How to Make the Fine Fit the Corporate Crime? An Analysis of Static and Dynamic Optimal Punishment Theories

Leung, Siu Fai

Working Paper No. 261
January 1991

How to Make the Fine Fit the Corporate Crime? An Analysis of Static and Dynamic Optimal Punishment Theories

Sui Fai Leung

Rochester Center for Economic Research
Working Paper No. 261

March 1990
Revised December 1990

How to Make the Fine Fit the Corporate Crime? An Analysis of Static and Dynamic Optimal Punishment Theories

Siu Fai Leung
University of Rochester

Rochester Center for Economic Research
Working Paper No. 261

Abstract:

This paper investigates dynamic optimal punishment theory. Previous studies of optimal punishment theory have all been confined to a static setup. The most celebrated result in the literature is Becker's (1968) optimal punishment theory which states that the optimal fine should be a multiple of the social cost of the crime. I show that this result is no longer valid in a dynamic environment. The implications of the dynamic model of optimal punishment are found to be strikingly different from those of the static model.

First Draft: March 1990
Current Version: December 1990

I am grateful to the referees for their comments and suggestions. This research was supported in part by a faculty research grant from the College of Arts and Science of the University of Rochester.

Address: Department of Economics, Harkness Hall, University of Rochester, Rochester, New York 14627, U.S.A.

1. Introduction

"How to make the punishment fit the crime?" is an ancient question still widely debated in many different disciplines of studies. While this ancient question remains pertinent in modern times, the frequency and the scope of crimes have continued to increase and expand. Perhaps the most recent and noticeable one is corporate crime (or business/white-collar /organizational crime). The emergence of corporate crimes has attracted considerable attention because they appear to be widespread and often involve huge amounts of money. For examples, insider tradings, bank frauds, large oil spills, and safety regulations violations frequently occupy news headlines. The $650 million fine on Drexel Burnham Lambert Inc. in a security fraud case in the United States in 1989 illustrates the large amount of money involved.

One of the most influential arguments in optimal punishment theory is the economic approach put forth by Becker (1968) who argued that social welfare will be increased if fines are used whenever feasible. Suppose corporate offenders are punished by fines rather than by imprisonments, how should the optimal fine be set? Becker (1968, p.199) suggested that "If compensation were stressed, the main purpose of legal proceedings would be to levy fines equal to the harm inflicted on society by constraints of trade." He then elaborated the argument in a footnote (p.199, fn.55): "Actually, fines should exceed the harm done if the probability of conviction were less than unity. The possibility of avoiding conviction is the intellectual justification for punitive, such as triple, damages against those convicted." Posner (1977, p.171) spelled out the principle more explicitly: "The optimum penalty is simply the social cost of the unlawful act divided by the desired probability that the penalty will in fact be imposed." Recently, Becker (1989) reiterated his principle: "The penalty should be a multiple of, rather than simply equal to, the harm caused. That is because many violators of the law go unpunished. If companies that are convicted pay only for the damages they cause, the deterrent effect of fines would be too weak, because companies that are not caught pay nothing." To give an example, suppose an oil tanker leaked oil and damaged the environment. If the social cost was $0.1 per gallon of oil leaked and the probability of conviction was 1/10, then the optimal fine should be $1 per gallon of oil leaked.

Becker's approach offers a simple way to set the optimal fine. It has been applied to explain treble and punitive damage awards [e.g., Becker (1968), Posner (1976), Ellis (1982)] because the optimal fine is a multiple of the damages caused. Many legal scholars and economists have advocated using Becker's approach to sanction corporate offenders [e.g., Posner (1980), Farber (1980), Levinson (1980), Ellis (1982)]. Twenty years after Becker proposed the approach, it began to gravitate toward the public policy-makers' circle in the United States. The approach was endorsed in a July 1988 discussion draft on business crime prepared by some members of the U.S. Sentencing Commission (Becker 1989). The subsequent preliminary draft on sentencing guidelines for organizational defendants (U.S. Sentencing Commission 1989) revealed the influence of Becker's approach on the Commission's proposals. The proposed guidelines have generated some serious concerns from the business community. Critics often argue that Becker's approach is unacceptable because the multiple fine imposes too heavy penalty on corporate offenders [Smart and Galen (1990), Mandel (1990)]. The debates illustrate the subtleties involved when economic theory is applied to important public policy-making [Polinsky and Shavell (1990)].

Despite the increasing popularity of Becker's approach, several critical issues have remained neglected by both the proponents and the opponents of the approach. The most notable one is the issue of dynamics. Becker's original analysis and the work that follows [e.g., Polinsky and Shavell (1984), Craswell and Calfee (1986)] have been confined to a static environment. All the models on the determination of the optimal fine in the literature are static one-period models. Although some work has been done on modeling dynamic criminal behavior,[1] none has rigorously investigated dynamic optimal punishment theory.[2] In this paper, I analyze how to set the optimal fine in a dynamic setting. The most striking finding of the analysis is that Becker's claim that the optimal fine should be a multiple of the social cost of the crime cannot be generalized to a dynamic environment. This casts serious doubts on whether the sentencing guidelines for corporate offenders proposed by the U.S. Sentencing Commission will generate optimal outcomes. The dynamic model examined in the paper reveals some important and subtle aspects of criminal behavior which have gone unnoticed in previous studies. It shows that setting the optimal fine in a general environment is a much more complicated task than that suggested by the static approach.

The plan of the rest of the paper is as follows. Section 2 briefly reviews Becker's model and then contrasts it with another static model.

---

[1] The earliest one is the dynamic income tax evasion model in Allingham and Sandmo (1972).

[2] Two exceptions that appear recently are Nash (1989) and Polinsky and Rubinfeld (1989). However, the approaches and the focuses of these two papers are different from those of the present paper.

Section 3 formulates a dynamic model of criminal behavior and derives the optimal fine from the model. The optimal fine is shown to be substantially different from the one derived from Becker's model. A detailed analysis on the distinctions between the static and the dynamic models is provided. Section 4 concludes the paper.

2. Two Static Models

In this section, I first present Becker's approach through a simple static model (S1) and then compare it with another static model (S2). The purposes of comparing the two static models are threefold. First, it reveals that Becker's approach is based on a restrictive assumption which has never been noticed in the literature. By making a different assumption, model (S2) shows that Becker's claim will no longer hold even in a static setup. Second, there are some shortcomings in model (S2) which serve to motivate the need for a dynamic analysis. Third, model (S2) captures some important features of the dynamic model and hence facilitates the comparisons between static and dynamic optimal punishment theories.

2.1. Model (S1)

Suppose an offender's returns from engaging in some criminal activity are given by $B(c)$, where c is the offense rate (crime rate). Without loss of generality, assume $c \in [0,1]$. A fine $\theta(c)$ will be levied if the offender is convicted. Let P be the probability of conviction, then the offender's decision problem is to

$$\text{(S1)} \qquad \max_{c} \, [B(c) - P\theta(c)]. \qquad (1)$$

Let $C(c)$ be the social cost (damage) of the crime and assume that $B(c)$ is also the social benefit of the crime, then the social planner's problem is to

$$\max_{c} [B(c) - C(c)]. \tag{2}$$

It follows that the social optimal offense rate is given by

$$c^* = \text{argmax} [B(c) - C(c)]. \tag{3}$$

Comparing (1) and (2), it is clear that the offender will choose c^* if the fine is set according to

$$\theta(c) = C(c)/P, \tag{4}$$

because the offender's objective function $B(c) - P[C(c)/P]$ will then be identical to the social planner's objective function $B(c) - C(c)$. Thus, the optimal fine is not just the social cost $C(c)$; a multiplier $1/P$ has to be included. Since P is usually strictly less than one, therefore the multiplier is always strictly greater than one. Hence, the optimal fine is a <u>multiple</u> of, instead of equal to, the damages done on the society. This multiple fine is the core of the debates between the proponents and the opponents of Becker's approach to optimal punishment.

Becker's approach can be conveniently labeled as the "static multiplier approach" since it is based on a static model and the multiplier is the most important and distinctive implication derived from the model. Model (S1) presents a simple and reasonable way to formally describe Becker's approach. Notice that enforcement costs are ignored here. It is easy to bring enforcement costs and other considerations (e.g. uncertainty) into the analysis, but the model may no longer be able to generate Becker's claim that the optimal fine is the social cost divided by the probability of conviction. In fact, Stern (1978), Polinsky and Shavell (1984), Craswell and Calfee (1986), and others have already evaluated the robustness of the results in Becker (1968) in many different ways. By focusing on dynamics (instead of the static setup in the literature), this paper takes another route to evaluate Becker's claim. To accomplish this goal, it is necessary to assume that the above static model is valid and use it as a starting point to examine whether the result on the multiplier can be carried over to a dynamic setup. Accordingly, the criticisms which previous studies have made that may alter Becker's results in the static setup will not be considered in the paper.

2.2. Model (S2)

Instead of (1), suppose the offender's decision problem is to

$$\text{(S2)} \qquad \max_{c} \; (1-P)B(c) - P\theta(c). \tag{5}$$

There are some crucial differences between models (S1) and (S2). Model (S1) assumes that the offender commits the crime and enjoys the benefits before he faces the possibility of punishment. The gains B(c) are guaranteed. The offender does not have to pay the fine if he is not arrested. He will gain B(c) with certainty and lose θ(c) with a probability of P. In model (S2), the gains B(c) are not guaranteed. The offender will gain B(c) with probability (1-P) and lose θ(c) with probability P. Put differently, the offender can be caught in the act so that he has to pay the fine and gains nothing at all. In contrast, model (S1) assumes that the offender will never be caught in the act.[3]

---

[3] As in footnote 17 in Becker (1968), one may interpret model (S1) by writing the objective function as $(1-P)B(c) + P[B(c)-\theta(c)]$, which allows for the possibility that the offender may be caught in the act. This implies that the offender will gain B(c) with probability 1-P and $B(c)-\theta(c)$ with probability P. Clearly, this interpretation is unappealing because it essentially assumes that if the offender is caught in the act, the law

What is the optimal fine in model (S2)? It is easy to see that the optimal fine is given by $\theta(c) = [(1-P)/P]C(c)$. Hence, the optimal multiplier is $(1-P)/P$, which is smaller than the optimal multiplier $1/P$ of model (S1). More importantly, the multiplier $(1-P)/P$ may be less than one, because $(1-P)/P > 1$ if $P < 1/2$, and $(1-P)/P < 1$ if $P > 1/2$. Therefore, the optimal fine in model (S2) may be a fraction, not a multiple, of the social cost of the crime. If the multiplier $1/P$ is applied to model (S2), then the sanction on the offender will be too heavy. As a result, the offender will overcomply, and the social optimum cannot be attained.[4]

One may criticize model (S2) because of the unrealistic assumption that the returns from the crime will all be lost if the offender is caught in the act. This can be dealt with by making an additional assumption that only a portion of the returns is lost. Let α $(0 \leq \alpha < 1)$ be the proportion of $B(c)$ that the offender has already reaped when he is caught, then the expected returns become $(1-P)B(c) + P[\alpha B(c)-\theta(c)]$. Clearly, the optimal multiplier is $(1-P+\alpha P)/P$, which again is smaller than the multiplier $1/P$ and is not necessarily greater than one. One can further relax the restriction that α is a constant by letting it be a random variable with distribution function $G(\alpha)$.[5] In this case, let $\bar{\alpha} = \int_0^1 \alpha dG(\alpha)$ be the mean of α, then the expected returns become $(1-P)B(c) + P[\bar{\alpha}B(c)-\theta(c)]$, i.e., the α in the previous case

---

enforcement agency will still pay $B(c)$ to him and fine him $\theta(c)$. In any case, the amount $B(c)$ is always guaranteed.

[4] The proof is straightforward. If the multiplier $1/P$ is applied to model (S2), the decision problem becomes $\text{Max}_c\ [(1-P)B(c) - C(c)]$, and the first-order condition is $(1-P)B'(c) = C'(c)$. With the standard assumptions $B''(c) < 0$ and $C''(c) > 0$, this implies $c < c^*$, where c^* is the social optimal offense rate defined by (3). Therefore, the offender will overcomply.

[5] This case was first analyzed by Heineke (1975). However, the focus of his analysis is different from the present one.

is now replaced by the mean $\bar{\alpha}$. Again the multiplier $(1-P+\bar{\alpha}P)/P$ is smaller than the multiplier $1/P$ and may be smaller than one. Notice that neither α nor $\bar{\alpha}$ in this type of one-period models is a fully specified quantity in the sense that they are not determined within the model. Any value of α or $\bar{\alpha}$ assumed is clearly arbitrary. These shortcomings can be removed in a dynamic analysis. It will be shown in the next section that the dynamic model offers a clear structural determination of α (or $\bar{\alpha}$) within the model and therefore does not suffer from the arbitrariness.

In reality, which model fits better the arrest process of a law enforcement agency? It seems clear that for most criminal activities, the law enforcement agency will not deliberately wait until the offender has completed the crime. This is particularly true for those criminal activities in which the flow of returns occurs over a period of time, so that there is a positive probability that the offender will be caught in the act. For example, speeding in highways is usually caught in the act; the gain from driving fast is lost and the fine is based on the speed of the automobile at the time the offender is caught. One may argue that in the case of pollution, a firm usually pollutes first and then faces the possibility of punishment. Hence, the firm has already reaped the gains from the pollution. It is easy to cite other crimes in which the offenders have enjoyed the gains from the crimes long before they are arrested and convicted. But the argument should not be based on the ex-post outcomes. The crucial point is ex-ante: whether the offender considers the possibility of being caught in the act in his decision. It appears that offenders do take this possibility into account. As long as an offender presumes that there is a non-zero probability (no matter how small) that he may be caught in the act, then model (S2) will be the right model to analyze.[6] This consideration points to the need for a rigorous dynamic analysis because the possibility of "caught in the act" clearly rests on the premise that the environment is dynamic.

3. A Dynamic Model

Instead of facing a one-period decision problem, the offender now has to choose a path of offense rate (crime rate) over an infinite horizon.[7] At any time t, an increase in the offense rate $c(t)$ ($c(t) \in [0,1]$) will not only increase the returns $\pi_1(c(t))$ from the illegal activity but also the hazard rate of arrest $h(c(t))$. If the offender is arrested, he will be convicted and has to pay a fine of $\theta(c(t))$, which depends on the offense rate at the time the illegal activity is detected.[8] After the offender is convicted, he can no longer participate in any illegal activity and will earn π_2 (a constant) from some legal activity thereafter. Therefore, the offender faces the following decision problem:

(D1)

$$\max_{c(.)} J = \int_0^\infty e^{-rt} \left\{ \pi_1(c(t))[1-F(t)] + \left[-\theta(c(t)) + \int_t^\infty e^{-r(s-t)} \pi_2 ds \right] f(t) \right\} dt \tag{6}$$

$$\text{subject to} \qquad f(t) = h(c(t))[1-F(t)], \tag{7}$$

---

6 In model (S1), if the law enforcement agency forces the offender to return the benefits from crime, then the total fine is $B(c)+C(c)$ and the objective function becomes $B(c) - P[B(c)+C(c)] = (1-P)B(c) - PC(c)$. Hence, if restitution is a component of the penalty, then model (S1) will become model (S2), and the optimal multiplier is $(1-P)/P$ (not $1/P$).

7 The assumption of an infinite horizon is justifiable in the present context since the offender is a corporation and not an individual.

8 For analytical convenience, detection, arrest, and conviction are assumed to happen at the same time. This assumption can be relaxed and the main results of the analysis will not be affected.

where r is the discount rate, f(t) and F(t) are respectively the density function and the distribution function of the time of arrest. The constraint (7) follows from the definition of the hazard function.[9] Problem (D1) is an infinite horizon optimal control problem with control variable c(.) and state variable F(.).[10] The objective function J can be simplified as follows. By Fubini's theorem,

$$\int_0^\infty e^{-rt}\left[\int_t^\infty e^{-r(s-t)}\pi_2 ds\right]f(t)dt = \int_0^\infty e^{-rt}\pi_2 F(t)dt$$
$$= -\int_0^\infty e^{-rt}\pi_2[1-F(t)]dt + \pi_2/r.$$

Substituting the last equality into (6), J becomes

$$J = \pi_2/r + \int_0^\infty e^{-rt}\left([\pi_1(c(t))-\pi_2] - \theta(c(t))h(t)\right)[1-F(t)]dt. \tag{8}$$

There is a nice interpretation for (8). If the offender participates in the illegal activity, the net gains at each time t will be $\pi_1(c(t)) - \pi_2 - \theta(c(t))h(t)$ with a probability of 1-F(t). Thus, the offender receives with certainty the amount π_2/r from the legal activity and an additional expected amount from the illegal activity given by the second expression on the right-side of (8). Let $B(c) = \pi_1(c) - \pi_2$ and $\tilde{J} = J - \pi_2/r$.[11] Since π_2/r does not depend on the control variable of the decision problem, therefore, problem (D1) is equivalent to

---

[9] The hazard rate (function) provides a convenient way for modeling dynamic behavior. See the surveys by Heckman and Singer (1986) and Kiefer (1988) for details.

[10] The mathematical structure of this dynamic model of criminal behavior is similar to the dynamic limit-pricing model formulated by Kamien and Schwartz (1971). Davis (1988) also uses a similar model to study some aspects of dynamic criminal behavior.

[11] For convenience, the symbols B(c) and C(c) in section 2 will also be used in this section. Strictly speaking, the variables B(c) and C(c) in section 2 are stock variables whereas they are flow variables in this section.

$$\text{(D2)} \quad \max_{c(.)} \ \tilde{J} = \int_0^{\infty} e^{-rt}\left[B(c(t)) - \theta(c(t))h(t) \right][1-F(t)]dt, \tag{9}$$

s.t. (7).

Applying the theorem in Leung (1990), the optimal control problem (D2) is equivalent to the optimization problem:

$$\text{(D3)} \qquad \max_{c} \ \frac{B(c) - \theta(c)h(c)}{r + h(c)}. \tag{10}$$

Problem (D3) is an important simplification of problem (D2) because the former is just an ordinary optimization problem which is clearly much simpler than the original infinite horizon optimal control problem. It also implies that the optimal path of offense rate $c(t)$, $t \in [0,\infty]$, that solves (D2) is constant over time; and the value of the constant can be found from (D3).[12]

Two separate cases are considered below. The first one deals with the case where the hazard rate of arrest is exogenously given, i.e., $h(c)$ is a positive constant h. This simple case serves to illustrate the economics of the problem and it reveals several important and subtle differences between the static and the dynamic models. The second case allows $h(c)$ to depend on

---

[12] There is a technical point which is worth mentioning. Kamien and Schwartz (1971) provide a proof of the equivalence of (D2) and (D3) in the context of a dynamic limit-pricing model. The proof has been used by Kamien and Schwartz and many others (e.g., Davis 1988) to solve similar infinite horizon problems in different contexts. It is shown in Leung (1990) that the Kamien-Schwartz proof is erroneous because it assumes that the transversality condition is a <u>necessary</u> condition for optimality. It is well known that the transversality condition is in general <u>not</u> a necessary optimality condition for an infinite horizon control problem. Although there are several results in the literature that provide sufficient conditions under which the transversality condition is a necessary optimality condition, Leung (1990) shows that none of these results provides a satisfactory justification for the transversality condition in this type of models. An alternative proof of the equivalence of the two optimization problems is offered in Leung (1990).

c. The analysis becomes more complicated and the implications are drastically different from the first case.

Case 1: The hazard rate of arrest is a constant

When $h(c) = h$ (a positive constant), problem (D3) becomes

$$\text{(D4)} \qquad \max_{c} \frac{B(c) - \theta(c)h}{r + h}. \qquad (11)$$

What is the optimal fine in this case? Since r and h are constants, it is obvious from (11) that the optimal fine is given by

$$\theta(c) = C(c)/h, \qquad (12)$$

where C(c) is the social cost of the crime. It follows that the optimal multiplier is 1/h, the inverse of the hazard rate of arrest. The most important implication of this result is that the optimal multiplier is not necessarily greater than one because the hazard rate h is not bounded between 0 and 1. Hence, the optimal fine may not be a multiple of the damages caused by the illegal activity. Therefore, Becker's approach will not generate the optimal outcome in a dynamic environment. In fact, it is highly likely that the approach will cause overcompliance because the multiple fine imposes too heavy penalty on the offender.

Why is there a difference between the multipliers in Becker's approach and the dynamic one? One may argue that Becker's claim fails in the dynamic case because of the continuous time formulation of the dynamic model, and therefore it is not surprising to find that the multiplier 1/P is replaced by 1/h because the hazard rate h is the continuous time analog of the probability P. This argument is incorrect for two reasons. First, the hazard rate is the conditional density and is not analogous to the probability. For any short time interval Δt, the probability of arrest in $[t,t+\Delta t]$ is $f(t)\Delta t$, whereas $h\Delta t = f(t)\Delta t/[1-F(t)]$ is the conditional probability of arrest in $[t,t+\Delta t]$. Thus, it may be more appropriate to regard $f(t)$ as the continuous time analog of P. Second, the discrete-continuous time distinction is not crucial. In fact, the dynamic model (D4) closely resembles the static model (S1) because $B(c)/(r+h)$ and $\theta(c)/(r+h)$ can respectively be regarded as the $B(c)$ and $\theta(c)$ in model (S1). The only real distinction lies in the terms P and h. Furthermore, the objective function (11) can be written as $\{[1-F(t)]B(c)-f(t)\theta(c)\}/\{(r+h)[1-F(t)]\}$, which is similar to the static model (S2) because the numerator $[1-F(t)]B(c) - f(t)\theta(c)$ is analogous to the objective function $(1-P)B(c) - P\theta(c)$ of model (S2).[13] This interpretation shows that the critical difference between the static and the dynamic models arises from the dynamic and stochastic nature of the arrest process, regardless of whether the process is in discrete time or in continuous time. In a dynamic environment, the flow of the gains from the crime can be sustained only if the offender has not yet been arrested. As a result of this conditioning, the probability in the static model has to be replaced by the conditional probability in the dynamic model, which leads to the differences between the implications of the two models.

The dynamic model does not suffer from the shortcomings of model (S2) because α is no longer an arbitrary number. The quantity $\alpha B(c)$ in model (S2) is now fully specified and determined within the dynamic model, and has a nice structural interpretation. It describes the gains that the offender

---

[13] This similarity in the objective functions is the most important feature of the dynamic model (D4) that is captured by the static model (S2). This is the main reason for constructing model (S2) to compare with model (S1).

obtained from the time he started engaging in the crime to the time when he is caught. In the dynamic model (D1), it is just the returns $\pi_1(c(t))$, $t \in [0,T]$, that the offender has already reaped before he is caught at time T. Hence, a dynamic model is preferable to a static model because the former formulates and addresses this aspect explicitly.

The two static models (S1) and (S2) illustrate the differences between the static and the dynamic models. The key lies in the assumptions on the timing of the benefits and costs of the illegal activity. In the dynamic model, the returns come in continuously over a period of time until the offender is caught. At each moment, the offender may be caught in the act so that the flow of returns will end immediately. The fine is then based on the offense rate when he is caught. The static model implicitly assumes that the flow of returns can be aggregated into a lifetime (or one-period) income. While one can justify the aggregation assumption in the case of the returns from legal activities, it is more difficult to justify it in the case of the returns from illegal activities. In the present setup, the key distinction between legal and illegal activities is that the flow of returns from illegal activities can be terminated suddenly at any time (when the offense is detected). Ex ante, it is impossible to aggregate the flow of returns from illegal activities into a lifetime or a one-period income because of the presence of uncertainty. To take into account the uncertainty, the conditional probability argument is required, and this drives the differences between the static and the dynamic models.

Case 2: The hazard rate of arrest depends on the offense rate

When the hazard rate of arrest is allowed to depend on the offense rate, the problem becomes more complicated and the results are strikingly altered. The first is that the optimal multiplier $1/h(c)$ in case 1 is no longer optimal here.[14] The difference arises because the discount factor in (11) $(r+h)$ is a constant, whereas the discount factor in (10) $(r+h(c))$ is endogenously determined.

What is the optimal fine in this case? To proceed, first notice that the first-order condition of the optimization problem (D3) is given by

$$[r+h(c)][B'(c)-\theta'(c)h(c)-\theta(c)h'(c)]-[B(c)-\theta(c)h(c)]h'(c) = 0, \tag{13}$$

which can be rearranged to yield

$$B'(c) = \frac{[B(c)+r\theta(c)]h'(c)}{r + h(c)} + \theta'(c)h(c).$$

Since $B'(c) = C'(c)$ at the social optimum $c*$, hence the optimal fine $\theta(c)$ must satisfy the equality

$$C'(c) = \frac{[B(c)+r\theta(c)]h'(c)}{r + h(c)} + \theta'(c)h(c) \tag{14}$$

at $c = c*$. Therefore, one way to determine the optimal fine is to find a function $\theta(.)$ such that the functional equation (14) is satisfied for every c. Although (14) is a complicated functional equation, it can be rearranged to yield a first-order nonhomogeneous differential equation:

$$\theta'(c) + M(c)\theta(c) = N(c), \tag{15}$$

where $M(c) = \dfrac{rh'(c)}{[r+h(c)]h(c)}$ and $N(c) = \dfrac{C'(c)}{h(c)} - \dfrac{B(c)h'(c)}{[r+h(c)]h(c)}$.

Consequently, the problem of finding the function $\theta(.)$ from (14) becomes more manageable because it amounts to solving the differential equation (15)

---

[14] This can easily be checked by substituting $\theta(c) = C(c)/h(c)$ into (D3) and verifying that the c that solves (D3) is different from the social optimal offense rate.

for $\theta(.)$. To solve the differential equation, first rearrange (15) to obtain

$$\frac{d}{dc}\left[\frac{h(c)\theta(c)}{r+h(c)}\right] = \frac{h(c)}{r+h(c)}\left[\frac{C'(c)}{h(c)} - \frac{B(c)h'(c)}{[r+h(c)]h(c)}\right].$$

Assume $h(0) = 0$ and integrate both sides of the last equation from 0 to c,

$$\theta(c) = \frac{[r+h(c)]}{h(c)}\int_0^c \left[\frac{C'(x)}{r+h(x)} - \frac{B(x)h'(x)}{[r+h(x)]^2}\right] dx. \tag{16}$$

Thus, the optimal fine $\theta(c)$ is related to the functions B(.), C(.), and h(.) in a complicated way. It is clear from (16) that the optimal multiplier $\theta(c)/C(c)$ is not necessarily greater than one.

Both the results from cases 1 and 2 demonstrate that the multiplier may be smaller than one in a more general dynamic setting. The dissimilarities between the results of cases 1 and 2 arise from the endogeneity of the hazard rate of arrest. In case 2, the term r+h(c) distorts both the gains and the costs of the crime because the gains and the costs are essentially $B(c)/[r+h(c)]$ and $\theta(c)/[r+h(c)]$ respectively. In case 1, r+h(c) is a constant so that there is no real distortion on the gains and the costs since r+h is just a scale factor. In case 2, the distortion is in effect because an increase in c raises both the gains and the hazard rate, and the results on the real gains $B(c)/[r+h(c)]$ and the real costs $\theta(c)/[r+h(c)]$ become ambiguous. These ambiguities transmit to the optimal fine function and result in the complicated expression in (16).

Since the implications of cases 1 and 2 are so different, it would be useful to know whether the hazard rate of arrest indeed depends on the offense rate. This is mainly an empirical question which depends largely on the operations of the law enforcement agencies and on the type of crimes involved. For example, consider a polluting factory and assumes that the factory will be indicted whenever the amount of pollutants discharged into the environment (the offense rate) is found to be above the legal standard. If the environmental agency has a regular schedule of inspecting the environment surrounding the factory, then the hazard rate of arrest will not depend on the offense rate. (Of course, the penalty may depend on the offense rate).[15] On the other hand, if the environmental agency does not have a regular inspection schedule and only inspects the environment when the agency receives complaints, then the hazard rate of arrest will increase with the offense rate because the likelihood of complaints increases with the amount of pollutants discharged. These considerations indicate the complications involved in setting the optimal fine in a dynamic environment.

4. Conclusion

Previous studies on optimal punishment have ignored dynamic considerations. In this paper, I have investigated dynamic optimal punishment theory. In particular, I demonstrate that the optimal fine derived from a dynamic model is drastically different from the one obtained from a static model. The analysis shows that Becker's (1968) celebrated result in the static setup (the optimal fine is a multiple of the damages done) is no longer valid in a dynamic setting. In fact, the multiple fine that Becker advocates may be too heavy and may lead to nonoptimal outcomes because of overcompliance. This result calls into question the validity of applying Becker's approach to sentence corporate offenders, as currently

[15] Speeding in highways is a similar case. In this case, offense rate = speed of the car - speed limit. Police will ticket the driver when the speed of the car exceeds the speed limit, and the hazard rate of arrest does not depend on the number of miles per hour above the speed limit (i.e., the offense rate).

proposed by the U.S. Sentencing Commission.

Even if one ignores the dynamic analysis, the results of the two static models suggest that Becker's approach should at least be amended in the following way. Since model (S2) is also a static model and may be a better description of reality than model (S1), the optimal multiplier should be set between (1-P)/P and 1/P. This implies that (1-P)C(c)/P and C(c)/P should be the lower and the upper bounds of the optimal fine. Using the oil spill example in section 1, the optimal fine should be between $0.9 and $1 per gallon of oil leaked. Although setting the bounds may not necessarily produce the optimal outcome, it seems to be a better strategy since it reduces the range of errors, given that both models (S1) and (S2) are plausible descriptions of reality.

In many circumstances, the crime process is a dynamic one and the dynamic model is a better description of reality than the static model. Hence, the dynamic analysis cannot be ignored. To set the optimal fine in a general dynamic setting, three crucial pieces of information are required: the gains from the crime, the social cost of the crime, and the hazard rate of arrest. Admittedly, little is known about these empirical aspects of criminal behavior. Therefore more research is needed before optimal punishment theory can be confidently applied to public policy-making. In view of the present state of knowledge, the widespread acceptance of Becker's approach seems to be premature.

References

Allingham, Michael and Agnar Sandmo, 1972, Income tax evasion: A theoretical analysis, Journal of Public Economics 1, 323-338.

Becker, Gary, 1968, Crime and punishment: An economic approach, Journal of Political Economy 76, 169-217.

Becker, Gary, 1989, Make the punishment fit the corporate crime, Business Week (March 13, 1989), 22.

Craswell, Richard and John Calfee, 1986, Deterrence and uncertain legal standards, Journal of Law, Economics, and Organization 2, 279-303.

Davis, Michael, 1988, Time and punishment: An intertemporal model of crime, Journal of Political Economy 96, 383-390.

Ellis, Dorsey, Jr., 1982, Fairness and efficiency in the law of punitive damages, Southern California Law Review 56, 1-78.

Farber, Daniel, 1980, Reassessing the efficiency of compensatory damages for breach of contract, Virginia Law Review 66, 1443-1484.

Heckman, James and Burton Singer, 1986, Econometric analysis of longitudinal data, in: Zvi Griliches and Michael Intriligator, eds., Handbook of econometrics, Vol. III. (Elsevier, New York) 1689-1763.

Heineke, J. M., 1975, A note on the criminal choice problem, Journal of Economic Theory 10, 113-116.

Kamien, Morton and Nancy Schwartz, 1971, Limit pricing and uncertain entry, Econometrica 39, 441-454.

Kiefer, Nicholas, 1988, Economic duration data and hazard functions, Journal of Economic Literature 26, 646-679.

Leung, Siu Fai, 1990, Transversality condition and optimality in a class of infinite horizon continuous time economic models, manuscript, University of Rochester. Forthcoming in the Journal of Economic Theory.

Levinson, Michael, 1980, Deterring air polluters through economically efficient sanctions: A proposal for amending the clean air act, Stanford Law Review 32, 807-826.

Mandel, Michael, 1990, Putting the fear of crime into corporations, Business Week (March 12, 1990), 28.

Nash, John D., 1989, To make the punishment fit the crime: The theory and statistical estimation of a multi-period optimal deterrence model, manuscript, World Bank. Forthcoming in the International Review of Law and Economics.

Polinsky, A. Mitchell and Daniel Rubinfeld, 1989, A model of optimal fines for repeat offenders, manuscript, Stanford Law School.

Polinsky, A. Mitchell and Steven Shavell, 1984, The optimal use of fines and imprisonment, Journal of Public Economics 24, 89-99.

Polinsky, A. Mitchell and Steven Shavell, 1990, Reviewing the rules of corporate punishment, Business Week (April 30, 1990), 5-6.

Posner, Richard, 1976, Antitrust law: An economic perspective. (University of Chicago Press, Chicago).

Posner, Richard, 1977, Economic analysis of law. (Little, Brown and Co., Boston).

Posner, Richard, 1980, Optimal sentences for white-collar criminals, American Criminal Law Review 17, 409-418.

Smart, Tim and Michele Galen, 1990, Soon, corporate crime may really not pay, Business Week (February 12, 1990), 35.

Stern, Nicholas, 1978, On the economic theory of policy towards crime, in: J. M. Heineke, ed., Economic models of criminal behavior (North Holland, Amsterdam) 123-152.

United States Sentencing Commission, 1989, Sentencing guidelines for organizational defendants, Preliminary draft, November 1989.