

NBER WORKING PAPER SERIES

THE OPTIMAL USE OF FINES AND IMPRISONMENT
WHEN WEALTH IS UNOBSERVABLE

A. Mitchell Polinsky

Working Paper 10761
<http://www.nber.org/papers/w10761>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2004

This research was supported by the John M. Olin Program in Law and Economics at Stanford Law School. I received helpful comments from Nuno Garoupa, Jason Johnston, Daniel Rubinfeld, Steven Shavell, Kathryn Spier, and Robert Weisberg. The views expressed herein are those of the author(s) and not necessarily those of the National Bureau of Economic Research.

©2004 by A. Mitchell Polinsky. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

The Optimal Use of Fines and Imprisonment When Wealth is Unobservable

A. Mitchell Polinsky

NBER Working Paper No. 10761

September 2004

JEL No. D31, D62, H23, K14, K42

ABSTRACT

This article studies the optimal use of fines and imprisonment when an offender's level of wealth is private information that cannot be observed by the enforcement authority. In a model in which there are two levels of wealth, I derive the optimal mix of sanctions, including the imprisonment sentence imposed on offenders who do not pay the fine – referred to as the “alternative” imprisonment sentence. Among other things, I demonstrate that if imprisonment sanctions are used, the optimal alternative imprisonment sentence is sufficiently high that high-wealth individuals prefer to pay a fine exceeding the wealth level of low-wealth individuals and bear a lower (possibly no) imprisonment sentence rather than to pretend to be low-wealth individuals. I also show that if the optimal enforcement system would rely exclusively on fines when wealth is observable, the inability to observe wealth is detrimental because higher fines then could not be levied on higher-wealth individuals. In this case, it may be desirable when wealth is unobservable to impose an imprisonment sentence on offenders who do not pay the fine – who will be low-wealth offenders – in order to induce high-wealth offenders to pay the fine. However, if the optimal enforcement system would employ both fines and imprisonment sentences when wealth is observable, the inability to observe wealth is not detrimental. In this case, the same sanctions would be chosen if wealth is unobservable and these sanctions lead high-wealth individuals to pay more than low-wealth individuals.

A Mitchell Polinsky

Stanford Law School

Stanford, CA 94305

and NBER

polinsky@stanford.edu

I. Introduction

Individuals have many ways of hiding assets from government enforcers, including by hoarding cash, transferring assets to relatives or friends, or moving money to offshore bank accounts. Notwithstanding these opportunities, prior analyses of optimal sanctions nearly always have assumed that an offender's level of wealth is costlessly observable by the enforcement authority. Knowing wealth levels, the enforcement authority then chooses the sanctions to impose, fines and/or imprisonment sentences. The contribution of the present article is to consider optimal sanctions when an offender's level of wealth is private information that cannot be observed by the enforcement authority.¹

My analysis is based on a model in which there are two types of risk-neutral potential offenders — a low-wealth type and a high-wealth type. If an offender is caught, he is sanctioned with a fine and, possibly, an imprisonment sentence. If he claims that he cannot pay the fine because he is a low-wealth individual, a higher imprisonment sentence might be imposed on him — referred to as the “alternative” imprisonment sentence. The enforcement authority's problem is to choose the following policy instruments to maximize social welfare: the fine and imprisonment sentence for committing the offense; the alternative imprisonment sentence if an offender claims that he cannot pay the fine; and the probability of detecting the offense.

The optimal enforcement system when wealth is unobservable takes one of three forms, depending on the cost of imprisonment. At one extreme, if the cost of imprisonment is relatively low, it may be optimal to impose both fines and imprisonment sentences on offenders and to employ an alternative imprisonment sentence that is high enough so that high-wealth individuals

¹ I discuss in Section V the most relevant prior literature on this topic — Levitt (1997), Chu and Jiang (1993), and Polinsky (2004) — and how this article relates to it.

prefer to pay the fine and bear the corresponding imprisonment sentence rather than to pretend to be low-wealth individuals. At the other extreme, if the cost of imprisonment is sufficiently high, it will be optimal to rely exclusively on fines to sanction offenders; because wealth is unobservable, the fine cannot exceed the wealth level of low-wealth individuals. An intermediate case also is possible, in which offenders who pay the fine are not also imprisoned, but those who do not pay the fine bear an (alternative) imprisonment sentence; in this case, too, the optimal alternative imprisonment sentence is high enough to discourage high-wealth individuals from pretending to be low-wealth individuals. In each of these cases, I derive the optimal sanctions and discuss whether low-wealth and high-wealth individuals are underdeterred or overdeterred relative to the first-best outcome.

Perhaps surprisingly, I show that information about wealth levels may or may not be helpful. The key point is that such information is useful only if the enforcement authority would want to impose higher sanctions on high-wealth individuals than on low-wealth individuals, for if this is the case, high-wealth individuals would pretend to be low-wealth individuals if wealth could not be observed. Conversely, if the enforcement authority would want to impose lower sanctions on high-wealth individuals than on low-wealth individuals, high-wealth individuals will voluntarily bear such sanctions. I show that both of these possibilities can occur.

A notable result from my analysis is that, if wealth is unobservable, it may be worthwhile to use an alternative imprisonment sentence to induce high-wealth offenders to pay more than low-wealth offenders even if imprisonment is so costly that that it would not be used if wealth were observable. In other words, when wealth is not observable, it may be desirable to impose a costly sanction — imprisonment sentences — on low-wealth offenders in order to better deter

high-wealth offenders through a cheap sanction — fines. If wealth were observable, it would not be necessary to incur this cost.

Section II presents the general framework used in this article. Section III derives the optimal enforcement system under different assumptions about the cost of imprisonment. Section IV compares the results to the case when wealth is observable. Section V discusses the prior literature in relation to the analysis here. Section VI considers various generalizations.

II. General Framework

In the model, individuals contemplate whether to commit an offense that causes harm. Each individual is identified by the benefit he would obtain from committing the offense and by his level of wealth. For simplicity, I assume that there are two levels of wealth and that an individual's benefit is independent of his wealth level.² An individual who commits the offense is detected with a probability that is determined by the enforcement expenditures of the state.³ If detected, he may be sanctioned with a fine and/or an imprisonment sentence. If he does not pay the fine, he may be subject to a higher alternative imprisonment sentence. I assume for simplicity that individuals are risk neutral with respect to fines and imprisonment sentences, and that they bear the same disutility from a given sentence.⁴ Fines are treated as socially costless

² I discuss in Section VI why my general results should continue to hold if there is a continuum of wealth levels.

³ This probability is assumed to be independent of the individual's level of wealth. This makes sense in many, if not most, enforcement contexts.

⁴ An individual who is risk neutral with respect to an imprisonment sentence only cares about the expected value of the sentence length. I explain in Section VI why my general results would not be affected if individuals are not risk neutral with respect to fines or imprisonment sentences, or if they differ in terms of their disutility from time in prison.

sanctions, while the cost of imprisonment to the state is assumed to be proportional to the length of the sentence.

The following notation will be used.

h = harm caused if the offense is committed; $h > 0$;

b = benefit from committing the offense; $b \geq 0$;

$r(b)$ = probability density of b ; r is positive for all $b \geq 0$;

w_L = level of wealth of low-wealth individuals; $w_L \geq 0$;

w_H = level of wealth of high-wealth individuals; $w_H > w_L$;

θ = fraction of individuals with a low level of wealth;

e = enforcement expenditures of the state; $e \geq 0$;

$p(e)$ = probability of detection; $p(0) = 0$; $p'(e) > 0$; $p''(e) < 0$;

f = fine for committing the offense; $f \geq 0$;⁵

s = imprisonment sentence for committing the offense

if the fine f is paid; $s \geq 0$;

s' = alternative imprisonment sentence for committing the

offense if less than f is paid; $s' \geq 0$; and

c = cost to the state per unit of imprisonment sentence; $c > 0$.

The state is assumed to know the distribution of wealth levels among the population, but not the wealth level of a particular individual. Imprisonment sentences are measured in units of time such that one unit of an imprisonment sentence corresponds to one dollar's worth of disutility to

⁵ This is the nominal fine. An individual might not have sufficient wealth to pay this fine.

an individual. Thus, an imprisonment sentence of length s imposes a cost on an individual equal to s .

Individual Behavior. Let

b_L = critical value of benefit below which a low-wealth individual will

not commit the offense and at and above which he will,⁶

and define b_H analogously. Thus, the higher b_L or b_H , the greater the level of deterrence.

If the fine f is less than or equal to a low-wealth individual's level of wealth w_L , then clearly everyone can be made to pay f and bear the imprisonment sentence s . Any alternative imprisonment sentence $s' \geq f + s$ will accomplish this. In this case,

$$b_L = b_H = p(e)(f + s). \quad (1)$$

If f exceeds w_L , a low-wealth individual will not be able to pay the fine. Such individuals can be made to pay w_L , however, if the state threatens to impose a sufficiently high imprisonment sentence, say $s'' \geq w_L + s'$, on anyone who does not pay at least w_L . A low-wealth individual therefore will pay w_L and bear the imprisonment sentence s' , in which case

$$b_L = p(e)(w_L + s'). \quad (2)$$

A high-wealth individual has the choice of paying the fine f and bearing the imprisonment sentence s ,⁷ or claiming that he is a low-wealth individual and paying w_L and bearing the alternative imprisonment sentence s' . Hence,

$$b_H = p(e)\min(f + s, w_L + s'). \quad (3)$$

⁶ I assume without loss of generality that he will commit the offense if he is indifferent.

⁷ It is obvious that there is no advantage to the enforcement authority from setting the fine in excess of the wealth level of the high-wealth individuals.

I assume without loss of generality that if $f + s = w_L + s'$, a high-wealth individual will pay the fine f and bear the imprisonment sentence s .

Social welfare. Social welfare is the sum of the benefits obtained by individuals who commit the offense, less the harm done, less the private and public cost of imprisonment, and less the cost of detection. It will be convenient to state social welfare in two cases.

If the fine f is less than or equal to w_L , then social welfare is

$$\int_{p(e)(f+s)}^{\infty} [b - h - p(e)(1+c)s]r(b)db - e. \quad (4)$$

Everyone whose benefit equals or exceeds $p(e)(f+s)$ commits the offense, resulting in a benefit b and a harm h . With probability $p(e)$ such individuals are detected and made to bear an imprisonment sentence s . The social cost of imprisonment is the sum of the private cost s and public cost cs . The state also incurs enforcement expenditures e . The fines that individuals pay do not enter social welfare directly because they are transfer payments.

If the fine f exceeds w_L and high-wealth individuals pay f , social welfare is

$$\begin{aligned} & \theta \int_{p(e)(w_L+s')}^{\infty} [b - h - p(e)(1+c)s]r(b)db \\ & + (1-\theta) \int_{p(e)(f+s)}^{\infty} [b - h - p(e)(1+c)s]r(b)db - e. \end{aligned} \quad (5)$$

The interpretation of (5) is the same as that of (4), though now low-wealth and high-wealth individuals face different sanctions because low-wealth individuals cannot pay f .⁸

The state's problem. The state's problem is to maximize social welfare through the choice of the fine f , the imprisonment sentence s , the alternative imprisonment sentence s' , and

⁸ It is not necessary to state social welfare for the case in which f exceeds w_L and high-wealth individuals pay w_L ; I demonstrate below that this case cannot occur in the optimal enforcement system.

enforcement expenditures e .⁹ Monetary payments by low-wealth and high-wealth individuals cannot exceed their respective levels of wealth. I do not impose an upper bound on imprisonment sentences; it will be evident below that there is no reason to expect the optimal imprisonment sentences to be maximal, so any corner solution resulting from limiting imprisonment sentences would not be particularly interesting.

Asterisks are used to denote the solution to the state's problem, which I assume is unique. I also assume that some enforcement is optimal, that is, $e^* > 0$; otherwise, the problem is uninteresting.

III. Analysis

There are three cases to consider, depending on whether imprisonment is used to sanction all offenders, or used just to sanction those who do not pay the fine, or not used at all.¹⁰ In this section I describe the optimal enforcement system through propositions that correspond to these cases. Following the proof of each proposition is an informal discussion of the results.

Proposition 1. If the public cost of imprisonment c is sufficiently low, it may be optimal to use imprisonment both to sanction offenders who pay the fine as well as those who do not; that is, both s^ and s'^* may be positive. In this case,*

(a) the optimal fine is maximal: $f^ = w_H$;*

(b) the optimal alternative imprisonment sentence (imposed if the fine is not paid) exceeds s^ : $s'^* > s^*$;*

⁹ Because my objective is to derive the socially optimal enforcement policy, I ignore the issue of whether the state's policy is credible.

¹⁰ Using imprisonment just to sanction offenders who pay the fine obviously is not feasible because then no one would pay the fine.

(c) high-wealth individuals pay the fine f^* rather than a payment equal to the wealth level of a low-wealth individual w_L ;

(d) the burden of the sanctions imposed on low-wealth individuals exceeds that imposed on high-wealth individuals: $w_L + s'^* > f^* + s^*$; and

(e) both groups may be underdeterred or overdeterred.

Proof: (i) I first show that if $s^* > 0$, it must be that $f^* \geq w_L$. Suppose otherwise, that $s^* > 0$ and $f^* < w_L$. Then it would be possible to raise f (with f remaining below w_L) and lower s so as to keep $f + s$ constant without affecting behavior, but raising social welfare by reducing the cost of imprisonment (see (4)).

(ii) Now suppose that $s^* > 0$ and $f^* = w_L$. This cannot be optimal either. It would be possible to set s' equal to s^* and then raise f above w_L and lower s by the same amount. This would increase social welfare because neither group's behavior would be affected, yet the cost of imprisoning high-wealth individuals would decline (see (5)).

(iii) Hence, if $s^* > 0$, it must be that $f^* > w_L$. But if $f^* < w_H$, social welfare can be increased by raising f and lowering s by the same amount, without changing s' ; again, the cost of imprisonment declines for high-wealth individuals. Hence, if $s^* > 0$, it must be that $f^* = w_H$, establishing part (a).

(iv) To prove part (b), suppose the contrary, that $s'^* < s^*$. Then, since $f^* = w_H > w_L$, high-wealth individuals would prefer to pay w_L and bear the imprisonment sentence s'^* , the same sanctions borne by low-wealth individuals. But this cannot be optimal by the type of argument used in step (ii). This establishes part (b). Part (c) follows from the same reasoning.

(v) I next show that if $s^* > 0$, s'^* and s^* are such that $w_L + s'^* > w_H + s^*$. If $s^* > 0$, then the first-order condition with respect to s derived from (5) is, with $f = w_H$ and after dividing through by $-(1 - \theta)p(e)$,

$$(1 + c)[1 - R(b_H)] + [b_H - h - p(e)(1 + c)s]r(b_H) = 0, \quad (6)$$

where $b_H = p(e)(w_H + s)$. The derivative of social welfare with respect to s' is

$$-\theta p(e)\{(1 + c)[1 - R(b_L)] + (b_L - h - p(e)(1 + c)s')r(b_L)\}, \quad (7)$$

where $b_L = p(e)(w_L + s')$. Consider this derivative at the value of s' such that $w_L + s' = w_H + s^*$, that is, at $s' = s^* + (w_H - w_L)$. At this value of s' , (7) can be written as

$$\begin{aligned} & -\theta p(e)\{(1 + c)[1 - R(b_L)] + (b_L - h - p(e)(1 + c)s^*)r(b_L)\} \\ & + \theta p(e)^2[(1 + c)(w_H - w_L)]r(b_L), \end{aligned} \quad (8)$$

where $b_L = p(e)(w_L + s') = p(e)(w_H + s^*)$. Since b_L at this s' equals b_H at s^* , (6) implies that the first term in (8) is zero. The second term in (8) clearly is positive, so (8) is positive at $s' = s^* + (w_H - w_L)$, implying that $s'^* > s^* + (w_H - w_L)$ or, equivalently, $w_L + s'^* > w_H + s^*$. This establishes part (d).

(vi) I next address whether individuals will be underdeterred or overdeterred relative to the first-best outcome. First consider low-wealth individuals and the derivative of social welfare with respect to s' (7). Evaluated at s' such that $b_L = p(e)(w_L + s') = h$, it is clear that the sign of (7) is the same as the sign of $p(e)s'r(h) - [1 - R(h)]$. The first term could be very low, for example if $p(e)$ is low, in which case the sign of (7) would be negative; this implies that s' should be chosen such that $b_L < h$, resulting in underdeterrence. Conversely, the second term could be very low — if $R(h)$ is close to one — in which case the sign of (7) would be positive and b_L would exceed h , so there would be overdeterrence. Similar observations apply to high-wealth individuals. This establishes part (e).

(vii) Finally, I demonstrate that there exists a sufficient condition for the optimal imprisonment sentence s^* to be positive. Given $f^* = w_H$, it can be shown from (5) that if the public cost of imprisonment c is zero, the derivative of social welfare with respect to s is positive at $s = 0$ if

$$h > w_H + \sup_{p \in [0, 1]} \{[1 - R(pw_H)]/r(pw_H)\}. \quad (9)$$

Thus, if (9) is satisfied and c is sufficiently low, $s^* > 0$. ■

It is obvious that it might be socially beneficial to use imprisonment despite its cost if the harm from the offense is sufficiently high. For example, if the harm far exceeds the wealth level of high-wealth individuals, using fines alone will result in substantial underdeterrence even if detection is certain; the additional deterrence from employing imprisonment will be socially desirable if the cost of imprisonment is not too great. Because the private cost of imprisonment per unit of deterrence is unavoidable,¹¹ this means that the public cost of imprisonment c must be sufficiently low.

That the fine should be maximal before imprisonment is used follows from well-known logic. Specifically, since fines are socially cheaper sanctions than are imprisonment sentences, fines should be used to their fullest extent before the state resorts to using imprisonment.

The key insight from Proposition 1 is that the optimal alternative imprisonment sentence imposed on individuals who do not pay the fine is high enough to make high-wealth individuals prefer to pay the fine. The reason it is optimal to impose such a high imprisonment sentence on individuals who do not pay the fine, all of whom are low-wealth individuals, can be explained as follows (in a way that parallels the proof). Consider an alternative imprisonment sentence s'

¹¹ To create a dollar's worth of disutility to an offender by imprisoning him necessarily causes a dollar's worth of social cost.

borne by low-wealth individuals that, combined with a monetary payment equal to their wealth level w_L , creates the same degree of deterrence as for high-wealth individuals, given their payment of the fine f and bearing of the imprisonment sentence s . Obviously, the imprisonment sentence borne by the low-wealth individuals must be higher. Now raise the imprisonment sentence borne by each group. The effect on social welfare from the additional deterrence is the same for each group since the level of deterrence initially is the same (and the distribution of benefits is assumed to be the same for both groups). But because the imprisonment sentence borne by low-wealth individuals is longer, the social benefit of raising the imprisonment sentence is greater for them — society saves more on imprisonment costs for each low-wealth individual deterred from committing the offense. Thus, it is optimal to impose a higher imprisonment sentence on low-wealth individuals than that which would equalize the deterrence of both groups. Consequently, high-wealth individuals prefer to pay the fine and thereby bear a lower total burden from the combined sanctions.

The explanation why both groups may be underdeterred or overdeterred is based on a familiar argument. Consider low-wealth individuals first, and suppose that their monetary payment and the (alternative) imprisonment sentence applicable to them are such that their expected combined sanction equals the harm. If the imprisonment sentence is raised, there is no first-order effect on social welfare with respect to the benefit and harm from committing the offense because the marginal individuals who are deterred from committing the offense were obtaining a benefit equal to the harm. But raising the sentence does have two effects on the cost of imprisonment. On one hand, the cost of imprisonment tends to increase because those individuals who continue to commit the offense now bear a longer sentence if they are detected. On the other hand, the cost of imprisonment tends to decrease because fewer individuals commit

the offense and serve imprisonment sentences. Either effect could dominate, possibly making it desirable to raise the imprisonment sentence — resulting in overdeterrence — or to lower the imprisonment sentence — resulting in underdeterrence. Analogous observations apply to high-wealth individuals.

The next proposition addresses the case in which offenders who pay the fine are not also subject to an imprisonment sentence, but offenders who do not pay the fine do bear an imprisonment sentence.

Proposition 2. It may be optimal to use imprisonment only to sanction offenders who do not pay the fine; in other words, it is possible that $s^ = 0$ and $s'^* > 0$. In this case,*

(a) the optimal fine exceeds the wealth level of low-wealth individuals but might not be maximal: $w_L < f^ \leq w_H$;*

(b) high-wealth individuals pay the fine f^ rather than a payment equal to the wealth level of a low-wealth individual w_L ;*

(c) the burden of the sanctions imposed on low-wealth individuals equals or exceeds that imposed on high-wealth individuals: $w_L + s'^ \geq f^*$; and*

(d) both groups may be underdeterred, but only low-wealth individuals may be overdeterred.

Proof: (i) I first show that if $s^* = 0$ and $s'^* > 0$, it must be that $f^* > w_L$. Suppose otherwise, that $s^* = 0$, $s'^* > 0$, and $f^* \leq w_L$. Given these sanctions, both groups pay f^* and do not bear an imprisonment sentence. Any value of s' , including $s' = 0$, would result in the same outcome, contradicting the uniqueness of s'^* . Hence, if $s^* = 0$ and $s'^* > 0$, it must be that $f^* >$

w_L . That f^* can be less than w_H is demonstrated in the analytical example in step (v) below; this example also can be modified to show that $f^* = w_H$ is possible.¹² This establishes part (a).

(ii) Suppose, given $s^* = 0$, $s'^* > 0$, and $f^* > w_L$, that high-wealth individuals prefer to pay w_L and bear the imprisonment sentence s'^* . In other words, suppose $w_L + s'^* < f^*$. Then both groups would be paying w_L and bearing the sentence s'^* . This cannot be optimal because it would be possible to improve social welfare by lowering f to $w_L + s'^*$, in which case high-wealth individuals now pay f and are deterred to the same extent, and the cost of imprisonment declines because they no longer bear imprisonment sentences. This proves part (b).

(iii) The prior step showed that $w_L + s'^* < f^*$ cannot hold. Hence, it must be that $w_L + s'^* \geq f^*$. The analytical example in step (v) below shows that strict inequality can occur. I show here that equality also is possible. The derivative of social welfare with respect to s derived from (5) is

$$-(1 - \theta)p(e)\{(1 + c)[1 - R(b_H)] + [b_H - h - p(e)(1 + c)s]r(b_H)\}, \quad (10)$$

where $b_H = p(e)(f + s)$. Since $s^* = 0$, (10) must be non-positive at $s = 0$, which implies that

$$(1 + c)[1 - R(b_H)] + [b_H - h]r(b_H) \geq 0, \quad (11)$$

where $b_H = p(e)f$. The derivative of social welfare with respect to s' is (7), where $b_L = p(e)(w_L + s')$. Consider this derivative at the value of s' such that $w_L + s' = f^*$, that is, at $s' = f^* - w_L$. At this value of s' , (7) can be written as

$$\begin{aligned} & -\theta p(e)\{(1 + c)[1 - R(b_L)] + (b_L - h)r(b_L)\} \\ & + \theta p(e)^2[(1 + c)(f^* - w_L)]r(b_L), \end{aligned} \quad (12)$$

¹² See note 13 below.

where $b_L = p(e)(w_L + s') = p(e)f^*$. Since b_L at this s' equals b_H at f^* , (11) implies that the first term in (12) is non-positive. The second term in (12) is positive. If (12) is non-positive, then it will not be optimal to raise s' above $s' = f^* - w_L$. A sufficient condition for (12) to be negative is that $r(b_L)$ is sufficiently low. Then $s'^* = f^* - w_L$, or $w_L + s'^* = f^*$. Thus, $w_L + s'^* \geq f^*$ can hold with equality or strict inequality, proving part (c).

(iv) The logic used in step (vi) of the proof of Proposition 1 implies here as well that the low-wealth individuals can be underdeterred or overdeterred. If they are underdeterred, then the high-wealth individuals also must be underdeterred because of part (c) of the present proposition. If the low-wealth individuals are overdeterred, however, it would never be optimal to set f such that the high-wealth individuals are overdeterred since f could be lowered to $h/p(e)$, leading high-wealth individuals to be first-best deterred.

(v) Finally, I provide an analytical example in which $s^* = 0$ and $s'^* > 0$. Suppose $w_L < h < w_H$ and, for simplicity, that the probability of detection can be set at any level at no cost (this assumption is not essential). Then, clearly, $p(e^*) = 1$; this allows high-wealth individuals to be first-best deterred by setting $f = h$ and results in low-wealth individuals being deterred to the greatest extent feasible by a monetary payment alone. If high-wealth individuals can be first-best deterred by a fine alone, the optimal imprisonment sentence would be $s = 0$. Low-wealth individuals, however, would be underdeterred by a monetary payment alone because $w_L < h$. To see whether it is desirable to employ an alternative imprisonment sentence s' sufficiently large so that high-wealth individuals prefer to pay $f = h$, consider the derivative of social welfare (5) with respect to s' evaluated at s' such that $w_L + s' = h$, that is, at $s' = h - w_L$. Assuming that $p = 1$, this derivative is positive at $s' = h - w_L$ if

$$(h - w_L)r(h) - [1 - R(h)] > 0. \quad (13)$$

It is evident that (13) can be satisfied — for instance, if $R(h)$ is sufficiently close to one — in which case $s'^* > h - w_L$. High-wealth individuals therefore prefer to pay $f^* = h$ rather than $w_L + s'^* > h$.¹³ Clearly, other examples in which $s^* = 0$ and $s'^* > h - w_L$ could be constructed similar to this one even if it is costly to raise the probability of detection. ■

That this case can occur should not be surprising. Because high-wealth individuals can be deterred by fines to a greater extent than low-wealth individuals, the value of deterring high-wealth individuals further by imposing an imprisonment sentence on them is lower. Thus, given the cost of imprisonment, it might not be worthwhile to impose an imprisonment sentence on high-wealth offenders who pay the fine, even though it may be desirable to impose an imprisonment sentence on low-wealth offenders who cannot pay the fine.

Although much of Proposition 2 parallels Proposition 1, there are three notable differences. First, the optimal fine no longer is necessarily maximal. It is maximal in the case described by Proposition 1 because high-wealth offenders are subject to an imprisonment term, and fines, being a socially cheaper sanction, should be used to the fullest extent possible before imprisonment is employed. Here, high-wealth individuals do not bear an imprisonment sentence, so this argument does not apply. It is possible, therefore, that the optimal fine is less than maximal. For example, given the probability of detection, it might not be necessary to raise the fine to the maximal level in order to achieve first-best deterrence.

Second, while in Proposition 1 the burden of the sanctions imposed on low-wealth offenders strictly exceeds that imposed on high-wealth offenders, here they can be equal. The

¹³ To see that this example can be modified so that $f^* = w_H$, suppose that the probability of detection can be raised to $h/w_H < 1$ at no cost, and that it is excessively costly to raise it higher; thus, $p^* = h/w_H$. Then setting $f = w_H$ and $s = 0$ leads to first-best deterrence of high-wealth individuals since $p^*f = h$, and analogous reasoning to that used in step (v) implies that $w_L + s'^* > w_H$ can hold, so high-wealth individuals again prefer to pay f^* .

earlier argument stemmed from the fact that when both low-wealth and high-wealth offenders are subject to imprisonment sentences, deterring a low-wealth offender generates greater savings in imprisonment costs because low-wealth offenders have to bear longer imprisonment sentences to achieve the same level of deterrence. This argument is not applicable here because high-wealth offenders are not subject to an imprisonment sentence. Now, starting from a situation in which both low-wealth and high-wealth individuals are equally deterred, with an imprisonment sentence imposed only on low-wealth offenders, it may not be desirable to raise the imprisonment sentence (whether it is depends on the desirability of the additional deterrence, the imprisonment cost savings from the additional deterrence, and the additional imprisonment costs due to undeterred low-wealth offenders serving longer sentences). If a higher imprisonment sentence is not desirable, then both groups will bear an equal burden from the sanctions.¹⁴

Finally, while both groups could be underdeterred or overdeterred in the case examined in Proposition 1, here the high-wealth group never is overdeterred. The only rationale for overdetering either group is to reduce the burden of imprisonment costs by reducing the number of offenders. Because high-wealth offenders here are not bearing an imprisonment sanction, this argument does not apply to them.

The last proposition considers the case in which fines alone are used. Although the results are straightforward, I include this case for completeness.

Proposition 3. If the public cost of imprisonment c is sufficiently high, it will be optimal to rely exclusively on fines to sanction offenders. In this case,

(a) the optimal fine equals the wealth level of low-wealth individuals: $f^ = w_L$;*

¹⁴ The low-wealth offenders cannot bear a lower burden, for then high-wealth offenders could not be induced to pay the fine f .

(b) everyone pays the fine f^* ; and

(c) both groups are undeterred.

Proof: (i) I first show that it must be that $f^* \geq w_L$. Suppose otherwise, that $f^* < w_L$. Then it would be possible to raise f and lower e so as to keep $p(e)f$ constant without affecting behavior, but reducing enforcement costs. If $f > w_L$, high-wealth individuals would not pay f unless there is a positive alternative imprisonment sentence s' , which would then be borne by low-wealth offenders, contradicting the premise that imprisonment costs make the imposition of imprisonment sentences socially undesirable. This proves part (a).

(ii) Everyone can be made to pay $f^* = w_L$ by the threat of an alternative imprisonment sentence s' such that $s' \geq w_L$. Since no one actually bears this sentence, the premise of this proposition is not contradicted. This establishes part (b).

(iii) The first-order condition for determining e^* from (4) can be written as $(h - p(e)f)[dR(p(e)f)/de] = 1$, which implies that $p(e)f < h$, proving part (c).

(iv) Starting from the outcome when fines alone are used, the marginal benefit of additional deterrence due to the use of an alternative imprisonment sentence s' is at most the harm h (it is $h - p(e^*)w_L$ for each low-wealth individual deterred by the imposition of s' and for each high-wealth individual deterred by setting $f = w_L + s'$). The marginal cost of employing an alternative imprisonment sentence, starting at $s' = 0$, is $\theta p(e^*)(1 + c)[1 - R(p(e^*)w_L)]$. Therefore, if the public cost of imprisonment c is high enough, the marginal cost of imprisonment at $s' = 0$ will exceed the marginal benefit of imprisonment, and it will be optimal to rely exclusively on fines to sanction offenders. ■

It is not surprising that it will be socially desirable to rely exclusively on fines if the cost of imprisonment is sufficiently great. The fine should be as high as possible because, by well-

known logic, this allows the probability of detection to be lowered, saving enforcement costs. But the fine cannot exceed the wealth level of low-wealth individuals because, to induce high-wealth individuals to pay a higher fine, it would be necessary to impose an (alternative) imprisonment sentence on low-wealth individuals; by assumption, an imprisonment sentence would not be desirable, given its cost. Given a fine equal to the wealth level of low-wealth individuals, everyone can be induced to pay the fine.

The explanation why both groups are underdeterred is based on a familiar argument. Suppose that the probability of detection is such that the expected fine equals the harm. If the probability is lowered, there is no first-order effect on social welfare due to the lower level of deterrence because the marginal individuals who are induced to commit the offense obtain a benefit equal to the harm. But lowering the probability saves enforcement costs. Hence, some underdeterrence is socially desirable.

IV. Comparison to the Case of Observable Wealth

If wealth is observable, the enforcement authority can choose the mix of sanctions for each group separately. In particular, to induce high-wealth individuals to pay more than low-wealth individuals, the enforcement authority no longer needs to employ an alternative imprisonment sentence and make the burden of the sanctions for low-wealth offenders at least as high as that for high-wealth offenders. If the state can observe wealth, it can simply command that individuals in each group pay any amount up to their wealth level and bear any corresponding imprisonment sentence.

It would seem, therefore, that information about wealth levels would be socially valuable. Perhaps surprisingly, I will show in this section that such information may or may not be helpful.

The key point is that information about wealth is useful only if the enforcement authority would want to impose higher sanctions on high-wealth individuals than on low-wealth individuals, for if this is the case, high-wealth individuals would pretend to be low-wealth individuals if wealth could not be observed. Conversely, if the enforcement authority would want to impose lower sanctions on high-wealth individuals than on low-wealth individuals, high-wealth individuals will voluntarily bear such sanctions. The following propositions show that both of these possibilities can occur.¹⁵

Proposition 4. If wealth is observable, suppose that the optimal enforcement system would rely exclusively on fines to sanction offenders. In this case, if wealth is not observable,

(a) social welfare would be lower; and

(b) the optimal enforcement system might rely exclusively on fines; or

(c) the optimal enforcement system might employ alternative imprisonment sentences as well as fines.

Proof: (i) Let f_L and f_H be the fines paid by low-wealth and high-wealth individuals, respectively, when wealth is observable. I show in this step that $f_L^* = w_L$.¹⁶ Assume, to the contrary, that $f_L^* < w_L$, and first suppose that $f_H^* < w_H$. Then there would exist a $p(e) < p(e^*)$, an $f_L > f_L^*$, and an $f_H > f_H^*$ such that $p(e)f_L = p(e^*)f_L^*$ and $p(e)f_H = p(e^*)f_H^*$. Since the behavior of low-wealth and high-wealth individuals will not have changed, but enforcement costs will have declined, $f_L^* < w_L$ could not have been optimal. Now suppose that $f_L^* < w_L$ and $f_H^* = w_H$. Observe

¹⁵ The first proposition below corresponds to Proposition 3, in which fines are used alone, while the second proposition below corresponds to Proposition 1, in which an imprisonment sentence and a monetary sanction are imposed on both low-wealth and high-wealth offenders. For brevity, I do not consider the case corresponding to Proposition 2, in which an imprisonment sentence and a monetary sanction are imposed on low-wealth offenders, but only a fine is imposed on high-wealth offenders. In this case, information about wealth may or may not be valuable, essentially for the reasons illustrated by the other cases.

first that it must be that $p(e^*)f_H^* \leq h$; otherwise, the resulting overdeterrence could be eliminated at no cost by reducing f_H . Since $w_L < w_H$, it must be that $p(e^*)w_L < h$. Thus, the underdeterrence associated with $f_L^* < w_L$ could be reduced at no cost by raising f_L , contradicting the presumed optimality of f_L^* . Hence, it must be that $f_L^* = w_L$.

(ii) I demonstrate in this step that $f_H^* > w_L$. Assume otherwise, that $f_H^* \leq w_L$. Suppose first that $f_H^* < w_L$. Note that $p(e^*)f_L^* = p(e^*)w_L \leq h$, since otherwise the resulting overdeterrence of low-wealth individuals could be eliminated at no cost by reducing f_L . This implies that $p(e^*)f_H^* < h$. But then the underdeterrence of high-wealth individuals could be eliminated at no cost by raising f_H , contradicting the presumed optimality of $f_H^* < w_L$. Now suppose $f_H^* = w_L$. Then by step (iii) of the proof of Proposition 3, $p(e^*)$ must be such that $p(e^*)w_L < h$. But this implies that the underdeterrence of the high-wealth individuals could be reduced at no cost by raising f_H , contradicting the presumed optimality of $f_H^* = w_L$. Therefore, $f_H^* > w_L$.

(iii) If wealth is not observable, it is now straightforward to see that social welfare must be lower than when wealth is observable. If fines are used alone when wealth is not observable, recall from Proposition 3 that all offenders pay $f^* = w_L$. This outcome is feasible when wealth is observable and fines are used alone, but step (ii) shows that it is not optimal. If imprisonment sentences are used when wealth is not observable, the resulting outcome also must be inferior because the same outcome could have been chosen when wealth is observable, but the premise of the present proposition is that the optimal enforcement system when wealth is observable relies exclusively on fines. This establishes part (a).

(iv) Clearly, part (b) holds if the public cost of imprisonment c is sufficiently high.

¹⁶ The results in this and the next step were demonstrated by Polinsky and Shavell (1984, pp. 96-97). Because the proofs are short, I reproduce them here for completeness.

(v) To establish part (c), let SW_O^F , SW_U^F , and SW_U^I be the limiting values of social welfare as the fraction of low-wealth individuals θ goes to zero when, respectively, wealth is observable and fines are used alone, wealth is unobservable and fines are used alone, and wealth is unobservable and an alternative imprisonment sentence is used in addition to fines. I claim that $SW_U^F < SW_U^I = SW_O^F$. It is obvious that $SW_U^F < SW_O^F$ since, if wealth is unobservable, all offenders pay a fine of w_L , whereas if wealth is observable, high-wealth offenders pay a fine of $f_H^* > w_L$. If wealth is unobservable and an alternative imprisonment sentence is used, set $s' = f_H^* - w_L$ and $f = f_H^*$. High-wealth offenders will pay $f = f_H^*$ (since $f = w_L + s'$) while low-wealth offenders will pay w_L and bear an imprisonment sentence of $s' = f_H^* - w_L$. In the limit, as θ goes to zero, the cost of imprisonment goes to zero. Thus, if the enforcement authority picks the same probability of detection when wealth is unobservable as when wealth is observable, $SW_U^I = SW_O^F$. This result holds regardless of the public cost of imprisonment c . Thus, even if fines are used alone when wealth is observable because of the high cost of imprisonment, it will be optimal to employ an alternative imprisonment sentence when wealth is unobservable if θ is sufficiently low.¹⁷ This proves part (c). ■

A notable aspect of the present proposition is that, if wealth is unobservable, it may be worthwhile to use an alternative imprisonment sentence to induce high-wealth offenders to pay more than low-wealth offenders even though imprisonment is so costly that that it would not be used if wealth were observable. In other words, when wealth is not observable, it may be desirable to impose a costly sanction — imprisonment sentences — on low-wealth offenders in

¹⁷ There is no inconsistency between the premise of Proposition 3 — that if the public cost of imprisonment c is sufficiently high, it will be optimal to rely exclusively on fines — and the result here that if the fraction of low-wealth individuals θ is low enough, it will be optimal to employ an alternative imprisonment sentence regardless of the cost of imprisonment. The premise of Proposition 3 involves raising c , holding θ constant, whereas the claim here involves lowering θ , holding c constant.

order to better deter high-wealth offenders through a cheap sanction — fines. If wealth were observable, it would not be necessary to incur this cost.

Proposition 5. If wealth is observable, suppose that the optimal enforcement system would rely both on imprisonment sentences as well as fines to sanction offenders. In this case, if wealth is not observable,

(a) social welfare is not affected; and

(b) the optimal enforcement system is the same.

Proof: Let s_L and s_H be the imprisonment sentences borne by low-wealth and high-wealth individuals, respectively, when wealth is observable. The premise of this proposition is that $s_L^* > 0$ and $s_H^* > 0$. Clearly, then, it must be that $f_L^* = w_L$ and $f_H^* = w_H$ (or the fine could be raised and the imprisonment sentence lowered). The argument used in step (v) of the proof of Proposition 1 implies that $w_L + s_L^* > w_H + s_H^*$. Hence, when wealth is not observable, if $s = s_H^*$, $s' = s_L^*$, and enforcement expenditures are the same as those made when wealth is observable, the outcome can duplicate that when wealth is observable. Clearly, this is the social-welfare-maximizing outcome when wealth is unobservable, proving parts (a) and (b). ■

V. Prior Literature

To my knowledge, the only prior article to have systematically studied optimal sanctions under the assumption that an offender's level of wealth is unobservable is Levitt (1997).¹⁸ His model differs in significant respects, however, from mine. In Levitt's framework, "rich" and

¹⁸ Several articles have studied optimal sanctions when variations in wealth are assumed to be costlessly observable. See Friedman (1981), Polinsky and Shavell (1984, pp. 95-98), Polinsky and Shavell (1991), and Garoupa (1998, pp. 484-87). (Garoupa examines the case in which the enforcement authority costlessly observes an underestimate of individuals' wealth levels.)

“poor” individuals are distinguished not in terms of their wealth levels (both types are assumed to be capable of paying any fine the state imposes), but instead in terms of their disutility from time in jail (rich individuals suffer greater disutility). He also assumes that the benefit from committing the offense is the same for every individual of a given type, but differs between the types, and that the benefit is less than the harm from the offense, so that ideal deterrence involves complete deterrence. In his formal analysis, Levitt considers the choice between an imprisonment sentence and a fine, not the optimal combination of the two, as I do here (though he does discuss this issue informally). My results show that whenever imprisonment sanctions are employed, monetary sanctions should be too.

Levitt’s main point is that, in contrast to the situation when individuals’ wealth levels can be costlessly observed, the availability of fines as an alternative to imprisonment might not result in higher social welfare. This is because, in his analysis, the introduction of a fine as an alternative to a jail term cannot increase deterrence — if the fine were more burdensome, the offender would choose the jail term. Consequently, Levitt suggests that the case for using imprisonment sentences as sanctions is stronger than is generally appreciated.

In an article slightly predating Levitt’s, Chu and Jiang (1993) examine optimal sanctions when there are three types of individuals in terms of wealth levels and a continuum of offenses that individuals can commit, each corresponding to a different level of harm. Because Chu and Jiang assume that the fine is proportional to the level of harm, they in effect assume that the enforcement authority cannot observe offenders’ wealth levels, though this is not the focus of their attention. Their main point is that, because of marginal deterrence considerations and differences among individuals in their responses to imprisonment (wealthier individuals have a higher opportunity cost of time in prison), it may be desirable to use imprisonment sentences

combined with less-than-maximal fines. Like Levitt, they emphasize the desirability of imprisonment sentences relative to fines. My analysis differs from theirs in that I do not impose any restrictions on the choice of fines (aside from wealth constraints), whereas, as noted, they assume that the fine is proportional to the level of harm. Our conclusions differ as well. I find, contrary to Chu and Jiang's result, that imprisonment sentences should not be used unless the party subject to the sentence also makes a monetary payment equal to his wealth level.¹⁹

In a companion article — Polinsky (2004) — I derive optimal fines when an offender's level of wealth can be determined after a costly audit. In contrast to the present analysis, I do not consider imprisonment sanctions. The focus of my other article is on characterizing the optimal audit rate and on deriving the optimal fine for misrepresenting one's wealth level, issues ignored here.

VI. Concluding Remarks

The key result from my analysis is that, even though offenders' wealth levels cannot be observed, the optimal mix of sanctions may induce high-wealth offenders to pay more than low-wealth offenders because they would bear more burdensome (or at least as burdensome) sanctions if they did not. In particular, this result holds whenever imprisonment sentences are employed, even if only for offenders who do not pay the fine.

I made several simplifying assumptions in order to keep the analysis tractable. The most notable ones are that individuals bear the same disutility from time in prison regardless of their

¹⁹ Also of some relevance to the present article is the analysis by Lott (1987) of the question whether individuals should be allowed to spend freely on their defense in criminal cases. In his model, offenders bear the same imprisonment sentence regardless of their wealth level, and higher-wealth individuals are assumed to suffer greater disutility from the sentence. Lott's argument is that allowing higher-wealth individuals to spend more on

wealth level; that individuals are risk-neutral with respect to fines and imprisonment; and that there are only two levels of wealth. I discuss these briefly here and suggest why modifying them would not change the point that if optimally chosen sanctions include imprisonment sentences, such sanctions will induce higher-wealth individuals to pay more than lower-wealth individuals.

Suppose that high-wealth individuals have a greater distaste for jail time than low-wealth individuals, as the other authors discussed in Section V assumed. The following argument shows that this assumption would not affect my principal conclusion. Given this assumption, suppose the sanctions were such that both groups chose to pay an amount of money equal to the wealth level of low-wealth individuals w_L and to bear the alternative imprisonment sentence s' . This outcome could not be optimal, for it would be possible to choose some fine $f > w_L$ and some imprisonment sentence $s < s'$ such that high-wealth individuals are indifferent between the primary (f and s) and alternative (w_L and s') sanctions. Hence, they could be induced to choose the primary sanctions, resulting in an increase in social welfare due to a reduction in the imprisonment sentence borne by high-wealth individuals. In other words, optimally chosen sanctions still will induce high-wealth individuals to pay a fine that exceeds the wealth level of low-wealth individuals.

If individuals are not risk neutral with respect to fines or imprisonment, the optimal sanctions would change, but the same argument used in the preceding paragraph demonstrates that my main result would not be affected. Suppose, given a different set of assumptions about the risk preferences of low-wealth and high-wealth individuals (but assuming that all high-wealth individuals have the same preferences and all low-wealth individuals have the same preferences),

their defense reduces the overdeterrence that otherwise would occur. For further discussion of this point, see Garoupa and Gravelle (2003).

that the sanctions were such that both groups chose to pay w_L and to bear the alternative imprisonment sentence s' . Again, this outcome could not be optimal, for it would be possible to choose some fine $f > w_L$ and some imprisonment sentence $s < s'$ such that high-wealth individuals would be just willing to choose the primary sanctions, thereby lowering imprisonment costs.

If wealth levels varied continuously among individuals, the optimal enforcement system would include a menu of choices for offenders — various combinations of fines and imprisonment sentences. Based on reasoning employed in this article, the optimal menu should have the property that the sum of the sanctions declines (or at least does not rise) with the fine paid, which would induce individuals to pay the highest fine possible. Specifically, start with a menu of sanctions — fines and imprisonment sentences — such that each combination in the menu generates the same level of deterrence. Obviously, the higher the fine, the lower the imprisonment sentence. The social benefit from raising the imprisonment sentences from these initial levels would always be greater the higher is the initial sentence since society saves more on imprisonment costs by deterring an individual who would have borne a longer sentence. This implies that in the optimal menu the sum of the sanctions declines with the fine paid, causing each offender to pay the highest fine he is capable of paying.

References

- Chu, C.Y. Cyrus, and Neville Jiang “Are Fines More Efficient than Imprisonment?” *Journal of Public Economics*, Vol. 51, No. 3 (July 1993), pp. 391-413.
- Friedman, David, “Reflections on Optimal Punishment, Or: Should the Rich Pay Higher Fines?” *Research in Law and Economics*, Vol. 3 (1981), pp. 185-205.
- Garoupa, Nuno, “Optimal Law Enforcement and Imperfect Information when Wealth Varies Among Individuals,” *Economica*, Vol. 65 (1998), pp. 479-90.
- Garoupa, Nuno, and Hugh Gravelle, “Efficient Deterrence Does Not Require that the Wealthy Should Be Able to Buy Justice,” *Journal of Institutional and Theoretical Economics*, Vol. 159, No. 3 (September 2003), pp. 545-552.
- Levitt, Steven, “Incentive Compatibility Constraints as an Explanation for the Use of Prison Sentences Instead of Fines,” *International Review of Law and Economics*, Vol. 17, No. 2 (June 1997), pp. 179-92.
- Lott, Jr., John R., “Should the Wealthy Be Able to ‘Buy Justice’?,” *Journal of Political Economy*, Vol. 95, No. 6 (December 1987), pp. 1307-16.
- Polinsky, A. Mitchell, “Optimal Fines and Auditing When Wealth is Costly to Observe,” Working Paper No. 286, John M. Olin Program in Law and Economics, Stanford Law School, August 2004.
- Polinsky, A. Mitchell, and Steven Shavell, “The Optimal Use of Fines and Imprisonment,” *Journal of Public Economics*, Vol. 24, No. 1 (June 1984), pp. 89-99.
- Polinsky, A. Mitchell, and Steven Shavell, “A Note on Optimal Fines When Wealth Varies Among Individuals,” *American Economic Review*, Vol. 81, No. 3 (June 1991), pp. 618-21.