



University of
Connecticut

Department of Economics Working Paper Series

**Deterrence and Incapacitation Models of Criminal Punishment:
Can the Twain Meet?**

Thomas J. Miceli
University of Connecticut

Working Paper 2009-25

August 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 focuses on the goal of deterrence, but actual punishment schemes, most notably recent three-strikes laws, seem to rely more on imprisonment than is prescribed by that model. One explanation is that prison also serves an incapacitation function. The current paper seeks to develop an economic model of law enforcement that combines the deterrence and incapacitation motives for criminal punishment. The resulting hybrid model retains the rationality assumption that is the basis of the pure deterrence model, but assumes that offenders face repeated criminal opportunities over their lifetimes. In this setting, deterrence and incapacitation emerge naturally as complementary motivations for imposing criminal punishment.

Journal of Economic Literature Classification: K14, K42

Keywords: Deterrence, incapacitation, law enforcement, prison

Deterrence and Incapacitation Models of Criminal Punishment: Can the Twain Meet?

1. Introduction

Economic models of law enforcement beginning with Becker (1968) have primarily focused on the role of criminal punishment in deterring crime. This approach to the determination of optimal criminal penalties relies on the *rational offender assumption*, which maintains that potential offenders decide whether or not to commit an illegal act by comparing the gain from commission to the expected punishment. Although some may doubt the validity of this assumption, there is a growing body of empirical evidence to support it (as reviewed in the next section).

One of the clearest policy implications emerging from this model is that fines should be relied upon to the maximum extent possible before imprisonment is used. The obvious reason is that, while fines and prison are equally capable of deterring rational offenders, fines are costless to impose while prison is costly. The use of prison should therefore be limited to those offenders whose lack of wealth makes the threat of a heavy fine ineffective as a deterrent (Polinsky and Shavell, 1984). The extensive use of prison in actual punishment schemes, however, appears to be inconsistent with this prescription.

One explanation for this practice is the desire for equal treatment of rich and poor offenders, given that the economically efficient punishment scheme would essentially allow rich offenders to “buy their way out of jail” (Lott, 1987). Another explanation is that prison serves an incapacitation function; that is, it allows the state to detain those offenders who are expected to commit further harmful acts if released. Recent three-strikes laws enacted by many states, which imprison certain repeat offenders for life, appear to be motivated primarily by this rationale

(Shepherd, 2002). Economists have devoted relatively little attention to incapacitation as basis for criminal punishment. An exception is Shavell (1987), who shows that the optimal incapacitation policy involves holding an offender in prison as long as the harm he is expected to impose if free exceeds the cost of imprisonment. The incapacitation model is silent, however, about why offenders choose to commit crimes in the first place.

As the literature stands, therefore, the deterrence and incapacitation models exist as separate theories of criminal punishment. The deterrence theory is the more refined and elegant of the two, and it clearly occupies the predominant position in the literature, but its inability to provide an adequate explanation for the actual use of prison undermines its status as a positive theory of law enforcement. The incapacitation theory, in contrast, is more ad hoc from a theoretical perspective, but it offers a more convincing explanation for certain imprisonment policies. Clearly, what is needed is a model of law enforcement that integrates the best of both theories. The purpose of this article is to outline such a model.

The resulting unified (or hybrid) model of criminal punishment retains the theoretical rigor of the deterrence model by assuming that offenders are fully rational. Thus, potential offenders make crime decisions based on the expected punishment, including the possibility of imprisonment. A role for incapacitation is introduced into this setting by allowing potential offenders to face repeated criminal opportunities over an infinite lifetime. The threat of imprisonment therefore deters some offenders from committing crimes in the first place, while the detention of previously convicted offenders prevents them from committing further crimes by depriving them of future criminal opportunities. In this way, the model embodies both the deterrence and incapacitation functions of imprisonment within a single, coherent framework.

The remainder of the article is organized as follows. Section 2 reviews the empirical evidence on the relationship between imprisonment policy and crime. While the evidence clearly shows that prison has a crime-reducing effect, this result is consistent with both the deterrence and incapacitation theories. By employing methods for disentangling the two effects, however, econometricians have shown that both are relevant. Section 3 reviews the standard economic models of deterrence and incapacitation, and discusses their compatibility. Section 4 then lays out the basic hybrid model, and shows that the pure deterrence and pure incapacitation models emerge as special cases. It goes on to show that, when prison is the only form of punishment, adding incapacitation can result in either a longer or a shorter prison term compared to the pure deterrence model. Intuitively, if there is underdeterrence in the pure deterrence model, then introducing incapacitation will cause the optimal prison term to be longer because it prevents offenders from committing further inefficient acts. Conversely, if there is overdeterrence, incapacitation will cause the optimal prison term to be shorter so that offenders are able to commit further efficient acts. When a fine is combined with prison and the fine is not constrained by the offender's wealth, then, as in the pure deterrence model, it is never optimal to use prison either for deterrence or incapacitation. The reason is that the optimal fine achieves the efficient (first-best) level of crime, so only efficient crimes are ever committed. Thus, there is no social gain from incapacitation. However, when the fine is limited by the offender's wealth, the optimal prison term is determined by the same factors as in the prison-only version of the model.

Section 5 extends the model in two ways. First, it allows the probability of apprehension to be endogenous. Second, it examines a version of the model in which the offender's utility is not counted as part of social welfare. This case is of interest because most crimes for which incapacitation is relevant are harmful to society and hence are not likely to be socially desirable.

In the prison-only version of this model, the optimal prison term is either finite (with deterrence and incapacitation offsetting each other), or infinite (with deterrence and incapacitation reinforcing each other). This last case seems most descriptive of the rationale for three-strikes laws. Finally, Section 6 offers concluding comments.

2. Empirical Evidence on the Impact of Imprisonment on Crime

Empirical analyses of the impact of imprisonment policies have focused on measuring their effects in reducing crime.¹ In particular, they ask whether increases in the use of prison as a criminal sanction—as reflected by more frequent use of prison and/or longer prison terms—are associated with a reduction in the crime rate, as predicted by both the deterrence and incapacitation models of crime. Although the hypothesis is a straightforward one, uncovering such a causal effect from aggregate crime data presents significant empirical problems. One of the difficulties is illustrated by Figure 1, which graphs the rate of violent crimes and the imprisonment rate in the United States from 1980-2006.² Depending on the particular time interval, one can observe either a positive or a negative correlation between the variables. Thus, by focusing on a particular time period, one can either conclude that an increase in the use of prison has not reduced the crime rate (and may actually have increased it), or that it has had the desired crime-reducing effect. The problem stems from a failure to account for multiple causal influences on the crime rate besides imprisonment, as well as the likely feedback effect of the observed crime rate on prison policy (i.e., the tendency for policymakers to respond to rising crime rates by enacting more stringent criminal policies, resulting in a reverse-causal effect).

[Figure 1 here]

¹ See the survey by Levitt and Miles (2007), on which the current section is based.

² The data were obtained from the Statistical Abstract of the United States, various years.

Recent studies have employed sophisticated empirical techniques to overcome these challenges. The results suggest that increases in the rate of imprisonment do in fact cause a decline in the crime rate. Recall, however, that this conclusion is consistent with both the deterrence and incapacitation models. In other words, the mere establishment of a causal connection between increased use of imprisonment and lower crime rates does not tell us whether this is due to a behavioral response of rational offenders who choose to commit fewer crimes for fear of punishment, or whether it is because offenders who otherwise would have committed crimes are deprived of the opportunity to do so because they are held in prison longer. In addition to its theoretical relevance, this distinction is important from a policy perspective because, as we will see, the specific prescriptions from the two models are different.

Fortunately, economists have found ways of disentangling the two effects. For example, Kessler and Levitt (1999) looked for changes in the crime rate immediately after California enacted Proposition 8, which provided for enhanced prison sentences for certain serious crimes. Since the incapacitating effect of the new law could only take effect after the standard prison term had run, any observed reduction in the crime rate before that time would have had to be solely due to deterrence. The authors in fact found that the crime rate fell more for affected offenses than for non-affected offenses in the year after adoption, showing that deterrence mattered. However, they also found that the crime rate fell twice as much in the three years after adoption as it did in the first year, suggesting that incapacitation had also contributed to the overall decline in the crime rate. Other studies have yielded similar results, confirming that deterrence and incapacitation effects are both relevant factors in assessing the impact of prison policies.

3. The Economic Theory of Crime and Punishment

The empirical finding that both deterrence and incapacitation effects matter sets the agenda for theoretical models seeking to explain the observed relationship between crime and punishment. This section reviews the basic versions of both the pure deterrence and pure incapacitation models by way of providing a context for the hybrid model to be developed in the next section.

3.1. The Economic Theory of Deterrence

As noted, the economic theory of criminal punishment is primarily based on the goal of deterrence. Although such a theory was discussed as early as the eighteenth century by Beccaria (1767) and Bentham (1789), the modern mathematical version was first developed by Becker (1968) and elaborated on by Polinsky and Shavell (2000, 2007). I will hereafter refer to this as the BPS model of deterrence. The key behavioral assumption underlying this model is the rational offender assumption, which maintains that would-be criminals decide whether or not to commit illegal acts in the same way that they would make any other economic decision; namely, by comparing the expected gain from committing the act to the expected punishment, where the latter consists of the probability of apprehension and conviction multiplied by the sanction (a fine and/or imprisonment term). If the expected gain exceeds the expected sanction, the offender commits the act; otherwise, he is deterred.

Summing over all offenders who choose to commit illegal acts yields the aggregate crime rate, which, by virtue of the rationality assumption, is decreasing in the severity of the expected sanction. In other words, increases in both the likelihood of apprehension and the severity of the sanction have the effect of reducing the crime rate. Based on this relationship, policymakers can choose the law enforcement policy that achieves the socially optimal crime rate. This is usually taken to be the crime rate that maximizes a social welfare function that depends on the net cost of

crime to society (consisting of the harm to victims less any acceptable benefits to offenders) and the cost of enforcement.

To see this formally, let

g = monetary gain from committing a criminal act;

$z(g)$ = density function reflecting the distribution of gains across offenders;

h = harm caused by a criminal act (assumed to be fixed);

p = probability of apprehension and conviction;³

$k(p)$ = cost of maintaining an apprehension rate of p , $k' > 0$, $k'' \geq 0$;

f = fine imposed on conviction;

s = length of the prison term imposed on conviction;

δ = unit cost of prison to the offender;

c = unit cost of prison to society.

In the case where the sanction consists of a fine and imprisonment, the expected sanction from the offender's perspective is given by $p(f + \delta s)$. After observing the gain, g , he will therefore commit the crime if and only if

$$g \geq p(f + \delta s) \equiv \hat{g}, \quad (1)$$

where \hat{g} is the threshold gain separating those offenders who commit crimes from those who are deterred. Thus, condition (1) is the embodiment of the rational offender assumption. Since the gain is distributed by $z(g)$ across potential offenders, the aggregate crime rate is given by $1 - Z(\hat{g})$, where Z is the distribution function associated with z (i.e., $Z' \equiv z > 0$). It follows that the

³ The model abstracts from the adjudication of guilt by assuming that all offenders who are apprehended are convicted. This of course ignores the possibility that some guilty defendants will be acquitted at trial, and some innocent defendants convicted. For a model that incorporates these errors, see Miceli (1991).

crime rate is decreasing in p, f , and s , reflecting the deterrent effect of increases in both the likelihood of apprehension and the severity of sanctions.

Social welfare in the BPS model consists of the net gain to offenders from committing illegal acts, minus the harm and enforcement costs. It is typically assumed in deterrence models that the gain to offender should count as part of social welfare, making the net social gain from a given crime $g-h$. This assumption is questioned by some (see, for example, Stigler (1970) and Lewin and Trumbull (1990)), and clearly is more reasonable for some types of crime (e.g., speeding, double parking) than for others (e.g., violent crimes). Rather than debate this point, however, I will maintain the standard assumption for most of the analysis. (However, Section 5.2 below examines the effect of excluding the offender's gains from welfare in the context of the hybrid model.)

Formally, welfare in the BPS model is given by

$$W = \int_{\hat{g}}^{\infty} (g - h - p(c + \delta)s)z(g)dg - k(p), \quad (2)$$

where the integral represents offender's gain minus the harm and expected punishment costs, summed over all offenses (where \hat{g} is defined by (1)), while $k(p)$ represents apprehension costs.

The enforcement authority is assumed to maximize this expression by its choice of the policy variables f, s , and p . It is easiest to see the optimum by first supposing that the probability of apprehension, and hence apprehension costs, are fixed.⁴ In this case, we first consider punishment by a fine alone ($s=0$), and then by prison alone ($f=0$).

In the fine-only punishment scheme, we set the derivative of (2) with respect to f equal to zero and solve for f to obtain

$$f^* = h/p. \quad (3)$$

⁴ This is the approach adopted by Polinsky and Shavell (2000, 2007).

Thus, the optimal fine equals the harm per crime, appropriately adjusted to reflect the uncertainty of apprehension. In this case, only efficient crimes (those for which $g > h$) are committed. To illustrate, suppose that the harm imposed per criminal act is \$500. Thus, only those offenders who expect a gain of more than \$500 should commit the act. Optimal deterrence therefore requires the expected fine, pf , to be set equal to \$500. If the probability of apprehension is $\frac{1}{2}$, this requires the *actual* fine to be set at \$1,000.

Obviously, the above policy is limited by the wealth of the offender, w , which will prevent the attainment of the first-best outcome for those offenders whose wealth is less than h/p . As will be shown below, this problem provides the economic rationale for use of imprisonment.

Before considering the combined use of fines and prison, however, we consider the optimal prison term when it is the only possible sanction. This is found by maximizing (2) with respect to s with $f=0$. Assuming an interior solution, the resulting first order condition is given by

$$(h + pcs)z(\hat{g})\delta = [1 - Z(\hat{g})](c + \delta). \quad (4)$$

The left-hand side of this condition represents the marginal benefit of a longer prison term in the form of the avoided harm plus the expected savings in punishment costs. The right-hand side is the marginal cost of punishment, consisting of the number of crimes multiplied by the incremental cost to society and the offender of lengthening the prison term. Unlike the case of a fine, there is no simple formula for the optimal prison term.

Now suppose that fines and prison can be combined. One of the key prescriptions of the BPS model of crime is that when both sanctions are available, prison should never be used unless the offender's wealth precludes setting the fine at the level prescribed by (3) (Polinsky and Shavell, 1984). This is easily proved by supposing initially that $f < w$ and $s > 0$. Now raise f and

lower s so that the critical gain, \hat{g} , remains constant. According to (1), the crime rate will remain unchanged, but expected punishment costs, $p(c+\delta)s$, will fall, implying that welfare must increase. Thus, the original scheme with $f < w$ could not have been optimal. Intuitively, it is never optimal in the BPS model to impose a prison term rather than a fine for deterrence purposes for the simple reason that the two sanctions are equally capable of deterring crimes, but fines are costless to impose while prison is costly. Thus, only when the offender's wealth does not allow the fine to be set at the first-best level in (3) is it possibly desirable to impose a prison term. In this case, the optimal prison term is found by maximizing (2) with respect to s with $f=w$. Since prison is costly, this may or may not result in a positive prison, depending on the magnitude of the marginal deterrence benefits compared to the marginal cost of imprisonment.

As an example, suppose that the social optimum entails a level of offender gains, \hat{g} , equal to \$4,000, meaning that it is only efficient for those offenders who receive a benefit of more than \$4,000 to commit crimes. Also, let the probability of apprehension be $\frac{1}{2}$, let the unit cost of prison to the offender be \$500 per month, and let the offender's wealth be \$2,000. Then from (1) we have $\hat{g} = p(f+\delta s) = \$4,000$. After substituting $p = \frac{1}{2}$, we obtain $f+\delta s = \$8,000$. Since this amount exceeds the offender's wealth, the optimal fine is maximal, or $f = \$2,000$, while the optimal prison term solves $\$2,000 + (\$500)s = \$8,000$, or $s = 12$ months. The socially optimal punishment scheme thus involves a fine of \$2,000 and one year in prison.

Finally, consider the effect of allowing the enforcement authority to choose the probability of apprehension along with the severity of sanctions. In this case, the optimal sanction, in both the fine-only and the prison-only schemes, is maximal.⁵ The proof of this proposition proceeds as above. Specifically, suppose initially that the sanction (the fine or the

⁵ In the case of prison, a maximal sanction might be life imprisonment, or some other statutorily determined maximum term.

prison term) is less than maximal. Now raise the sanction and lower p so as to hold the critical gain, \hat{g} , fixed. Since this lowers $k(p)$ while holding the integral term in (2) fixed, welfare must increase. Thus, the initial punishment scheme could not have been optimal. This conclusion makes intuitive sense in the fine-only case since, as before, increasing the fine is costless while increasing p is costly. It is less obvious, however, why the prison term should be maximal. The reason is that only those offenders who are caught are imprisoned. Thus, by lowering p , fewer offenders are caught and imprisoned, thereby lowering (or at least not raising) expected punishment costs.

When fines and imprisonment are combined, the optimal fine is still maximal, but the optimal prison term may not be maximal. This is true because, when the fine is set equal to the offender's wealth, simultaneously raising the prison term and lowering the probability of apprehension so as to hold the expected cost of prison fixed will *reduce* deterrence because the expected fine, pw , falls. Thus, it is not necessarily welfare-enhancing to continue to raise s while proportionately lowering p . The optimal prison term in this case depends, as before, on the particular relationship between the marginal deterrence benefits and the marginal cost of imprisonment.

3.2. *The Economic Theory of Incapacitation*

Incapacitation protects society from the harm caused by criminals, not by deterring them, but by depriving them of the opportunity to commit crimes. Imprisonment is therefore the primary form of incapacitation. Unlike the deterrence model, however, incapacitation is not concerned with an offender's decision about whether or not to commit a crime. Instead, it takes the crime rate as given and asks whether social costs are lower if an offender, once apprehended, is detained or released. Specifically, the comparison is between c , the cost of holding the

criminal in prison, and h , the expected harm that he would impose if released, where both are defined per unit of time. If $c > h$, the offender should be released, but if $c < h$, he should be imprisoned and detained for as long as the inequality continues to hold, possibly for the remainder of his life. In the likely case where the harm an offender would impose declines with his age, the optimal policy is therefore to release him as soon as the threatened harm falls below the cost of holding him in prison (Shavell, 1987).

3.3. Are the Economic Models of Deterrence and Incapacitation Compatible?

At present, the deterrence (BPS) and incapacitation models represent distinct strands in the economics of crime literature, with the BPS model being the predominant paradigm. As noted, however, there is a somewhat troubling disconnect between the prescriptions of the BPS model and actual criminal policy, especially as regards the use of prison. In particular, actual practice seems to involve a significantly greater reliance on prison than the BPS model prescribes, especially as exemplified by the recent spate of three-strikes laws that impose life sentences on repeat offenders for certain crimes. The goal of incapacitation would seem to be a better explanation for such a policy.

In principle, however, there is no reason why deterrence and incapacitation cannot co-exist as complementary economic theories of criminal punishment. From a deterrence perspective, the threat of punishment should prevent some offenders from committing dangerous acts in the first place, while from an incapacitation perspective, the imprisonment of those offenders who are not deterred, or who are not deterrable (for whatever reason), will prevent them from committing further harmful acts. The literature, however, has yet to offer a fully integrated model that captures both of these approaches to crime prevention (though Ehrlich (1981) represents an early effort).

A necessary first step in developing such an integrated model is to make the BPS model dynamic so as to introduce the time dimension that is inherent in the incapacitation motive. Several recent efforts along these lines have been made by way of investigating the pervasiveness of escalating penalty schemes for repeat offenders,⁶ but none of these models has explicitly addressed the question of incapacitation. Still, these studies have revealed an important insight that sheds light on the compatibility of deterrence and incapacitation models. In particular, they have shown that escalating penalties can never be optimal in a pure deterrence model when penalties are structured so as to achieve efficient (first-best) deterrence. The reason is that, in such a regime, only efficient crimes are committed, so there would be no social gain from increasing the punishment on those offenders who commit them repeatedly. Doing so would be like charging a higher price for repeat customers.⁷ By the same logic, there would seem to be no social gain from incapacitating offenders who are expected to commit further efficient crimes once they are released.

The foregoing argument suggests that, in order to accept incapacitation as a basis for imprisoning offenders, one must either believe that some offenders are undeterrable and hence can only be prevented from committing inefficient crimes by detaining them, or that the optimal punishment policy involves some underdeterrence. Regarding the first of these possibilities, while it is likely that some offenders are in fact undeterrable, this does not provide a very satisfying answer to the compatibility question because it suggests that incapacitation can never be relevant for rational offenders. (It also requires the court to be able to distinguish deterrable and underdeterrable offenders at the time of sentencing.)

⁶ See, for example, Polinsky and Rubinfeld (1991), Polinsky and Shavell (1998), Emons (2004), and Miceli and Bucci (2005).

⁷ See Polinsky and Shavell (2000, p. 67) and Posner (2003, pp. 227-228).

Regarding the second possibility, it turns out that the BPS model does generally result in imperfect deterrence when punishment is costly (i.e., when it takes the form of prison). In particular, the socially optimal prison term as defined by equation (4) above may entail either overdeterrence or underdeterrence (Polinsky and Shavell, 2000, p. 50). In other words, the gain enjoyed by the marginal offender may be larger or smaller than the harm caused by the act (i.e., \hat{g} may be larger or smaller than h at the optimum). This is true because the optimal prison term must account for the expected cost of punishment, which depends both on the number of offenders punished and the length of the prison term. Thus, adjusting the prison term downward, for example, will reduce the cost per offender but will raise the number of offenders. And since the effect on expected costs is ambiguous, this may or may not be socially desirable.

The model to be developed in the next section exploits the fact that there is imperfect deterrence in the BPS model to develop a fully rational, dynamic model of law enforcement in which both deterrence and incapacitation arise naturally as rationales for possibly imprisoning convicted offenders.

4. Deterrence and Incapacitation: A Hybrid Model

This section lays out the hybrid model of criminal punishment. It first shows that the pure deterrence and pure incapacitation models emerge as special cases, and then examines the optimal enforcement policy in the general version.

4.1 The Basic Model

The model to be developed in this section retains the rational offender assumption, but extends the standard BPS model to make it dynamic. In particular, offenders are assumed to have infinite life spans and potentially to 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 an illegal act to the expected punishment. If an offender chooses to commit a crime, he does so continuously until he is apprehended. Then, after serving his prison term, he confronts another criminal opportunity immediately upon release and makes a new calculation. This sequence of crime and punishment repeats itself throughout the offender's infinite lifetime.

To model this formally, we again let g be the gain from a criminal act, which continues to be distributed across offenders by the density function $z(g)$. For simplicity, we will assume that the value of g that an offender draws at time zero (his first criminal opportunity) defines his “type” throughout the remainder of his life (i.e., whenever he encounters a criminal opportunity).⁸ Also, since time is continuous, we now define g to be the gain per instant of time up to the date when the offender is apprehended. Thus, if t is the date of apprehension, the present value of the gain from committing the initial crime, as of time zero (the commission date), is given by

$$\int_0^t g e^{-r\tau} d\tau = \frac{g}{r} (1 - e^{-rt}), \quad (5)$$

where r is the instantaneous discount rate. If we normalize the gain from not committing crimes to be zero, then expression (5) represents the gross gain from commission of a single criminal act as a function of the apprehension date, t .

⁸ This reflects the idea that the group of individuals who become criminals and those who refrain from criminal acts are distinct and remain constant over time. The results would be unaffected if individuals took a new draw of g at each criminal opportunity, in which case the identity of criminals would change over time.

We model the apprehension technology in a manner first suggested by Davis (1988) in his intertemporal model of crime.⁹ In particular, let the apprehension date, t , be a random variable that is distributed exponentially with density function

$$v(t) = pe^{-pt}, \quad (6)$$

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

Since we are interested in incapacitation, we focus on prison as the form of punishment, possibly combined with a fine. Thus, at the time of apprehension, the offender is assessed a fine (if any), denoted by f , and is imprisoned for a length of time s . Then, at the date of release, which occurs at time $t+s$, the offender immediately confronts another criminal opportunity, and the process begins again. Given the above assumption that the offender's type remains fixed throughout his life, and assuming a time-invariant punishment policy,¹⁰ the offender will make the same decision at each opportunity. Thus, those offenders who find crime profitable at time zero will become repeat (habitual) criminals, while those who are initially deterred will never commit crimes. Figure 2 depicts the time sequence of events as just described.

[Figure 2 here]

Following Polinsky and Shavell (2000, 2007), we first examine the optimal punishment scheme assuming a fixed probability of apprehension. We begin with a prison-only scheme, and then introduce the possibility of a fine combined with prison. In Section 5 below, we extend the model to allow an endogenous probability of apprehension.

4.2. *Punishment by Prison Only*

⁹ Loury (1979) and Mortensen (1982) use a similar approach to model the uncertain discovery of a technological innovation.

¹⁰ To keep the model simple, we do not consider enforcement policies that condition an offender's punishment on his offense history. For models that do allow such a policy, see the references in footnote 6 above.

We have already derived the gross gain for an individual who chooses to commit a crime. We now need to combine that with the expected punishment cost. In the prison-only scheme, this involves the expected present value of the cost of imprisonment from the date of apprehension, t , up to the date of release, $t+s$. Proceeding as above, we calculate the punishment cost for the offender's initial crime as a function of t and s to be

$$\int_t^{t+s} \delta e^{r\tau} d\tau = \frac{\delta}{r} (e^{-rt} - e^{-r(t+s)}), \quad (7)$$

where, recall, δ is the unit cost of imprisonment to the offender. The net benefit for the initial offense is thus given by the difference between (5) and (7), or

$$\frac{g}{r} (1 - e^{-rt}) - \frac{\delta}{r} (e^{-rt} - e^{-r(t+s)}). \quad (8)$$

Since the apprehension date is a random variable, we need to compute the expected value of this expression. Thus, weighting (8) by the density function in (2) and integrating over all values of t , we obtain

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

This expression represents the net expected gain to the offender from committing the first criminal act. If the opportunity were one-time, he would choose to commit the act if and only if this expression is positive, or if and only if

$$g \geq \frac{p\delta}{r} (1 - e^{-rs}), \quad (10)$$

where the right-hand side is the critical gain. As in the standard BPS model, this condition indicates that the offender is less likely to commit a crime as the length of the prison term and

the likelihood of apprehension are increased. The only difference here is that the cost of imprisonment is expressed in present value terms.¹¹

For purposes of integrating deterrence and incapacitation, the crucial extension of the standard model is the assumption that offenders face repeated criminal opportunities over an infinite horizon. This is captured in the above framework by computing the present value of lifetime net benefits, given the assumption of time invariance. Formally, this is done by forming the recursive equation

$$\Gamma(g, s, p) = G(g, s, p) + \beta(s)\Gamma(g, s, p), \quad (11)$$

where $\beta(s)$ is the expected discount factor. According to this expression, the offender expects to receive a net gain of G from every episode of crime and punishment over an infinite number of periods. The discount factor is in expected terms because it depends on the random date of apprehension. The expected value of this factor is thus computed as follows:

$$\beta(s) = \int_0^{\infty} e^{-r(t+s)} p e^{-pt} dt = \frac{p e^{-rs}}{p + r}. \quad (12)$$

Substituting this expression into (11) and solving for $\Gamma(g, s, p)$ yields:

$$\Gamma(g, s, p) = \frac{1}{p(1 - e^{-rs}) + r} \left[g - \frac{p\delta}{r} (1 - e^{-rs}) \right]. \quad (13)$$

A potential criminal at time zero will commit the initial crime if this expression, which represents the present value of his expected income from a life of crime, is positive. Note that the condition for (13) to be positive is given by (10), which was the condition for the first crime to be profitable. This makes sense since, given time invariance, if the first crime is profitable, all

¹¹ As a result, proportionally raising s and lowering p (or vice versa) will *not* leave deterrence unaffected, as was true in the BPS model (Davis, 1988).

subsequent crimes will be profitable as well. The threshold level of g separating criminals and non-criminals is thus given by the right-hand side of (10), or

$$\tilde{g}(s, p) \equiv \frac{p\delta}{r}(1 - e^{-rs}), \quad (14)$$

which, as noted, is decreasing in s , reflecting the deterrence function of prison.

Now consider the social cost of crime. This consists of three components, (1) the harm suffered by victims, (2) the cost to society of imprisoning offenders, and (3) the cost of apprehension. In the current version of the model where p is treated as a parameter, apprehension costs, $k(p)$, are fixed. The harm suffered by victims is again denoted by h , but, like the gain enjoyed by offenders, this cost is now measured per unit of time that the offender is free and committing crimes. Similarly, the cost of imprisonment, c , is defined to be the cost society incurs per unit of time that the offender is incarcerated. The harm plus imprisonment costs for the offender's first episode of crime and punishment are thus computed as follows:

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

Proceeding as above, we compute the expected value of this expression to obtain

$$\begin{aligned} C(s, p) &= \int_0^\infty \left[\frac{h}{r}(1 - e^{-rt}) + \frac{c}{r}(e^{-rt} - e^{-r(t+s)}) \right] p e^{-pt} dt \\ &= \frac{1}{p+r} \left[h + \frac{pc}{r}(1 - e^{-rs}) \right]. \end{aligned} \quad (15)$$

Finally, we can convert this to the present value of social costs over the lifetime of the offender by again using the recursive equation

$$\Psi(s, p) = C(s, p) + \beta(s)\Psi(s, p). \quad (16)$$

Substituting from (12) and solving yields the present value of harm plus punishment costs

$$\Psi(s, p) = \frac{1}{p(1 - e^{-rs}) + r} \left[h + \frac{pc}{r} (1 - e^{-rs}) \right]. \quad (17)$$

Total social costs consist of this expression plus the fixed costs of apprehension, or $\Psi(s, p) + k(p)$.

Before proceeding with the analysis of the hybrid model, we note that the pure deterrence (BPS) and pure incapacitation models emerge from this general model as special cases.

4.2.1. The Pure Deterrence Model. The pure deterrence model emerges from the above formulation by considering only the initial episode of crime and punishment. In other words, instead of assuming that the offender commits repeated crimes throughout his life, we assume his crime decision is one-time. Obviously, imprisonment cannot serve an incapacitation function in this case because there is no threat that the offender will commit further harmful acts. Thus, the only possible function of prison is to deter the initial crime.

Social welfare in this case consists of the net benefits to the offender from the first criminal act, given by (9), minus the corresponding social costs, given by (15), summed over those offenders who choose to commit the crime (i.e., those for whom $g \geq \tilde{g}$), minus the fixed cost of apprehension. The resulting welfare function is

$$\begin{aligned} W_d &= \int_{\tilde{g}(s, p)}^{\infty} [G(g, s, p) - C(s, p)] z(g) dg - k(p) \\ &= \int_{\tilde{g}(s, p)}^{\infty} \frac{1}{p + r} \left[g - h - \frac{p(c + \delta)}{r} (1 - e^{-rs}) \right] z(g) dg - k(p). \end{aligned} \quad (18)$$

Note that this expression corresponds closely to the welfare function for the BPS model in (2).

The resulting first order condition for the optimal prison term is given by

$$\left[h + \frac{pc}{r} (1 - e^{-rs}) \right] z(\tilde{g}) \delta = [1 - Z(\tilde{g})](c + \delta), \quad (19)$$

which has the same interpretation as (4) and only differs by the fact that the punishment cost savings on the left-hand side of (19) are in present value terms.

4.2.2. The Pure Incapacitation Model. The pure incapacitation model of Shavell (1987) also emerges from the above formulation by choosing the prison term that minimizes the present value of harm plus imprisonment costs, holding the crime rate fixed. In this case, the repeated-offense model is relevant. Thus, the problem is to choose s to minimize (17). The derivative of this expression with respect to s is given by

$$\frac{\partial \Psi}{\partial s} = \frac{pre^{-rs}(c-h)}{[p(1-e^{-rs})+r]^2}, \quad (20)$$

the sign of which depends on a comparison of c and h . If $c > h$, (20) is positive, implying that costs are increasing in the length of the prison term. In this case, the optimal prison term is zero; that is, offenders should face no prison time. Intuitively, if the cost of imprisonment exceeds the harm that the offender would impose on society if free, then it is inefficient to detain them. In contrast, if $c < h$, (20) is negative, implying that the optimal prison term is infinite. In this case, the cost of imprisonment is less than the harm that offenders would impose if free, so they should be imprisoned for life.

The simple model involves a corner solution (a zero or infinite prison sentence) because h and c are both assumed to be constant. More generally, if the offender's danger to society declines over time—either because a criminal's propensity to commit crime naturally declines with age, or because prison has a rehabilitative effect—then it becomes optimal to release the offender at the point where h falls below c (Shavell, 1987).

4.2.3. The Hybrid Model. Finally, consider the general model that encompasses both deterrence and incapacitation. Social welfare in this case consists of the present value of net gains to offenders over their infinite lifetimes, given by (13), minus the present value of social

costs, given by (17), both summed over those crimes that are committed, minus fixed apprehension costs. The resulting welfare function is given by

$$\begin{aligned}
W_h &= \int_{\tilde{g}(s,p)}^{\infty} [\Gamma(g,s,p) - \Psi(s,p)] z(g) dg - k(p) \\
&= \int_{\tilde{g}(s,p)}^{\infty} \frac{1}{p(1-e^{-rs}) + r} \left[g - h - \frac{p(c+\delta)}{r} (1-e^{-rs}) \right] z(g) dg - k(p). \quad (21)
\end{aligned}$$

The optimal prison term maximizes this expression. The relevant first order condition, assuming an interior solution, is given by

$$\left[h + \frac{pc}{r} (1-e^{-rs}) \right] z(\tilde{g}) \delta = \frac{1}{p(1-e^{-rs}) + r} \int_{\tilde{g}(s,p)}^{\infty} (c + \delta + g - h) z(g) dg. \quad (22)$$

The left-hand side of this condition is identical to the left-hand side of (19) and again represents the marginal deterrence benefit of increasing the length of the prison term. However, the right-hand side of (22), the marginal cost of a longer prison sentence, is different. As in the pure deterrence model, it includes the marginal cost to society and to the offender of a longer sentence, captured by the $c+\delta$ term in the integral, appropriately adjusted to reflect the repeated nature of crime and punishments. But in addition, the marginal cost of punishment includes a term to reflect the incapacitation effects of imprisonment. Specifically, the $g-h$ term in the integral represents the foregone net social benefits of those crimes that the offender is unable to commit because he is detained in prison for a longer period of time, conditional on the fact that he would continue to commit crimes if set free, given $g \geq \tilde{g}$. Notice, however, that this incapacitation term may be positive or negative, depending on whether the expected value of g for those offenders who find crime profitable is larger or smaller than h . If it is positive, the marginal cost of imprisonment is larger compared to the pure deterrence model, implying that the optimal prison term should be shorter. In contrast, if it is negative, the marginal cost of

imprisonment is smaller compared to the pure deterrence model, implying that the optimal prison term should be longer.

The intuitive explanation for these results is as follows. Suppose that the prison sentence is initially set at the length that would be optimal from a pure deterrence perspective. Then, if the expected gain for offenders who commit crimes is less than the social harm that they cause, then on average they are committing *inefficient crimes*. That is, $E[g - h \mid g \geq \tilde{g}] < 0$. In this case, there is a net social gain from detaining them longer in prison to prevent them from committing additional crimes at the margin. In other words, incapacitation dictates that the optimal prison sentence should be longer than would be indicated by a pure deterrence model. In contrast, if the expected gain for those offenders who commit crimes is larger than the harm they impose (i.e., if $E[g - h \mid g \geq \tilde{g}] > 0$), then on average they are committing *efficient crimes*. In this case, there would be a social loss from detaining them for a longer period of time because it deprives them of the opportunity to commit further efficient crimes. In other words, holding offenders in prison for purposes of incapacitation actually results in a net social loss. The optimal prison term is therefore shorter than would be prescribed under a policy of pure deterrence.

As a general rule, it is not possible to tell from (22) whether crimes are efficient or inefficient at the optimum. All we can say is that the right-hand side of (22) must be positive, which requires that

$$c + \delta > h - E[g \mid g \geq \tilde{g}]. \quad (23)$$

Note that this condition necessarily holds if crimes are on average efficient (i.e., if the right-hand side is negative), but it can also hold if crimes are inefficient (i.e., if the right-hand side is positive). Thus, we can only conclude that incapacitation can either raise or lower the optimal prison term compared to a regime based on deterrence alone.

This possibility that incapacitation can actually lower the optimal prison term is obviously a consequence of the assumption that the offender's benefit from crime counts in social welfare. Given the controversial nature of this assumption (especially for violent crimes), Section 5.2 will therefore examine the implications of relaxing it.

4.3. Prison and Fines

In the prison-only model analyzed in the previous section, the optimal prison term had to balance deterrence and incapacitation. We now extend the model to allow the use of fines along with prison. Fines obviously can have no incapacitation effect, but they can deter offenders, thus allowing the use of prison solely for incapacitation purposes. We noted above that in the pure deterrence (BPS) model, when fines and prison are both available, fines should be maximal (i.e., equal to the offender's wealth) before the use of prison is considered. The question is whether this conclusion continues to hold in the hybrid model. To provide an answer, we assume initially that there is no limit on the offender's ability to pay the fine.

Assume that the fine, f , is imposed as a lump sum amount at the instant the offender is apprehended. For the initial crime, the present value of the fine as of time zero is therefore $e^{-rt}f$, which is subtracted from the offender's net benefit in (8). Calculating the expected value of this expression as above yields

$$G(g, f, s, p) = \frac{1}{p+r} \left[g - pf - \frac{p\delta}{r} (1 - e^{-rs}) \right]. \quad (24)$$

Converting this to the present value of gains over the offender's infinite lifetime yields

$$\Gamma(g, f, s, p) = \frac{1}{p(1 - e^{-rs}) + r} \left[g - pf - \frac{p\delta}{r} (1 - e^{-rs}) \right]. \quad (25)$$

As in the prison-only model, the offender commits the first crime if and only if this expression is positive, or if and only if

$$g \geq pf + \frac{p\delta}{r}(1 - e^{-rs}) \equiv \tilde{g}(f, s, p), \quad (26)$$

which differs from the threshold in (14) by the addition of the expected fine, pf , on the right-hand side.

The social cost of crime also needs to be amended to account for the fine revenue received by the government. Proceeding as above, we obtain the following expression for the present value of expected social costs over the offender's lifetime:

$$\Psi(f, s, p) = \frac{1}{p(1 - e^{-rs}) + r} \left[h + \frac{pc}{r}(1 - e^{-rs}) - pf \right]. \quad (27)$$

Finally, we form the social welfare function by subtracting social costs in (27) from offender gains in (25), integrating over the set of offenders who commit crimes, and subtracting fixed apprehension costs:

$$W_h^f = \int_{\tilde{g}(f, s, p)}^{\infty} \frac{1}{p(1 - e^{-rs}) + r} \left[g - h - \frac{p(c + \delta)}{r}(1 - e^{-rs}) \right] z(g) dg - k(p). \quad (28)$$

Note that the fine revenue drops out of this expression since it is simply a transfer payment.

Thus, welfare in this case differs from the expression in the prison-only case in (21) only by the lower limit of the integration, which here depends on the expected fine as well as the prison term. This reflects the fact, noted above, that the fine only affects deterrence (i.e., the number of crimes committed).

Consider first the optimal fine, which is found by maximizing (28) with respect to f .

Under the assumption that there is no limit on the offender's ability to pay, we obtain

$$f^* = h/p + \frac{c}{r}(1 - e^{-rs}), \quad (29)$$

which says that the optimal fine equals the harm suffered by victims, appropriately inflated to reflect uncertain apprehension, plus the present value of imprisonment costs that the offender

imposes on society. In the special case where there is no imprisonment, this expression reduces to the optimal fine in the fine-only version of the BPS model, as shown in (3). If deterrence were the only consideration, there would be no reason to impose a prison term since the fine achieves perfect deterrence. The question is whether there is a role for prison in the hybrid model for purposes of incapacitation.

To answer this question, we set $f=f^*$ in (28) and take the derivative with respect to s . The result is

$$\frac{\partial W_h^f}{\partial s} \Big|_{f=f^*} = \frac{pre^{-rs}}{[p(1-e^{-rs})+r]^2} \int_{\tilde{g}(f^*,s,p)}^{\infty} (-g - \delta - c + h)z(g)dg, \quad (30)$$

where

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

Now evaluate this derivative at $s=0$:

$$\frac{\partial W_h^f}{\partial s} \Big|_{f=f^*,s=0} = \frac{p}{r} \int_h^{\infty} (-g - \delta - c + h)z(g)dg < 0, \quad (31)$$

where the sign follows from the fact that the integration is over the range where $g \geq h$, which implies that the entire term inside the integral must be negative. It follows that $s^*=0$; that is, no prison term should be imposed. Since the fine is unconstrained by the offender's wealth, it can be set to achieve the efficient (first-best) level of deterrence. Thus, although offenders will continue to commit crimes continuously throughout their lifetimes (since they are never imprisoned), those crimes are socially efficient, so there is no social gain from incapacitating them.

Note that this conclusion is consistent with the results in the previous section, where the gain from incapacitation (if any) arose from the possibility of underdeterrence when prison was

the only available sanction. In other words, because prison alone cannot generally achieve first-best deterrence due to the cost of punishment, there is a potential gain from adjusting the prison term for purposes of incapacitation when criminals are known to be repeat offenders. In contrast, when the fine can be adjusted at no cost to achieve perfect deterrence, there is no gain from incapacitation.

Of course, this conclusion would be different if the fine were constrained by the offender's wealth. In that case, it is easy to show that the optimal fine is maximal, or $f=w$ (as in the BPS model), and prison may now be desirable for purposes of both deterrence and incapacitation. The trade-off is identical to that in the prison-only model. That is, a positive prison term is optimal if the expected deterrence benefits exceed the expected punishment costs.

5. Extensions of the Model

This section examines two extensions of the above model. First, we consider the case where the probability of apprehension is endogenous, and second, we consider the implications of not counting the offender's gains as part of social welfare.

5.1. Endogenous Probability of Apprehension

Consider first the case where prison is the only available sanction. Recall that in the BPS model, the optimal prison term is maximal in this case. It turns out that this result continues to hold in the hybrid model. To see why, consider the welfare function in (21) and suppose initially that s is less than maximal. Now raise s and lower p so as to hold the term $p(1-e^{-rs})$ constant. Since the integral term is unchanged but apprehension costs fall, welfare must rise, implying that welfare could not have been maximized under the initial policy. Thus, for any $p>0$, welfare cannot be maximized if s is less than maximal. The fact that offenders (potentially) commit an

infinite number of crimes in the hybrid model does not affect this conclusion because each crime is an exact replay of the previous one, and the optimal policy with respect to the first crime remains the optimal policy throughout time. Besides, a maximal prison term implies that the offender would be imprisoned for life on the first offense, so he would have no opportunity to commit further crimes.

When the threatened prison term is infinite, the threshold gain in (26) reduces to $\tilde{g} = p\delta/r$. Thus, the threat of life imprisonment does not generally result in complete deterrence in the hybrid model. In this case, the welfare function in (21) becomes

$$W_h = \int_{p\delta/r}^{\infty} \frac{1}{p+r} \left[g - h - \frac{p(c+\delta)}{r} \right] z(g) dg - k(p). \quad (32)$$

The optimal apprehension rate is found by maximizing this expression with respect to p .

Assuming an interior solution, we obtain the following first order condition

$$\frac{1}{p+r} \left(h + \frac{pc}{r} \right) z(\tilde{g}) \frac{\delta}{r} = k'(p) + \frac{1}{(p+r)^2} \int_{p\delta/r}^{\infty} (c + \delta + g - h) z(g) dg. \quad (33)$$

The left-hand side is the marginal deterrence benefit of a higher apprehension rate in the form of the reduced harm to victims and saved punishment costs. The right-hand side is the marginal cost of increasing p (the cost of hiring more police officers, for example), plus the increased punishment costs incurred as more offenders are caught and imprisoned. Note that the marginal punishment cost term (the second term on the right-hand side) includes the incapacitation effect described above (represented by the $g-h$ term in the integral), reflecting the foregone net gains from those crimes that offenders are unable to commit due to the higher apprehension rate. As before, this may be positive or negative at the optimum, and so may increase or decrease the marginal cost of raising p .

Finally, consider the case where both fines and prison are available when p is endogenous. As was true in the case where p was fixed, it is never optimal to use prison unless the fine is first set at its maximal level. Thus, after setting $f=w$, the prison term and probability of apprehension are chosen simultaneously to maximize welfare in (28). In this case, the prison term is not necessarily maximal. To see why this is true, suppose that s is less than maximal, and then proceed as above to raise s and lower p so as to hold $p(1-e^{-rs})$ fixed. In this case, apprehension costs fall, but so does deterrence because the expected fine (given by pw) falls. Thus, welfare is not necessarily increased. As usual, the desirability of imposing a prison term in this case depends on the marginal deterrence benefits compared to the marginal cost.

The basic conclusions in this section are qualitatively similar to those in the standard BPS model. They differ only by the inclusion of the incapacitation effect in the marginal cost of imprisonment, as discussed above.

5.2. The Effect of Excluding the Offender's Gain from Social Welfare

To this point we have maintained the standard practice of counting the offender's gain as a component of social welfare. Early on, however, Stigler (1970, p. 527) questioned the propriety of this practice when he asked, "What evidence is there that society sets a positive value upon the utility derived from murder, rape, or arson? In fact, the society had branded the utility from such activities as illicit." But the issue is not a simple one since some acts that society labels as "crimes" can yield benefits to the offender that most people would consider a valid component of social welfare. Consider, for example, a man who exceeds the speed limit to get his pregnant wife to the hospital, or a lost hiker who breaks into a cabin for food and shelter. Further, as Friedman (2000, p. 230) observes, once we start sorting criminals into "the deserving and the undeserving," we make the error of "assuming our conclusions" about the appropriate

treatment of criminals. For these reasons, economists have for the most part retained the standard assumption of counting the offender's utility in welfare.

Still, it is almost certainly the case that for those offenses where incapacitation is a relevant consideration, like dangerous crimes, Stigler's point is a valid one. Thus, in order to get a true sense of the interaction between deterrence and incapacitation, it would seem worthwhile to consider a version of the above model in which the offender's gain is excluded from welfare. For this purpose, it is sufficient to focus on the prison-only version of the model and to assume a fixed probability of apprehension. (Thus, in this section we ignore the fixed cost of apprehension.)

The measure of social welfare in this case simply consists of the harm suffered by victims plus the present value of expected punishment costs, summed over the range of offenders.¹² The relevant cost expression is thus given by (17), integrated over $g \geq \tilde{g}(s)$, where $\tilde{g}(s)$ is defined by (14). Thus, the optimal prison term in this case is chosen to minimize the following cost expression:

$$SC = \int_{\tilde{g}(s)}^{\infty} \frac{1}{p(1 - e^{-rs}) + r} \left[h + \frac{pc}{r}(1 - e^{-rs}) \right] z(g) dg. \quad (34)$$

Note that this choice problem differs from that in the pure incapacitation model described in Section 4.2.2 only by the endogeneity of the crime rate, as embodied by the threshold gain, $\tilde{g}(s)$.

The derivative of (34) with respect to s is given by

$$\frac{\partial SC}{\partial s} = \frac{-rpe^{-rs}}{p(1 - e^{-rs}) + r} \left\{ \left[h + \frac{pc}{r}(1 - e^{-rs}) \right] z(\tilde{g}) \frac{\delta}{r} + \frac{1}{p(1 - e^{-rs}) + r} \int_{\tilde{g}(s)}^{\infty} (h - c) z(g) dg \right\}. \quad (35)$$

¹² For consistency, we do not count the offender's disutility from imprisonment as part of punishment costs.

Note that the first term in braces is identical to the left-hand side of (22) and again represents the marginal deterrence benefit of increasing the prison term. However, the second term in braces, the marginal cost of increasing s , is different from the right-hand side of (22). Note in particular that it may be positive or negative, depending on the relationship between the harm suffered by victims, h , and the cost of imprisonment, c . Thus, this term reflects the pure net benefit (cost) of incapacitation.

Suppose initially that $h < c$, or that the cost of imprisonment exceeds the harm from crime. In this case, the second term in (35) is negative, meaning that increasing the prison term imposes a net cost on society. The optimal prison term will therefore will occur at the point where the derivative in (35) equals zero, or where the marginal deterrence benefit equals the marginal incapacitation cost. The resulting prison term will therefore generally be of finite length. Although imprisonment is undesirable from a pure incapacitation perspective because the cost of holding the offender in prison exceeds the harm that he would impose if free, it is still socially desirable to impose some prison time on offenders because of the deterrence benefits.¹³ The optimal prison term in this case thus represents a trade-off between deterrence and incapacitation.

Suppose in contrast that $h > c$, or that the harm caused by the offender exceeds the cost of holding him in prison. In this case, the second term in (35) is positive, implying that the entire derivative is negative. Thus, social costs are strictly decreasing in s . As a result, the optimal prison term is infinite (maximal). (Note that there is no possibility of “overdeterrence” in this case because all crimes are assumed to be inefficient.) In this case, incapacitation and deterrence reinforce each other and indicate that the prison term should be as long as possible.

¹³ Even if fines were available, some prison time might still be desirable, given that optimal deterrence in this case is complete deterrence. Thus, any finite level of wealth would be a binding constraint.

In terms of policy relevance, this last outcome seems to be most descriptive of the prototypical case where deterrence and incapacitation serve as complementary reasons for imprisoning dangerous offenders. Specifically, the threat of imprisonment deters some offenders from committing dangerous crimes in the first place, while those offenders who reveal their predilection to commit crimes in spite of the threatened punishment should be imprisoned for life on their first apprehension in order to prevent them from having further criminal opportunities. This logic seems to be the motivation underlying three strikes laws, though the current model, with its assumption of a time-invariant policy, does not account for the gradual progression in such policies toward a maximal prison term. Indeed, the current model with fully rational offenders provides no basis at all for waiting until the third (or even the second) offense to impose the maximal sentence. Explaining this provision of the law therefore requires further elaboration of the basic model.

6. Conclusion

The economic theory of law enforcement has traditionally focused on deterrence as the primary motivation for criminal punishment. Since fines and imprisonment are equally capable of deterring crime under this theory, the model prescribes that prison should never be used unless the limited wealth of offenders prevents the attainment of the desired level of deterrence. To the extent that the actual use of prison seems to be more extensive than is warranted by this prescription, however, the economic model falls short as a positive theory of criminal policy. In addition, the economic model offers no rationale for punishing offenders who are undeterrable. The theory of incapacitation, on the other hand, addresses both of these shortcomings, but it offers no theory of criminal behavior. Unfortunately, the law and economics literature has yet to

find a way to incorporate these two theories into a coherent model. Filling that gap has been the goal of this paper.

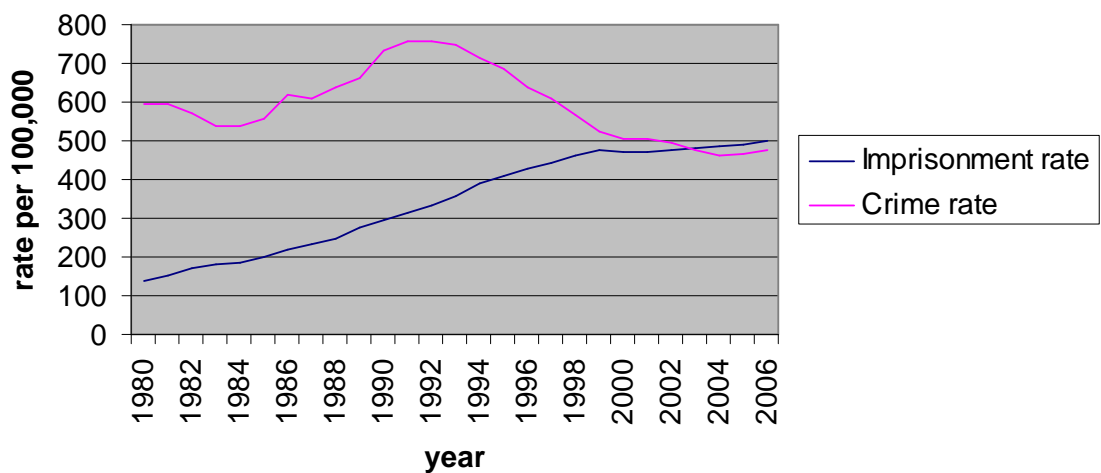
The hybrid model outlined herein showed that, in a dynamic setting where fully rational offenders face recurrent criminal opportunities throughout their lifetimes, deterrence and incapacitation emerge naturally as complementary motives for criminal punishment. In particular, the threat of imprisonment (and/or a fine) deters some potential offenders from ever committing crimes, while the actual imposition of a prison sentence on convicted offenders prevents them from committing further, inefficient crimes by detaining them in jail. The optimal prison sentence thus embodies both approaches to harm prevention. In most respects, the hybrid model does not prescribe fundamentally different policies as compared to the pure deterrence model. Still, by combining the two motives for punishment into a coherent framework, the hybrid model provides a more compelling theory of actual punishment policies without having to abandon the theoretical appeal of the standard economic model.

References

- Beccaria, Cesare (1767) [1995] *On Crimes and Punishments, and Other Writings*, Cambridge: Cambridge Univ. Press (R. Bellamy, ed).
- Becker, Gary (1968) "Crime and Punishment: An Economic Approach," *Journal of Political Economy* 76: 169-217.
- Bentham, Jeremy (1789) [1973] *An Introduction to the Principles of Morals and Legislation*, in *The Utilitarians*, Garden City, N.J.: Anchor Books.
- 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.
- Friedman, David (2000) *Law's Order: What Economics has to do with Law and Why it Matters*, Princeton, NJ: Princeton Univ. Press.
- Kessler, Daniel, and Steven Levitt (1999) "Using Sentence Enhancements to Distinguish Between Deterrence and Incapacitation," *Journal of Law and Economics* 17: 343-363.
- 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.
- Lott, John (1987) "Should the Wealthy be Able to Buy Justice?" *Journal of Political Economy* 95: 1307-1316.
- 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.
- Posner, Richard (2003) *Economic Analysis of Law*, 6th Edition, New York: Aspen Law and Business.
- 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.
- Stigler, George (1970) "The Optimum Enforcement of Laws," *Journal of Political Economy* 78: 526-536.

Figure 1: Crime and imprisonment rates (per 100,000), 1980-2006



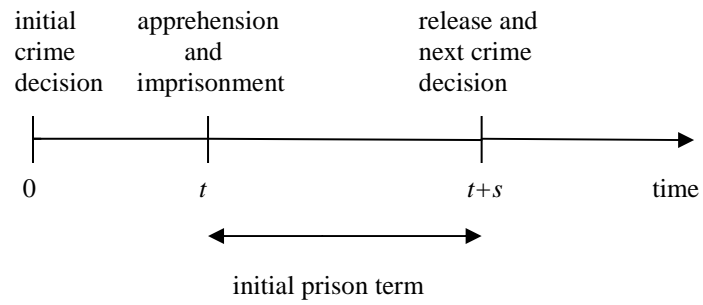


Figure 2. Time line of crime and punishment over an infinite horizon.