March 2009

Deterrence and Incapacitation: Towards a Unified Theory of Criminal Punishment

Thomas J. Miceli
University of Connecticut

Department of Economics Working Paper Series

Deterrence and Incapacitation: Towards a Unified Theory of Criminal Punishment

Thomas J. Miceli
University of Connecticut

Working Paper 2009-11

March 2009

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

Abstract

Economic models of crime have focused primarily on the goal of deterrence; the goal of incapacitation has received much less attention. This paper adapts the standard deterrence model to incorporate incapacitation. When prison only is used, 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. In contrast, when a fine is available and it is not constrained by the offender's 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.

Journal of Economic Literature Classification: K14, K42

Keywords: Career criminals, deterrence, incapacitation, law enforcement

Deterrence and Incapacitation:
Towards a Unified Theory of Criminal Punishment

Economic models of law enforcement since Becker (1968) have focused primarily on the goal of deterrence (Polinsky and Shavell, 2000, 2007). Much less examined by economists has been the incapacitation function of imprisonment (exceptions are Ehrlich (1981) and Shavell (1987)). Yet actual law enforcement policies almost certainly combine both motives for punishment, as reflected by the seeming "overuse" of prison from the perspective of deterrence models, which prescribe the use of fines up to the limit of a defendant's wealth before imposing prison time. Recent three-strikes laws, which imprison an offender for life on a third offense, are an example. (See, for example, Shepherd (2002).)

The purpose of this note is to develop an economic model of law enforcement that combines the deterrence and incapacitation motives for punishment. Analytically, this requires adding a dynamic element to the standard Becker-Polinsky-Shavell (BPS) deterrence model, which treats the potential offender's decision of whether or not to commit a crime as one-time. In contrast, incapacitation envisions some offenders as being habitual (undeterrable) in the sense that they commit repeated crimes whenever free, irrespective of the threat of punishment. In the hybrid model, offenders are infinitely lived and potentially commit crimes throughout their lives, but they may also be deterred by the threat of punishment. In this setting, prison potentially serves the dual functions of deterring some offenders from ever committing crimes, and incapacitating those who do.

I. The Model

Potential offenders have infinite life spans. At time zero, they each take a random draw of the monetary gain from committing a crime, g, which is distributed by the density function $z(g)$. Each offender's realized g will remain his "type" throughout his life. (Thus, each offender will make the same choice each time he is confronted with a criminal opportunity.) Time is continuous, and g is therefore defined to be the gain per instant of time that an offender is free and committing crimes.[1]

Having drawn g at time zero, an offender makes his first criminal decision instantly. If he commits the crime, he enjoys the gain g up to the time he is apprehended, time t. Once apprehended, he is imprisoned for a length of time s and is then released. (Below we also consider fines as a form of punishment.) He then immediately confronts another criminal opportunity, and the process begins again. Figure 1 illustrates the time line of events.

The apprehension technology is described as follows.[2] The time until apprehension, t, is a random variable with an exponential density function

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

where p is the instantaneous probability of apprehension. The expected time until apprehension is therefore equal to $1/p$. Throughout the analysis, we treat p as fixed and focus on the optimal level of punishment (prison term and/or fine).

A. The Offender's Gain

We first derive the expected lifetime gain to the offender at the point that he makes his initial crime decision. As noted, if he commits the crime he enjoys the gain g continuously up to his capture at time t, is imprisoned for s periods, and then is released at time $t+s$. Letting δ be the disutility of prison per unit of time, we obtain the following net return for the offender's first episode of crime and punishment as a function of g and t:

[1] This assumption is not essential. We could also assume that the entire gain g is realized at the first instant the offender commits the crime.
[2] This formulation follows Davis (1988). It is also the approach used by Loury (1979) and Mortensen (1982) to describe the process by which firms make technological advances.

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

where r is the discount rate. We next compute the expected value of this expression by weighting it by the density function in (1) and integrating over all t. This yields the expected net gain from the initial crime and punishment as a function of the offender's type:

$$G_1(g) = \int_0^\infty \left[\frac{g}{r}(1-e^{-rt}) - \frac{\delta}{r}(e^{-rt} - e^{-r(t+s)}) \right] pe^{-pt} dt$$

$$= \frac{1}{p+r}\left[g - \frac{p\delta}{r}(1-e^{-rs}) \right]. \tag{3}$$

As noted, the offender will repeat his commission decision as soon as he is released over an infinite horizon. Given time invariance, the present value of his lifetime net gains from crime, denoted $G(g)$, is therefore

$$G(g) = \frac{1}{p+r}\left[g - \frac{p\delta}{r}(1-e^{-rs}) \right] + \beta(s)G(g), \tag{4}$$

where $\beta(s)$ is the expected discount factor. It is computed as follows:

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

Substituting (5) into (4) and solving for $G(g)$ yields

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

A potential offender will commit a crime at time zero (and at each subsequent criminal opportunity) if $G(g)>0$. Thus, those offenders for whom

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

become career criminals (i.e., the commit crimes whenever free). Note that $\partial \hat{g} / \partial s = p\delta e^{-rs} > 0$, which provides the basis for the deterrence function of prison. It will also be useful to note that

$$\frac{\partial G}{\partial s} = \frac{-pre^{-rs}(g+\delta)}{[p(1-e^{-rs})+r]^2} < 0. \tag{7}$$

Thus, as expected, increasing the length of the prison term reduces the lifetime return from crime, both because it imposes marginal disutility of δ, but also because it deprives the offender of the incremental gains from additional crimes.

B. The Social Cost of Crime

Now consider the cost of crime to society. This consists of the usual three components: the harm suffered by victims at each instant that a criminal is free and committing crimes, denoted by h; the cost per unit of time of imprisoning an offender, denoted by c;[3] and the cost of enforcement, $k(p)$, which we take as fixed. As described above, the harm to victims is incurred continuously up to the time when the offender is apprehended, t, while the cost of imprisonment is incurred from t up to $t+s$. Further, these costs are repeated during all future episodes of crime and imprisonment. The cost of enforcement, however, is incurred only once. Proceeding as above, we calculate the present value of harm plus imprisonment costs:

$$C = \frac{1}{p+r}\left[h + \frac{pc}{r}(1-e^{-rs})\right] + \beta(s)C\ . \tag{8}$$

Again, substituting for $\beta(s)$ from (5) and solving for C yields

$$C = \frac{1}{p(1-e^{-rs})+r}\left[h + \frac{pc}{r}(1-e^{-rs})\right]. \tag{9}$$

Taking the derivative of (9) with respect to s yields

[3] The disutility of imprisonment suffered by offenders will also be included as a social cost in the welfare function of the hybrid model below.

$$\frac{\partial C}{\partial s} = \frac{pre^{-rs}(c-h)}{[p(1-e^{-rs})+r]^2}, \tag{10}$$

which is ambiguous in sign, depending on the relative magnitudes of *c* and *h*. Specifically, if $c>h$, (10) is positive, and costs are minimized by setting $s=0$. In this case, the cost of imprisonment is greater than the harm from crime, so society is better off not imprisoning offenders. Conversely, if $c<h$, (10) is negative, and costs are minimized by making *s* infinite. In this case, the cost of imprisonment is less than the harm from crime, so society is better off imprisoning offenders for life after their first offense. This trade-off represents the economic approach to incapacitation as studied by Shavell (1987).[4]

C. Optimal Punishment in the Hybrid Model

Social welfare in for the hybrid model consists of the gain to offenders less the social costs of crime, summed over those offenders who commit crimes.[5] Formally,

$$W = \int_{\hat{g}(s)}^{\infty} [G(g)-C]z(g)dg - k(p).$$

$$= \int_{\hat{g}(s)}^{\infty} \frac{1}{p(1-e^{-rs})+r}\left[g-h-\frac{p(c+\delta)}{r}(1-e^{-rs})\right]z(g)dg - k(p). \tag{12}$$

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

$$\frac{\partial W}{\partial s} = -\left[G(\hat{g})-C\right]z(\hat{g})\left(\frac{\partial \hat{g}}{\partial s}\right) + \int_{\hat{g}}^{\infty}\left(\frac{\partial G}{\partial s}-\frac{\partial C}{\partial s}\right)z(g)dg = 0. \tag{13}$$

[4] Shavell also considers the case where the harm caused by an offender, *h*, decreases over his lifetime. Thus, assuming $c<h$ initially, the offender should be imprisoned but then released when $c>h$.

[5] I make the usual assumption that the gains to the offender should be counted as part of social welfare. See the discussion of this point in Polinsky and Shavell (2000, p.48, note 12).

As usual, the net gain to the marginal offender in this expression, $G(\hat{g})$, is zero and therefore drops out. Substituting for the $\partial\hat{g}/\partial s$, $\partial G/\partial s$, and $\partial C/\partial s$ terms from above and simplifying yields

$$\left[h+\frac{pc}{r}(1-e^{-rs})\right]z(\hat{g})\delta=\frac{1}{p(1-e^{-rs})+r}\int_{\hat{g}}^{\infty}(g+\delta+c-h)z(g)dg\ . \qquad (14)$$

The left-hand side of this condition is the marginal deterrence benefit of increasing *s*, as reflected by the savings in harm to victims plus punishment costs. This term is essentially identical to that in the standard BPS deterrence model; the only difference is that the saved punishment costs are in present value terms.

The right-hand side of (14) is the marginal cost of increasing *s*. This consists of the usual cost of lengthening the prison term in the form of the incremental cost of imprisonment, *c*, plus the disutility to the offender, δ. In the standard BPS model, these would be the only components of the marginal cost term.[6] Here, however, there is an additional term, *g–h*, which represents the foregone net benefits from those crimes that career criminals (i.e., those who are not deterred) are unable to commit because they are kept in prison longer. This reflects the incapacitation effects of imprisonment.[7] Generally, the sign of this term may be positive or negative, depending on whether the expected gain to undeterred offenders is larger or smaller than the harm they impose on victims, given optimal punishment. If this term is negative, the marginal cost of prison is reduced, making the optimal prison term longer. In this case, prison serves an incapacitation function by reducing the amount of time offenders have to commit further crimes, which on average are inefficient (given that $E(g|g\geq\hat{g}(s^*))<h$).

[6] See, for example, Polinsky and Shavell (2007), equation (6).
[7] Another way to interpret the right-hand side of (14) is that $g+\delta$ represents the marginal cost of a longer prison term to the offender in terms of the foregone gains from crime plus the disutility of prison, while the *c–h* term represents the social trade-off associated with incapacitation.

In contrast, if this extra term is positive, then the optimal prison term should actually be reduced compared to the pure deterrence model. The reason for this anomalous effect is that, if undeterred offenders on average gain more than the harm they are imposing, there is a net social cost of keeping them in prison longer because it prevents them from committing further "efficient" crimes. Taken together, the preceding results show, surprisingly, that introducing incapacitation into the economic model of crime can result in either a higher or a lower optimal prison sentence compared to the pure deterrence model.

D. Fines and Prison

Now let us add fines as an additional form of punishment. Obviously, fines can have no effect on incapacitation, but they may reduce (or eliminate) the need to rely on prison for deterrence, in which case it can used solely for incapacitation. Let *f* be the amount of the fine, which is imposed on an offender each time he is apprehended.

Note first that the gain to the offender has to be altered to reflect the fine. Proceeding as above, we obtain

$$G'(g) = \frac{1}{p(1-e^{-rs})+r}\left[g - pf - \frac{p\delta}{r}(1-e^{-rs})\right] \tag{15}$$

as the lifetime gain from crime for an offender of type *g*. The threshold level of *g* above which the offender becomes a career criminal is thus

$$\hat{g}(f,s) = pf + \frac{p\delta}{r}(1-e^{-rs})\,. \tag{16}$$

Social costs must also be amended to reflect the fine revenue. Again, proceeding as above we obtain

$$C' = \frac{1}{p(1-e^{-rs})+r}\left[h + \frac{pc}{r}(1-e^{-rs}) - pf\right] \tag{17}$$

as the present value of harm plus net punishment costs (i.e., the cost of imprisonment less the expected fine revenue). The resulting expression for welfare is

$$W' = \int_{\hat{g}(f,s)}^{\infty} [G'(g) - C']z(g)dg - k(p)$$

$$= \int_{\hat{g}(f,s)}^{\infty} \frac{1}{p(1-e^{-rs})+r}\left[g - h - \frac{p(c+\delta)}{r}(1-e^{-rs})\right]z(g)dg - k(p)\,, \qquad (18)$$

where, notice, the fine revenue drops out since it is merely a transfer payment. Thus, (18) differs from (12) only by the lower limit of integration.

The optimal punishment scheme involves choosing *f* and *s* to maximize (18). The usual result in the standard BPS deterrence model in this case is that it is never optimal to impose a prison sentence unless the fine is constrained by the offender's wealth. The intuition is that, since both the fine and prison are equally effective in deterring crime, it would never be optimal to use the costly tool (prison) until the costless tool (a fine) has been fully exhausted.[8] The question is whether this same result holds in the hybrid model.

To answer this question, we first derive the optimal fine, assuming no limit on the offender's wealth. Setting the derivative of (18) with respect to *f* equal to zero and solving for *f* yields

$$f^* = h/p + \frac{c}{r}(1-e^{-rs})\,. \qquad (19)$$

Thus, the optimal fine equals the social cost of a crime, which consists of the harm suffered by victims plus the present value of the expected cost of imprisoning the offender, both appropriately adjusted to reflect the probability of apprehension. In the case where the prison term is zero, (19) reduces to the usual expression for the optimal fine, $f^*=h/p$.

[8] See Polinsky and Shavell (2007, p. 411) for a more formal proof.

Now differentiate (18) with respect to *s*, given *f*=*f**:

$$\frac{\partial W'}{\partial s}\Big|_{f=f^*} = \frac{pre^{-rs}}{[p(1-e^{-rs})+r]^2} \int_{\hat{g}(f^*,s)}^{\infty} (-g-\delta-c+h)z(g)dg \, . \tag{20}$$

To determine whether a positive prison term is optimal, we evaluate this derivative at *s*=0 and check whether it is positive or negative. Note first that $\hat{g}(f^*,0)=h$, in which case (20) becomes

$$\frac{\partial W'}{\partial s}\Big|_{f=f^*,s=0} = \frac{p}{r}\int_{h}^{\infty} (-g-\delta-c+h)z(g)dg < 0, \tag{21}$$

where the negative sign follows by the fact that the integration is over *g*≥*h*. Thus, the expression in parentheses must be negative. It follows that *s**=0. Intuitively, since the fine is set to achieve optimal (first-best) deterrence, only efficient crimes are committed. Thus, from an economic perspective, there is no social gain from incapacitating repeat offenders. In the prison-only formulation above, in contrast, it was possible that there were incapacitation benefits since the optimal prison term possibly resulted in underdeterrence at the optimum.

In the case where the fine is limited by the offender's wealth (i.e., *w*<*f**), the foregoing logic implies that a positive prison term may be desirable for purposes of incapacitation. The trade off is the same as that for the prison-only case.

As a final point, it is interesting to note that the conclusions from the hybrid model shed light on the difficulty that the literature has had in explaining the pervasiveness of escalating penalty schedules for repeat offenders. If deterrence is efficient, as in the model with both fines and prison, then there is no reason to punish repeat offenders more severely, either for purposes of deterrence or incapacitation, because they are in fact acting efficiently. An escalating penalty scheme can therefore only be optimal if there is underdeterrence of first-time offenders.[9]

---

[9] See, for example, Polinsky and Rubinfeld (1991), Polinsky and Shavell (2000, p. 67), Shavell (2004, p. 529), Miceli and Bucci (2005), and Polinsky and Shavell (2007, p. 438).

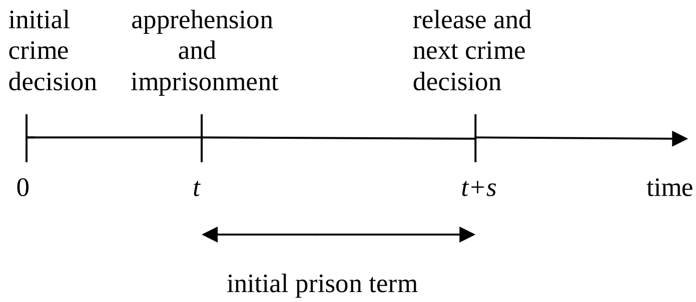

Figure 1. Time line of crime and punishment over an infinite horizon.

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.

Loury, Glenn (1979) “Market Structure and Innovation,” *Quarterly Journal of Economics* 93: 295-310.

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.

Shavell (2004) *Foundations of Economic Analysis of Law*, Cambridge, MA: Belknap.

_______ (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.