

3.2. Prison and fines

We now add the possibility that offenders can be subject to a fine as well as prison if caught. Obviously, fines can have no incapacitation effect, but they can deter offenders, thus possibly leaving prison to be used solely for purposes of incapacitation. Let f be the amount of the fine, which is imposed as a lump-sum amount at the time of

apprehension. (For repeat offenders, it will be imposed for each offense.) As noted in the introduction, it is well-known that in the pure deterrence (BPS) model, it is never optimal to use prison unless an offender's wealth is insufficient to cover the optimally deterring fine (and even then, prison may not be optimal, depending on the marginal cost and benefit). To see whether this is also true in the hybrid model, we assume initially that offenders are not wealth-constrained.

The first step in incorporating a fine into the above model is to amend the offender's expected gain. For the initial episode of crime and punishment, we first subtract the present value of the fine, $e^{-rt}f$, from the expression in (2). Then, proceeding as above, we obtain the present value of the expected lifetime gain for the offender (the analog to (6)):

$$\Gamma(g, f, s, p) = \frac{1}{p(1 - e^{-rs}) + r} \left[g - pf - \frac{p\delta}{r}(1 - e^{-rs}) \right]. \quad (18)$$

As before, the offender chooses to become a career criminal if the term in square brackets is positive, or if

$$g \geq pf + \frac{p\delta}{r}(1 - e^{-rs}) \equiv \hat{g}(f, s), \quad (19)$$

which differs from (7) only by the addition of the expected fine on the right-hand side of the inequality.

The expression for the social cost of crime in (10) also needs to be amended to reflect the expected revenue from the fine. Again, proceeding as above, we obtain

$$C(f, s, p) = \frac{1}{p(1 - e^{-rs}) + r} \left[h + \frac{pc}{r}(1 - e^{-rs}) - pf \right] \quad (20)$$

as the present value of the net costs of crime to society. Finally, combining the offender's gain with the social costs for those crimes that are actually committed, we obtain the expression for social welfare:

$$W_f = \int_{\hat{g}(f,s)}^{\infty} \frac{1}{p(1 - e^{-rs}) + r} \left[g - h - \frac{p(c + \delta)}{r} (1 - e^{-rs}) \right] z(g) dg - k(p). \quad (21)$$

Note that the fine revenue drops out because it is just a transfer payment. As a result, this expression for welfare is identical to that in (15) except for the threshold value of g , which now depends on the magnitude of the fine as shown in (19).

Consider first the optimal fine, which is found by setting the derivative of (21) with respect to f equal to zero. The result is

$$f^* = h/p + \frac{c}{r} (1 - e^{-rs}). \quad (22)$$

The optimal fine thus equals the harm suffered by victims, plus the present value of imprisonment costs, both discounted to reflect the uncertainty of apprehension. Note that in the case where the imprisonment term is zero, the optimal fine reduces to the usual expression, $f^* = h/p$. If deterrence were the only enforcement objective and offenders were not wealth-constrained, this would represent the first-best solution (in which case $s^* = 0$.) The question is whether, when incapacitation is also a goal, it is optimal to combine the fine with a positive prison term.

To answer this question, we maximize welfare with respect to s , given $f = f^*$. Thus, taking the derivative of (21) with respect to s and evaluating it at $f = f^*$ yields

$$\frac{\partial W_f}{\partial s} \Big|_{f=f^*} = \frac{pre^{-rs}}{[p(1 - e^{-rs}) + r]^2} \int_{\hat{g}(f^*,s)}^{\infty} (-g - \delta - c + h) z(g) dg, \quad (23)$$

where

$$\hat{g}(f^*, s) = h + \frac{p(c + \delta)}{r} (1 - e^{-rs}). \quad (24)$$

Now set $s=0$, in which case (23) becomes

$$\frac{\partial W_f}{\partial s} \Big|_{f=f^*, s=0} = \frac{p}{r} \int_h^\infty (-g - \delta - c + h) z(g) dg. \quad (25)$$

Note that this expression must be negative given that the integration is over the range where $g \geq h$. It follows that $s^*=0$, or the optimal prison term is zero. Intuitively, since the fine is unconstrained, it can be set to achieve the first-best level of deterrence, which means that only efficient crimes (i.e., only those for which $g \geq h$) are ever committed.¹⁰ Thus, there is no reason to imprison offenders, either for purposes of deterrence or incapacitation. Note that this conclusion is consistent with the results in the previous section, where it was shown that the gain from incapacitation (if any) was due to the existence of underdeterrence in the prison-only model. It follows that *when perfect deterrence can be achieved by means of a fine, there is no social gain from incapacitation*. Thus, the hybrid model yields that same conclusion as the BPS model; namely, that when the fine is not constrained by the offender's wealth, it is never optimal to use prison.

The conclusion may be different if the fine is limited by the offender's wealth (i.e., if $w < h/p$), for in that case, perfect deterrence cannot be achieved by a fine alone. When this is true, the situation is essentially the same as in the prison-only case. That is, there may be some underdeterrence at the optimum, and given $f=w$, prison may now be desirable for purposes of both deterrence and incapacitation. The specific direction of the

¹⁰ Note, in particular, that when $s=0$, $\hat{g}(f^*, 0) = h$.

incapacitation effect on the optimal prison term is determined by the same factors described above.

4. Endogenous Probability of Apprehension

In this section, we extend the model to the case where the probability of apprehension is a choice variable along with the severity of punishment. Thus, the cost of apprehension is no longer fixed but is a convex function of p ; that is, $k' > 0$, $k'' \geq 0$.

First, in the prison-only version of the hybrid model, we can show that, as in the pure deterrence (BPS) model, when p is endogenous, the optimal prison term is maximal. The proof proceeds as usual. First, suppose that s is less than maximal. Now raise s and lower p so as to hold $p(1 - e^{-rs})$ constant. Since the integral term in (15) remains constant and enforcement costs fall, welfare must rise, implying that the less-than-maximal prison term could not have been optimal. In the context of the hybrid model, a maximal prison term would, in theory, be infinite, meaning that an offender would be imprisoned for life on the first offense.¹¹ In that case, the welfare function in (15) reduces to

$$W_h = \int_{\hat{g}(s)}^{\infty} \frac{1}{p+r} \left[g - h - \frac{p(c+\delta)}{r} \right] z(g) dg - k(p), \quad (26)$$

where, from (7), $\lim_{s \rightarrow \infty} \hat{g}(s) = p\delta/r$.¹² The optimal apprehension rate in this case is found by maximizing (26) with respect to p . The resulting first order condition (again, assuming an interior solution) is

$$\frac{1}{p+r} \left(h + \frac{pc}{r} \right) z(\hat{g}) \frac{\delta}{r} = \frac{1}{(p+r)^2} \int_{\hat{g}}^{\infty} (c + \delta + g - h) z(g) dg + k'(p). \quad (27)$$

¹¹ Considerations of fairness, however, might dictate a finite maximal sentence. See Miceli (1991).

¹² Thus an infinite prison term will not generally result in complete deterrence.

The left-hand side of this condition is the marginal deterrence benefit of increasing p , as reflected by the present value of prevented harm plus saved punishment costs, while the right-hand side is the marginal cost of increasing p , including marginal punishment plus enforcement costs. Note that the first term on the right-hand side, the marginal punishment cost, has the same interpretation as the right-hand side of (16). That is, the marginal cost of increasing the certainty of punishment includes the social and private cost of imprisonment, $c+\delta$, but also the foregone gains (or opportunity cost) of prison, $g-h$, which may be positive or negative at the optimum.

Consider next the case of fines and prison when p is endogenous. As shown in the model above where p was fixed, it is never optimal to use prison if the fine is unconstrained by the offender's wealth. Thus, the only case of interest is when the fine is constrained. In that case, the probability of apprehension and the prison term are chosen simultaneously to maximize the welfare function in (21). However, it is no longer necessarily true that the optimal prison term is maximal. The reason is that if s is raised and p lowered to keep $p(1-e^{-rs})$ constant, deterrence is reduced because the expected fine, pw , falls (see (19)). Thus, welfare is not necessarily increased by continually raising s and lowering p . The qualitative results here are identical to those from the standard BPS deterrence model as described in Polinsky and Shavell (2000, pp. 55-56).

5. Offender's Utility does not Count in Social Welfare

To this point, we have maintained the standard assumption that the offender's utility should be counted as part of social welfare. While this may make sense for crimes like speeding or double parking, it is less appealing for violent crimes. In addition,

offenders who commit violent crimes are the ones that society would most likely want to incapacitate. Thus, it is of interest to extend the model to the case where offenders' utility is not included in social welfare. For simplicity, we limit the analysis to the prison-only, fixed- p case. (Hence, in what follows we can ignore the fixed cost of apprehension.)

Social welfare in this case consists of the harm caused by criminals plus punishment costs, summed over offenses.¹³ The optimal prison term is therefore chosen to minimize this cost. The relevant cost expression is given by (11), summed over offenses, or

$$SC = \int_{\hat{g}(s)}^{\infty} \Psi(s, p) z(g) dg = \int_{\hat{g}(s)}^{\infty} \frac{1}{p(1 - e^{-rs}) + r} \left[h + \frac{pc}{r} (1 - e^{-rs}) \right] z(g) dg, \quad (28)$$

where $\hat{g}(s)$ continues to be defined by (7). Taking the derivative of this expression with respect to s and rearranging yields

$$\frac{\partial SC}{\partial s} = \frac{-rpe^{-rs}}{p(1 - e^{-rs}) + r} \left\{ \left[h + \frac{pc}{r} (1 - e^{-rs}) \right] z(\hat{g}) \frac{\delta}{r} + \frac{1}{p(1 - e^{-rs}) + r} \int_{\hat{g}(s)}^{\infty} (h - c) z(g) dg \right\}. \quad (29)$$

Note that the first term in braces is identical to the left-hand side of (27), and again represents the marginal deterrence benefit of a longer prison term. The second term, however, is different from the right-hand side of (27). Note in particular that it may be positive or negative, depending on the magnitudes of h and c . This term therefore reflects the social desirability of incapacitation.

Suppose first that $h < c$, or that the harm caused by offenders is less than the cost of imprisoning them. Then, the second term in braces in (29) is negative. In that case,

¹³ To be consistent, we do not include the offender's disutility of imprisonment as part of social costs.

the optimal prison term will generally be finite, and will occur at the point where the marginal deterrence benefit of a longer prison term equals the net marginal cost of imprisonment (i.e., the point where (29) equals zero). Even though prison costs more than the harm offenders impose when free—i.e., even though imprisonment is not desirable from a strict incapacitation perspective—it is still socially beneficial to imprison offenders for some amount of time because of its deterrence benefits. In this case, deterrence and incapacitation work in opposing directions.

Alternatively, suppose that $h > c$, or that the harm imposed by offenders exceeds the cost of imprisoning them. The second term in braces in (29) is now positive, meaning that the entire expression is negative. As a result, social costs are minimized by an infinite (maximal) prison term. In this case, deterrence and incapacitation work in the same direction and indicate that the prison term should be as long as possible. Note that there is no such thing as “overdeterrence” in this model because all crimes are assumed to be inefficient. Thus, it is desirable to deter as many crimes as possible by the threat of prison, and, given that complete deterrence is not possible, to imprison those who do commit them for life (or as long as legally permitted) on the first offense.

This last outcome seems most descriptive of the motivation underlying three-strikes laws, which tend to be aimed at offenders who repeatedly commit undesirable acts. Of course, the current model does not account for the gradual progression to a maximal penalty under these laws. Thus, extending the current framework to allow for escalating penalties is a potential direction for future work.

7. Conclusion

This paper represents the first attempt to incorporate the motives of deterrence and incapacitation into a unified economic model of law enforcement. This is desirable from the perspective of a positive economic theory of law since there seems little doubt that both motives operate in the design of actual criminal policy. The effort presents certain methodological difficulties, however, because the standard economic model of law enforcement relies on the assumption that offenders are rational, whereas incapacitation implicitly assumes that offenders are undeterrable, and so can only be prevented from committing crimes by imprisonment. The model developed here gets over this conceptual hurdle by considering an explicitly dynamic setting where fully rational offenders make criminal decisions throughout their infinite lifetimes. In this setting, we showed that deterrence and incapacitation naturally emerge as complementary enforcement methods (assuming punishment by prison is socially desirable). Specifically, some offenders are deterred by the threat of punishment, while those who are not deterred are (possibly) incapacitated to prevent them from committing further inefficient acts.

The specific conclusions from this model are summarized as follows. (1) When prison alone is used as punishment and the probability of detection is fixed, incapacitation may raise or lower the optimal prison term compared to the pure deterrence model. Incapacitation raises the prison term when there is underdeterrence at the optimum, but lowers it when there is overdeterrence at the optimum. (2) When fines can be combined with prison and the fine is not constrained by the offender's wealth, the optimal fine achieves first-best deterrence, and the optimal prison term is therefore zero. In this case, there is no social gain from incapacitation because only efficient crimes are

ever committed. When the fine is limited by the offender's wealth, however, the optimal fine is maximal, and the optimal prison term is determined by the same trade-off as in the prison-only model. (3) When the probability of apprehension is endogenous, the optimal prison term is maximal if it is the only form of punishment. But if fines are also available, the optimal fine is maximal, and the optimal prison term is no longer generally maximal. (4) In the model where the offender's utility is not counted as part of social welfare, if prison is the only form of punishment and the probability of apprehension is fixed, the optimal prison term is either finite (if the harm from crime is less than the cost of imprisonment), or infinite (if the harm from crime is greater than the cost of imprisonment). In the former case, deterrence and incapacitation work in opposite directions, while in the latter case they reinforce each other.

References

- Becker, Gary (1968) "Crime and Punishment: An Economic Approach," *Journal of Political Economy* 76: 169-217.
- Davis, Michael (1988) "Time and Punishment: An Intertemporal Model of Crime," *Journal of Political Economy* 96: 383-390.
- Ehrlich, Isaac (1981) "On the Usefulness of Controlling Individuals: An Economic Analysis of Rehabilitation, Incapacitation and Deterrence," *American Economic Review* 71: 307-332.
- Emons, Winand (2004) "Subgame Perfect Punishment for Repeat Offenders," *Economic Inquiry* 42: 496-502.
- Levitt, Steven and Thomas Miles (2007) "Empirical Study of Criminal Punishment," in *Handbook of Law and Economics*, A. Mitchell Polinsky and Steven Shavell, eds., Vol. 1, pp. 455-495. Amsterdam: North-Holland.
- Lewin, Jeffrey and William Trumbull (1990) "The Social Value of Crime?" *International Review of Law and Economics* 10: 271-284.
- Loury, Glenn (1979) "Market Structure and Innovation," *Quarterly Journal of Economics* 93: 295-310.
- Miceli, Thomas (1991) "Optimal Criminal Procedure: Fairness and Deterrence," *International Review of Law and Economics* 11: 3-10.
- Miceli, Thomas and Catherine Bucci (2005) "A Simple Theory of Increasing Penalties for Repeat Offenders," *Review of Law and Economics* 1: 71-80.
- Mortensen, Dale (1982) "Property Rights in Mating, Racing, and Related Games," *American Economic Review* 72: 968-979.
- Polinsky, A Mitchell and Daniel Rubinfeld (1991) "A Model of Optimal Fines for Repeat Offenders," *Journal of Public Economics* 46: 291-306.
- Polinsky, A Mitchell and Steven Shavell (2007) "The Theory of Public Enforcement of Law," in *Handbook of Law and Economics*, A. Mitchell Polinsky and Steven Shavell, eds., Vol. 1, pp. 403-454. Amsterdam: North-Holland.
- _____ (2000) "The Economic Theory of Public Law Enforcement," *Journal of Economic Literature* 38: 45-76.
- _____ (1998) "On Offense History and the Theory of Public Law Enforcement," *International Review of Law and Economics* 18: 305-324.

_____ (1984) “The Optimal Use of Fines and Imprisonment,” *Journal of Public Economics* 24: 89-99.

Shavell, Steven (1987) “A Model of Optimal Incapacitation,” *American Economic Review* 77: 107-110.

Shepherd, Joanna (2002) “Fear of the First Strike: The Full Deterrence Effect of California’s Two- and Three-Strikes Legislation,” *Journal of Legal Studies* 31: 159-201.

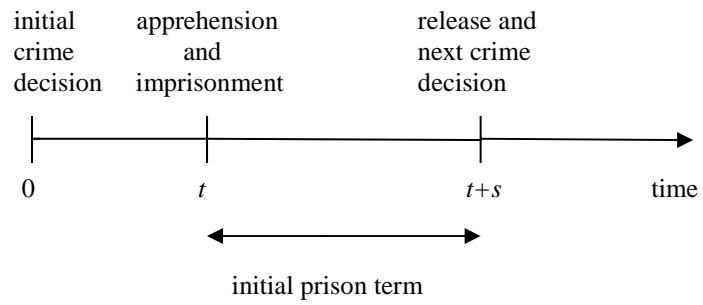


Figure 1. Time line of crime and punishment over an infinite horizon.