

LE
8/3 8:45
Polinsky

Preliminary Draft
May 2002

OPTIMAL SANCTIONS WHEN WEALTH CAN BE AUDITED

A. Mitchell Polinsky*

Abstract: In nearly all discussions of optimal sanctions it is assumed that an offender's level of wealth is known to the enforcement authority without cost. One notable exception is Levitt (1997), who made the polar opposite assumption that an offender's wealth level is private information that cannot be obtained by the enforcement authority at any cost. This article studies optimal sanctions in a more general setting, when an offender's level of wealth is private information that can be obtained by the enforcement authority after a costly audit. In the model employed, which assumes that parties are risk neutral, I derive the optimal audit rate, the optimal sanctions for misrepresenting one's wealth level, the optimal sanctions for the underlying offense as a function of offenders' reported wealth levels, and the optimal probability of detecting the offense. Both fines and imprisonment are considered as sanctions. I demonstrate that the optimal enforcement system can be characterized in one of two ways. In the first, only fines are used to control the offense and auditing of wealth levels is required to induce high-wealth individuals to pay a fine greater than that paid by low-wealth individuals. The fine for the offense generally is not maximal, though the fine and the imprisonment sentence for misrepresenting one's wealth level are maximal. In the second case, both fines and imprisonment sentences are used to control the offense. In this case the fine for the offense is maximal and the imprisonment sentence for individuals who do not pay this fine is sufficiently high that high-wealth individuals will prefer to pay the fine and bear a lower imprisonment sentence. Auditing is of no value in this case because it is not needed to induce high-wealth individuals to pay more than low-wealth individuals; imprisonment sentences serve, in effect, as a substitute for auditing.

* Stanford University and National Bureau of Economic Research. 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, Daniel Rubinfeld, and Steven Shavell.

I. Introduction

Almost without exception, prior analyses of optimal sanctions have assumed that an offender's level of wealth is costlessly observable by the enforcement authority. Knowing that it can observe wealth levels, the enforcement authority chooses the probability of detection and the sanctions to impose, fines and/or imprisonment. In contrast, Steven Levitt (1997) has analyzed the choice of sanctions when the offender's level of wealth is private information and cannot be discovered by the enforcement authority at any cost. The contribution of the present article is to consider optimal sanctions in a more general setting, in which the enforcement authority can determine an offender's level of wealth after a costly audit. Both the conventional analysis and the Levitt analysis are special cases of the present model — the former when the audit cost is zero, and the latter when the audit cost is sufficiently high.

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. The enforcement authority decides how much to invest in detecting offenders. If an offender is detected, he is sanctioned with a fine and, possibly, an imprisonment sentence. If he claims that he cannot pay the fine, a higher imprisonment sentence might be imposed on him. Additionally, he may be audited. If the audit determines that he misrepresented his wealth level, he can be sanctioned further, both by a fine and an imprisonment sentence. The enforcement authority's problem is to choose the following policy instruments to maximize social welfare: the probability of detecting the underlying offense; the fine and imprisonment sentence for committing the offense; the imprisonment sentence to impose on an offender who claims that he cannot pay the fine for the offense; the probability of auditing offenders who do not pay the fine for the offense; and the fine and imprisonment sentence to impose on offenders who, as a result of an audit, are found to have misrepresented their wealth level.

I demonstrate that the optimal enforcement system can be characterized in one of two ways.¹ In the first, only fines are used to control the offense and auditing of wealth levels is required to induce the high-wealth individuals to pay a fine greater than that paid by the low-wealth individuals. Although fines are usually treated as socially costless sanctions because they

¹ There is a third outcome, but because it is a hybrid of the two that I am about to describe, I omit reference to it here. It is discussed in detail in Section III below.

are mere transfers of wealth, they are no longer socially costless if auditing is required to induce higher-wealth individuals to pay higher fines. In effect, there is a marginal social cost incurred from imposing a higher fine due to the need for a higher audit rate to induce the payment of the fine. As a consequence, the fine for the offense generally should not be maximal, though it often will exceed the wealth level of the low-wealth individuals. In this enforcement system, high-wealth individuals are deterred from committing the offense more than low-wealth individuals because they are paying higher fines, but, for reasons I explain, are underdeterred relative to the first-best outcome.

Although fines alone are used to control the offense in this case, a fine as well as an imprisonment sentence should be used to sanction individuals who misrepresent their wealth level. Moreover, these sanctions should be maximal because higher sanctions for misrepresentation allow the audit probability to be lowered while still inducing high-wealth individuals to pay the fine for the offense. There is no offsetting increase in imprisonment costs because, as I demonstrate, high-wealth individuals will refrain from misrepresenting their wealth level in the optimal enforcement system, so these sanctions will not actually be imposed.

In the second characterization of the optimal enforcement system, both fines and imprisonment are used to control the offense. In this case I show that the fine for the offense is maximal, equal to the wealth level of the high-wealth individuals. Thus, the conventional result that fines should be exhausted before resorting to imprisonment still holds despite the fact that fines are socially costly in the sense described above. Additionally, and most importantly, I demonstrate that the optimal imprisonment sentence for individuals who do not pay the fine for the offense (that is, who claim to be low-wealth individuals) is sufficiently high that high-wealth individuals will strictly prefer to pay the fine for the offense and bear the corresponding imprisonment sentence. This means that auditing is of no value in this case because it is not needed to induce high-wealth individuals to pay more than low-wealth individuals. The optimal audit probability is zero. In this enforcement system, high-wealth individuals are deterred less than low-wealth individuals from committing the offense, the opposite of the result in the previous case. I also show that both groups may be underdeterred or overdeterred relative to the first-best outcome.

The preceding discussion shows that auditing of wealth levels and the imposition of higher imprisonment sentences on those who do not pay the fine for the offense may be viewed as alternative ways of ameliorating the disadvantages to the enforcement authority of not being able to costlessly observe individuals' wealth levels. In some circumstances, auditing wealth levels but not using imprisonment sentences to control the offense is the better way to induce high-wealth individuals to pay high fines. In other circumstances, employing imprisonment, with sufficiently high sentences on those who claim to be low-wealth individuals, is the preferred method. Thus, either because wealth levels are audited and misrepresentation is punished, or because high imprisonment sanctions are imposed on offenders who claim to have a low level of wealth, high-wealth individuals will not want to misrepresent their level of wealth in the optimal enforcement system.²

Section II presents the general framework used in the article. Section III derives the optimal enforcement system. Section IV discusses the prior literature in relation to my analysis. Section V provides some concluding remarks.

II. General Framework

In the model, risk-neutral 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.³ An individual who commits the offense is detected with some probability; this probability is determined by the enforcement expenditures of the state.⁴ If an offender is detected, he may be sanctioned with a fine and/or an imprisonment sentence. If he claims that he cannot pay the fine because he has insufficient wealth, he may be subject to a higher imprisonment sentence. Additionally, he may be audited to determine his wealth level. Audits cost the state a fixed amount of money. If an audit

² In the third outcome referred to in the preceding footnote, it may be optimal to both audit wealth levels and impose an imprisonment sentence on individuals who claim to have a low level of wealth. See Section III below.

³ I discuss in Section V below why my general results should continue to hold if there is a continuum of wealth levels.

⁴ The probability of detection is assumed to be independent of the individual's level of wealth. This makes sense in many, if not most, enforcement contexts, though there are exceptions (for example, detecting tax evasion).

determines that the individual was lying about his wealth level, he may be sanctioned further, by a fine and/or an imprisonment sentence. The fine imposed on an individual, whether for the offense or for misrepresenting his wealth level, cannot exceed his level of wealth. Fines per se are socially costless, though, as will be seen, they may have an associated auditing cost. I assume for simplicity that the maximum possible imprisonment sentence is the same for all individuals, that they are risk neutral with respect to imprisonment sentences, and that they bear the same disutility from a given sentence.⁵ Also, 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; $p'(e) > 0$; $p''(e) < 0$;
- f_0 = fine for committing the offense; $f_0 \geq 0$;⁶
- s_0 = imprisonment sentence for committing the offense
if the fine f_0 is paid; $s_0 \geq 0$;
- s'_0 = imprisonment sentence for committing the offense
if $w_L < f_0$ is paid; $s'_0 \geq 0$;
- \bar{s} = maximum possible imprisonment sentence; $\bar{s} > 0$;
- c = cost to the state per unit of imprisonment sentence;
 $c > 0$;
- q = probability of an audit; $q \geq 0$;

⁵ By risk neutrality with respect to imprisonment sentences I mean that individuals care only about the expected value of the sentence length. In Section V below I explain why my general results would not be affected if individuals are not risk neutral with respect to imprisonment sentences, or if they differ in terms of their disutility from time in prison.

⁶ This is the nominal fine. An individual might not have sufficient wealth to pay this fine.

k = cost to the state of an audit; $k > 0$;
 f_M = fine for misrepresenting one's wealth level; $f_M \geq 0$; and
 s_M = imprisonment sentence for misrepresenting one's wealth
 level; $s_M \geq 0$.

If an individual is found to have misrepresented his wealth level, the sanctions f_M and s_M are imposed instead of, not in addition to, the sanctions that otherwise would be applicable. 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 an individual. Thus, an imprisonment sentence of length s imposes a cost on an individual equal to s . The population is normalized to one.

The state is assumed to know the distribution of wealth levels among the population but not, without an audit, the wealth level of a particular individual. However, the state does know, without having to conduct an audit, that everyone in the population has wealth of at least w_L . I assume, therefore, that the state can costlessly collect any fine up to and including w_L .⁷

Behavior of low-wealth individuals. First consider the behavior of a low-wealth individual. 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.⁸

In other words, b_L represents the degree to which low-wealth individuals are deterred from committing the offense — the higher b_L is, the greater the level of deterrence.

If the fine for the offense f_O is less than or equal to a low-wealth individual's level of wealth w_L , then the individual can be made to pay the fine f_O without an audit (for the reason discussed above) and bear the imprisonment sentence s_O . In this case,

$$b_L = p(e)(f_O + s_O). \quad (1)$$

⁷ This assumption could be justified as follows. For any fine up to and including w_L the enforcement authority could threaten to impose an additional imprisonment sentence (that is, in addition to the sentence that is optimal assuming the fine is paid) equal to the fine. Then every individual would be willing to pay this fine and the additional sentence would not be imposed. Obviously, this argument presumes that the upper bound on the length of the imprisonment sentence \bar{s} is not binding.

⁸ I assume without loss of generality that he will commit the offense if he is indifferent.

If f_O exceeds w_L , he will claim that he cannot pay the full fine, and his claim will be found to be valid if he is audited. He therefore will pay w_L and bear the imprisonment sentence s'_O in which case

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

Behavior of high-wealth individuals. Similarly, let

b_H = critical value of benefit below which a high-wealth individual will not commit the offense and at and above which he will.

If the fine f_O is less than or equal to w_L , a high-wealth individual also will pay f_O and bear the imprisonment sentence s_O . Thus, b_H equals the right-hand side of (1).

If f_O exceeds w_L (but not w_H ⁹) and the high-wealth individual does not misrepresent his wealth level, he will pay this fine and bear the imprisonment sentence s_O . Then

$$b_H = p(e)(f_O + s_O). \quad (3)$$

If he does misrepresent his wealth level, he either will pay w_L and bear the imprisonment sentence s'_O if he is not audited or pay f_M and bear the sentence s_M if he is audited. In this case,

$$b_H = p(e)[(1 - q)(w_L + s'_O) + q(f_M + s_M)]. \quad (4)$$

Thus, a high-wealth individual will misrepresent his wealth level if and only if¹⁰

$$(1 - q)(w_L + s'_O) + q(f_M + s_M) < f_O + s_O. \quad (5)$$

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

If the fine for the offense f_O is less than or equal to w_L , then $b_L = b_H = p(e)(f_O + s_O)$, and social welfare is

$$\frac{\int_0^{\infty} [b - h - p(e)(1 + c)s_O]r(b)db - e}{p(e)(f_O + s_O)} \quad (6)$$

Everyone whose benefit level equals or exceeds $p(e)(f_O + s_O)$ 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

⁹ 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.

¹⁰ I assume without loss of generality that he will not misrepresent his wealth if he is indifferent.

imprisonment sentence s_O . The social cost of imprisonment is the sum of the private cost s_O and public cost cs_O . The fine individuals pay does not enter social welfare directly because it is a transfer payment. The state also incurs enforcement expenditures e . No auditing is necessary in this case because the state knows that everyone can pay the fine.

If the fine for the offense f_O exceeds w_L and high-wealth individuals do not misrepresent their wealth level, social welfare is

$$\begin{aligned} & \theta \int_{b_L}^{\infty} \{b - h - p(e)[qk + (1 + c)s'_O]\} r(b) db \\ & + (1 - \theta) \int_{b_H}^{\infty} [b - h - p(e)(1 + c)s_O] r(b) db - e, \end{aligned} \quad (7)$$

where b_L is given by (2) and b_H by (3). The interpretation of (7) is similar to that of (6), though now low-wealth and high-wealth individuals are deterred to different degrees from committing the offense. This is because low-wealth individuals cannot pay the fine f_O and are therefore subject to a different imprisonment sentence. Also, conditional on detection of the offense, individuals who do not pay the fine f_O are audited with probability q . In the present case, only low-wealth individuals will be audited.

If the fine for the offense f_O exceeds w_L and high-wealth individuals do misrepresent their wealth level, social welfare is

$$\begin{aligned} & \theta \int_{b_L}^{\infty} \{b - h - p(e)[qk + (1 + c)s'_O]\} r(b) db \\ & + (1 - \theta) \int_{b_H}^{\infty} \{b - h - p(e)[qk + (1 + c)((1 - q)s'_O + qs_M)]\} r(b) db - e, \end{aligned} \quad (8)$$

where b_L is given by (2) and b_H by (4). Now both high-wealth and low-wealth individuals are audited. Also, the imprisonment sentence borne by a high-wealth individual depends on whether he is audited.

The state's problem. The state's problem is to maximize social welfare through the choice of the fine f_O and imprisonment sentence s_O for the offense, the imprisonment sentence for the offense s'_O if the individual claims to be a low-wealth individual and pays only w_L , enforcement expenditures e , the audit probability q , and the fine f_M and imprisonment sentence

s_M applicable if an audit determines that an individual was misrepresenting his wealth level.¹¹ The fines actually paid by low-wealth and high-wealth individuals cannot exceed their respective levels of wealth. The imprisonment sentences cannot exceed the upper bound \bar{s} .

Asterisks will be 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

In this section I describe the optimal enforcement system through a series of propositions. Following the proof of each proposition is a brief informal discussion of the result.¹²

Proposition 1. *The optimal enforcement system results in high-wealth individuals paying the fine for the offense f_O rather than misrepresenting their level of wealth.*

Proof: If f_O is less than or equal to the low-wealth level w_L , the enforcement authority can costlessly collect the fine from both low-wealth and high-wealth individuals (for the reason discussed in Section II). Assume, therefore, that $f_O > w_L$. A high-wealth individual will misrepresent his wealth level if and only if (5) holds, that is, if and only if $(1 - q)(w_L + s_O) + q(f_M + s_M) < f_O + s_O$. If he does misrepresent his wealth level, $b_H = p(e)[(1 - q)(w_L + s_O) + q(f_M + s_M)]$. I will show that if (5) holds, it is possible to choose f_O and s_O in such a way that social welfare rises because auditing costs decline. Specifically, set $f_O = (1 - q)w_L + qf_M$ and $s_O = (1 - q)s_O + qs_M$. With these values, $(1 - q)(w_L + s_O) + q(f_M + s_M) = f_O + s_O$, so high-wealth individuals will pay the fine f_O and bear the imprisonment sentence s_O . Although b_H now equals $p(e)(f_O + s_O)$, b_H is unaffected because $f_O + s_O = (1 - q)(w_L + s_O) + q(f_M + s_M)$. The social cost of imprisonment also is unaffected, because $s_O = (1 - q)s_O + qs_M$. The only effect on social welfare is to eliminate the cost of auditing high-wealth individuals; compare (7) and (8). Thus, in the optimum it must be that $(1 - q)(w_L + s_O) + q(f_M + s_M) \geq f_O + s_O$, in which case a high-wealth individual will prefer to pay the fine for the offense f_O rather than misrepresent his wealth level. ■

¹¹ Because my objective is to derive the socially optimal enforcement policy, I will ignore the issue of whether the state's policy is credible in a game-theoretic sense.

¹² Although some of the proofs convey useful intuition about the result, they may be skipped with little loss of continuity.

The essence of the argument behind this proposition can be stated as follows. If the enforcement system made it attractive to high-wealth individuals to misrepresent their wealth level, it would be possible to lower the sanctions for the offense without reducing the expected liability of high-wealth individuals, since the high-wealth individuals were not bearing these sanctions. In particular, the sanctions could be lowered to a level that just induces the high-wealth individuals to not misrepresent their wealth level. This is preferable because the cost of auditing these individuals is saved, while the expected liability of high-wealth individuals is unaffected.

Although the proof of the proposition employs sanctions for the offense that lead high-wealth individuals to be indifferent concerning whether to misrepresent their wealth level, the optimal enforcement system is not necessarily characterized by this feature. Proposition 1 only rules out the case in which (5) holds. As will be seen below, the optimal enforcement system might cause high-wealth individuals to strictly prefer to pay the fine for the offense.

Note that Proposition 1 implies that if the optimal fine for the offense exceeds w_L , the relevant expression for social welfare is (7) rather than (8).

Proposition 2. The optimal fine and the optimal imprisonment sentence for misrepresenting one's wealth level are maximal: $f_M^ = w_H$ and $s_M^* = \bar{s}$.*

Proof: This claim is relevant only if the fine for the offense f_O exceeds w_L since otherwise the enforcement authority can costlessly collect the fine. Suppose f_M and s_M are not maximal and let $q_1 > 0$ be the original value of the audit probability. By Proposition 1, it must be that $(1 - q_1)(w_L + s_O) + q_1(f_M + s_M) \geq f_O + s_O$, so high-wealth individuals pay the fine for the offense f_O and bear the corresponding imprisonment sentence s_O .

If $f_M + s_M \geq w_L + s_O$ raise f_M to w_H and s_M to \bar{s} and lower the audit probability to q_2 such that $(1 - q_2)(w_L + s_O) + q_2(w_H + \bar{s}) = (1 - q_1)(w_L + s_O) + q_1(f_M + s_M)$. If $f_M + s_M < w_L + s_O$ raise f_M to w_H and s_M to \bar{s} and lower the audit probability to $q_2 = 0$, in which case $(1 - q_2)(w_L + s_O) + q_2(w_H + \bar{s}) = w_L + s_O$. Since the expected liability of a high-wealth individual who misrepresents his wealth level either remains the same or rises, high-wealth individuals continue to pay the fine f_O and bear the imprisonment sentence s_O . Hence, the relevant expression for social welfare is (7). It is clear from (7) that the only effect of raising f_M

and s_M and lowering q in the way described is to lower auditing costs and thereby raise social welfare.*

It is well known that if individuals are risk neutral and have the same level of wealth, the optimal fine for the offense is maximal; if it were not, the fine could be raised and the probability of detection lowered without sacrificing deterrence but saving enforcement costs. Similar logic is used to derive the present result. It is possible to raise the sanctions for misrepresenting one's wealth level and to lower the probability of an audit so as to keep the resulting expected liability the same, or to raise it, while reducing auditing costs. Here, however, there is a potentially countervailing consideration. The expected imprisonment sentence, $(1 - q)s_O^* + qs_M$, could rise as a result of this change and result in higher social costs of imprisonment. But because the optimal enforcement system induces high-wealth individuals to not misrepresent their wealth level, imprisonment sentences for misrepresenting one's wealth level are not actually imposed, so this consideration is not applicable.

The next three propositions characterize the optimal mix of fines and imprisonment to control the offense, as well as the associated optimal rate of auditing. As will be seen, it is always optimal to use fines to some extent. Whether imprisonment should be used as well to control the offense depends on the public cost of imprisonment. For reasons that will become clear, auditing tends to be less valuable when imprisonment is used.

I assume throughout the remainder of this section that if imprisonment is employed to control the offense, the optimal imprisonment sentences s_O^* and s_M^* are not constrained by the upper bound on the length of the imprisonment sentence \bar{s} . As will be evident below, there is no general reason to expect the optimal imprisonment sentences for the offense to be maximal, as there was for the imprisonment sentence for misrepresenting one's wealth level. By assuming that s_O^* and s_M^* are less than \bar{s} , I avoid having to consider some corner solutions that are not particularly interesting.

Proposition 3. If the public cost of imprisonment c is sufficiently high, fines alone should be used to sanction offenders. Then

(a) the optimal probability of an audit may be zero or positive, but is bounded below one; specifically, $0 \leq q^ \leq (w_H - w_L)/(w_H - w_L + \bar{s}) < 1$;*

(b) the optimal audit probability is positive if the cost of auditing is sufficiently low;

(c) the optimal fine for the offense may be as low as the wealth level of low-wealth individuals or as high as the wealth level of high-wealth individuals; specifically, $f^{\circ} = (1 - q)w_L + q(w_H + \bar{s})$; and

(d) if the audit probability is positive, the expected liability of low-wealth individuals is less than the expected liability of high-wealth individuals, with both groups being underdeterred: $b_L < b_H < h$.

Proof: That imprisonment should not be used to sanction offenders if the public cost of imprisonment c is high enough is obvious from (7) since, for any positive value of s_O or s_D , social welfare goes to $-\infty$ as c goes to ∞ .

I first prove that the optimal fine for the offense f° is at least as high as the wealth level of low-wealth individuals w_L . Suppose $f_O < w_L$. Then the relevant expression for social welfare is (6) with $s_O = 0$. Let e_1 be the original value of enforcement expenditures. Raise f_O to w_L and lower enforcement expenditures to e_2 such that $p(e_2)w_L = p(e_1)f_O$. It is clear from (6) that social welfare rises because enforcement expenditures decline. Thus, $f^{\circ} \geq w_L$.

It is convenient to prove next that $f^{\circ} = (1 - q)w_L + q(w_H + \bar{s})$. If fines alone are used to control the offense, Propositions 1 and 2 imply, from (5), that $(1 - q)w_L + q(w_H + \bar{s}) \geq f_O$. If this expression were to hold with strict inequality, it would be possible to lower q until it holds with equality, while still inducing high-wealth individuals to pay the fine for the offense f_O . It is easily seen from (7) that social welfare would rise because auditing costs decline. Thus, $f^{\circ} = (1 - q)w_L + q(w_H + \bar{s})$.

It is clear that $f^{\circ} = (1 - q)w_L + q(w_H + \bar{s})$ is increasing in q , equals w_L if $q = 0$, and equals w_H if $q = (w_H - w_L)/(w_H - w_L + \bar{s}) < 1$. Since high-wealth individuals will pay w_H if $q = (w_H - w_L)/(w_H - w_L + \bar{s})$, any higher q obviously is inferior because it merely results in higher auditing costs. Thus, $q^* \leq (w_H - w_L)/(w_H - w_L + \bar{s})$.

I next show that q^* can be as low as zero or as high as $(w_H - w_L)/(w_H - w_L + \bar{s})$. If $s_D = s_O = 0$, the derivative of social welfare (7) with respect to the audit probability q is

$$-\theta p(e)k[1 - R(b_L)] - (1 - \theta)[b_H - h]r(b_H)[p(e)(w_H - w_L)], \quad (9)$$

where $b_L = p(e)w_L$ and $b_H = p(e)f_O = p(e)[(1 - q)w_L + q(w_H + \bar{s})]$. The first term in (9) is the cost of raising q , in terms of increased auditing costs of low-wealth individuals. The second term is the benefit (if $b_H < h$, which will be demonstrated two paragraphs below) of raising q in terms of

greater deterrence of high-wealth individuals from committing the offense. It is clear that if k is sufficiently high, (9) will be negative for all q between 0 and $(w_H - w_L)/(w_H - w_L + \bar{s})$, in which case $q^* = 0$. To see that $q^* = (w_H - w_L)/(w_H - w_L + \bar{s})$ is possible, assume that $w_H < h$. Then for all $q \leq (w_H - w_L)/(w_H - w_L + \bar{s})$, $b_H - h < 0$ and $-(1 - \theta)[b_H - h]r(b_H)[p(e)(w_H - w_L)] > 0$. If k is sufficiently low, (9) will be positive for all q between 0 and $(w_H - w_L)/(w_H - w_L + \bar{s})$, in which case $q^* = (w_H - w_L)/(w_H - w_L + \bar{s})$. Note that since $f_O^* = (1 - q)w_L + q(w_H + \bar{s})$, $f_O^* = w_L$ if $q^* = 0$ and $f_O^* = w_H$ if $q^* = (w_H - w_L)/(w_H - w_L + \bar{s})$.

The preceding establishes parts (a) and (c). To demonstrate part (b), I will show that if auditing is costless, the optimal audit probability is positive. Suppose $q = 0$. Then $f_O^* = w_L$ and the relevant expression for social welfare is (6) with $s_O = 0$. The first-order condition with respect to enforcement expenditures e can be written as

$$-[p(e)w_L - h]r(p(e)w_L)p'(e)w_L = 1. \quad (10)$$

For (10) to hold, it must be that $p(e)w_L < h$, that is, there is some underdeterrence. Clearly, if it were costless to do so, it would be desirable to raise the fine for the offense paid by high-wealth individuals up to f_O such that $p(e)f_O = h$, that is, up to $f_O = h/p(e)$. To induce high-wealth individuals to pay this fine would require an audit probability q satisfying $h/p(e) = (1 - q)w_L + q(w_H + \bar{s})$, that is, $q = (h - p(e)w_L)/p(e)(w_H - w_L + \bar{s}) > 0$ (the inequality follows from having shown that $p(e)w_L$ is less than h). This q might exceed the previously derived upper bound on the audit probability, $(w_H - w_L)/(w_H - w_L + \bar{s})$. Thus, if auditing were costless, $q^* = \min[(h - p(e)w_L)/p(e)(w_H - w_L + \bar{s}), (w_H - w_L)/(w_H - w_L + \bar{s})] > 0$. By continuity, the optimal audit probability approaches this value as the cost of auditing k goes to zero, which establishes part (b).

Finally, to prove part (d) observe that if $q > 0$, $f_O = (1 - q)w_L + q(w_H + \bar{s}) > w_L$, which implies that $b_L < b_H$. To see that $b_H < h$, first suppose that q^* is an interior optimum; in this case (9) equals zero, which can occur only if $b_H < h$. If $q^* = (w_H - w_L)/(w_H - w_L + \bar{s})$, (9) must be non-negative, which also can occur only if $b_H < h$. Hence, $b_L < b_H < h$.

A key to understanding Proposition 3 is the observation that, in the optimal enforcement system when imprisonment sentences are not used to control the offense, high-wealth individuals are indifferent between paying the fine f_O and misrepresenting their wealth level. This result follows in essence from two observations. If high-wealth individuals strictly preferred to pay the

fine for the offense rather than misrepresenting their wealth level, the audit probability could be lowered while still deterring misrepresentation. Conversely, if high-wealth individuals strictly preferred to misrepresent their wealth level, the fine for the offense could be lowered so as to make them indifferent between misrepresenting their wealth level and not. This is beneficial because lowering the fine to this extent does not affect the expected liability of high-wealth individuals, but reduces auditing costs by inducing them to pay the fine for the offense rather than misrepresent their wealth level.¹³ In the optimal enforcement system, therefore, it must be the case that high-wealth individuals are indifferent between paying the fine for the offense and misrepresenting their wealth level. As explained in the proof, this means that $f_O = (1 - q)w_L + q(w_H + \bar{s})$. Accordingly, it will be useful to treat the audit probability q as a policy instrument and to set f_O according to this formula.

If the audit probability were zero, the formula just stated implies that the fine for the offense would equal the wealth level of low-wealth individuals. Both low-wealth and high-wealth individuals would pay this fine. The optimal probability of detection of the offense would be such that the expected fine is less than the harm caused (for if the expected fine equaled the harm, lowering the probability of detection would reduce enforcement costs but have no first-order effect on social welfare in terms of the harm and benefit from the offense).¹⁴ In this case, then, both groups would be underdeterred to the same extent. Raising the audit probability above zero is beneficial because it results in high-wealth individuals paying a higher fine for the offense, thereby reducing their underdeterrence. If the cost of auditing is not too great, it will be desirable to do this, including possibly to raise the audit probability to the level that induces high-wealth individuals to pay a fine equal to their wealth level. However, if the cost of auditing is sufficiently great, then the optimal audit probability will be zero and the optimal fine for the offense will equal the wealth-level of low-wealth individuals.

Note that because of the threat of an imprisonment sentence for misrepresenting one's wealth level, in addition to the maximal fine, the audit probability needed to induce high-wealth individuals to pay a fine for the offense equal to their wealth level is less than one.

¹³ This argument is a special case of the point made in the discussion immediately following the proof of Proposition 1.

¹⁴ This result is well known in the standard literature on enforcement in which auditing of offenders' wealth levels is ignored.

I explained two paragraphs above why both low-wealth and high-wealth individuals will be underdeterred to the same extent if the optimal audit probability is zero. If the optimal audit probability is positive, however, high-wealth individuals will be paying a higher fine, equal to $f_O = (1 - q)w_L + q(w_H + \bar{s})$, than low-wealth individuals, and so will be deterred from committing the offense to a greater extent. They still will be underdeterred, however, for a reason analogous to that mentioned parenthetically in the discussion above. Specifically, if the audit probability were such that the expected fine for the offense borne by high-wealth individuals equaled the harm, lowering the probability of an audit would reduce auditing costs but have no first-order effect on social welfare in terms of the harm and benefit from the offense.

Proposition 4. *If the public cost of imprisonment c is sufficiently low, it may be optimal to use imprisonment to sanction offenders; that is, s_O^* may be positive. If so, then*

(a) *the optimal probability of an audit is zero: $q^* = 0$;*

(b) *the optimal fine for the offense is maximal: $f_O^* = w_H$;*

(c) *the optimal imprisonment sentence for the offense if w_L is paid instead of f_O exceeds s_O^* ; that is, $s_O^* > s_O^*$; and*

(d) *the expected liability of low-wealth individuals exceeds the expected liability of high-wealth individuals: $b_L > b_H$; both groups may be underdeterred or overdeterred.*

Proof: I will first prove part (b) by contradiction. Suppose $f_O < w_H$ and $s_O > 0$. Then it would be possible to raise f_O and lower s_O while keeping $f_O + s_O$ constant. Regardless of whether social welfare is given by (6) or (7), social welfare rises because the behavior of individuals is not affected but the social cost of imprisonment declines. (The case in which social welfare is described by (8) is ruled out by Proposition 1.) Thus, if $s_O^* > 0$, it must be that $f_O^* = w_H$.

I next show that if $s_O^* > 0$, s_O^* exceeds s_O^* and is such that $w_L + s_O^* > w_H + s_O^*$. If $s_O^* > 0$, then the first-order condition with respect to s_O derived from (7) is, with $f_O = w_H$ and after dividing through by $-(1 - \theta)p(e)$,

$$(1 + c)[1 - R(b_H)] + [b_H - h - p(e)(1 + c)s_O]r(b_H) = 0, \quad (11)$$

where $b_H = p(e)(w_H + s_O)$. The first term is the cost of raising s_O due to the increased social costs of imprisonment with respect to the inframarginal high-wealth individuals. The second term is the benefit (since it must be that $b_H < h + p(e)(1 + c)s_O$) of raising s_O due to the greater deterrence of high-wealth individuals from committing the offense.

The derivative of social welfare with respect to s_{ϕ} is

$$-\theta p(e)\{(1+c)[1-R(b_L)] + (b_L - h - p(e)[qk + (1+c)s_{\phi}])r(b_L)\}, \quad (12)$$

where $b_L = p(e)(w_L + s_{\phi})$. Consider this derivative at the value of s_{ϕ} such that $w_L + s_{\phi} = w_H + s_{\phi}^*$, that is, at $s_{\phi} = s_{\phi}^* + (w_H - w_L)$. At this value of s_{ϕ} , (12) can be written as

$$\begin{aligned} &-\theta p(e)\{(1+c)[1-R(b_L)] + (b_L - h - p(e)(1+c)s_{\phi}^*)r(b_L)\} \\ &+ \theta p(e)^2[qk + (1+c)(w_H - w_L)]r(b_L), \end{aligned} \quad (13)$$

where $b_L = p(e)(w_L + s_{\phi}) = p(e)(w_H + s_{\phi}^*)$. Since b_L at this s_{ϕ} equals b_H at s_{ϕ}^* , (11) implies that the first term in (13) is zero. The second term in (13) clearly is positive, so (13) is positive at $s_{\phi} = s_{\phi}^* + (w_H - w_L)$, implying that $s_{\phi}^* > s_{\phi}^* + (w_H - w_L)$. Thus, $s_{\phi}^* > s_{\phi}^*$ and $w_L + s_{\phi}^* > w_H + s_{\phi}^*$, as claimed. This establishes part (c) and the result in part (d) that $b_L > b_H$.

I next prove that $q^* = 0$ if $s_{\phi}^* > 0$. It is clear from (11) that s_{ϕ}^* does not depend on q , and from (12) that s_{ϕ}^* does depend on q . Suppose $q > 0$ and let s_{ϕ}^* be the resulting optimal value of s_{ϕ} . Social welfare then is given by (7) with $b_L = p(e)(w_L + s_{\phi}^*)$ and $b_H = p(e)(w_H + s_{\phi}^*)$. Reducing q to zero while leaving s_{ϕ}^* at its original value clearly increases (7). The only remaining issue is whether reducing q will cause high-wealth individuals to misrepresent their wealth level. The result in the preceding paragraph that $w_L + s_{\phi}^* > w_H + s_{\phi}^*$ holds for the $q > 0$ in question (it holds for any q). Thus, high-wealth individuals will prefer to pay $f_{\phi}^* = w_H$ rather than w_L even if $q = 0$; see (5). This establishes part (a).

I next address the issue of whether individuals will be underdeterred or overdeterred. First consider low-wealth individuals and the derivative of social welfare with respect to s_{ϕ} (12). Evaluated at s_{ϕ} such that $b_L = p(e)(w_L + s_{\phi}) = h$, and using the result that $q^* = 0$, it is clear that the sign of (12) is the same as the sign of $p(e)s_{\phi}r(h) - [1 - R(h)]$. The first term could be very low, for example if $p(e)$ is low, in which case the sign of (12) would be negative; this implies that s_{ϕ} 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 (12) would be positive and b_L would exceed h , so there would be overdeterrence. Similar observations apply to high-wealth individuals and will not be repeated. This establishes the second claim in part (d).

Finally, I will demonstrate that there exists a sufficient condition which, if satisfied, implies that the optimal imprisonment sentence for the offense, s_{ϕ}^* , is positive. Given $f_{\phi}^* = w_H$, it

can be shown from (7) that if the public cost of imprisonment c is zero, the derivative of social welfare with respect to s_O is positive at $s_O = 0$ if

$$h > w_H + \sup_{p \in [0, 1]} \{[1 - R(pw_H)]/r(pw_H)\}. \quad (14)$$

Thus, if (14) is satisfied and c is sufficiently low, $s_O^* > 0$.¹⁵

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 exceeds the wealth level of high-wealth individuals, using fines alone will result in underdeterrence; the additional deterrence from employing imprisonment will be socially desirable if the social cost of imprisonment is not too great. Because the private cost of imprisonment per unit of deterrence is unavoidable,¹⁵ this means that the public cost of imprisonment c must be sufficiently low.

The key insight from Proposition 4 is that, if it is optimal to use imprisonment to control the offense, the optimal imprisonment sentence imposed on individuals who claim to be low-wealth individuals is high enough to make high-wealth individuals prefer to pay the fine for the offense rather than attempt to misrepresent their wealth level. As a result, it is not necessary, or desirable, to audit individuals if they claim to have a low level of wealth.

The reason it is optimal to impose such a high imprisonment sentence on individuals who claim to have a low level of wealth can be explained as follows (in a way that parallels the proof). Consider an imprisonment sentence on low-wealth individuals that, combined with a fine equal to their wealth level, creates the same degree of deterrence of the offense as for high-wealth individuals, given their fine and imprisonment sentence. Obviously, the imprisonment sentence on the low-wealth individuals must be higher. Now raise the imprisonment sentence for each group. The effect on social welfare from the additional deterrence of the offense 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 is higher for low-wealth individuals, 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

¹⁵ In other words, to create one dollar's worth of disutility to an offender by imprisoning him necessarily causes him, and society, to incur one dollar in private cost.

sentence on low-wealth individuals than that which would equalize the deterrence of both groups. This, as noted, makes auditing neither necessary nor desirable.

That the fine for the offense should be maximal before imprisonment is used follows from well-known logic. Specifically, fines are socially costless sanctions whereas imprisonment is socially costly, so fines should be used to their fullest extent before one resorts to imprisonment. The only possible qualification to this point that arises in the present context concerns the need to use higher audit probabilities to induce high-wealth individuals to pay higher fines when fines are used alone to control the offense (see part (c) of Proposition 3). As previously observed, this makes fines socially costly due to auditing costs. This qualification is inapplicable, however, when imprisonment is used to control the offense for the reason just discussed — the optimal imprisonment sentence imposed on individuals who claim to be low-wealth individuals is high enough to make high-wealth individuals prefer to pay the fine for the offense equal to their wealth level. Thus, when it is optimal to use imprisonment to control the offense, auditing is not needed to deter high-wealth individuals from misrepresenting their wealth level, in which case the usual logic about using fines to the fullest extent possible before resorting to imprisonment continues to apply.

The explanation why both groups may be underdeterred or overdeterred also is based on a familiar argument. Consider low-wealth individuals first, and suppose that the fine and the imprisonment sentence applicable to them are such that their expected liability 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 social 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 end up serving imprisonment sentences. Either effect could dominate, making it desirable to raise the imprisonment sentence — resulting in overdeterrence — or lower the imprisonment sentence — resulting in underdeterrence. Analogous observations apply to high-wealth individuals.

The next, and final, proposition addresses an intermediate case, in which fines alone are used to sanction offenders who pay the fine for the offense, but an imprisonment sentence (together with a lower monetary payment) is imposed on offenders who claim to be low-wealth individuals.¹⁶ Because the optimal enforcement system in this case could involve either a positive audit probability, as in Proposition 3, or no auditing, as in Proposition 4, I will use the following proposition merely to establish that this case can exist, and then discuss informally the possible outcomes.

Proposition 5. *The public cost of imprisonment c may be such that it is optimal to use imprisonment to sanction offenders only if they claim to be low-wealth individuals; in other words, it is possible that $s^* = 0$ and $s^* > 0$.*

Proof: I will provide an analytical example that generates the result that $s^* = 0$ 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 and that auditing of wealth levels is costless (the latter two assumptions are not essential to the example). Then, clearly, $p^* = 1$; for 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 fines alone. The optimal imprisonment sentence for the offense therefore is $s^* = 0$. Low-wealth individuals are still underdeterred, however, since $w_L < h$. To see whether it is desirable to impose an imprisonment sentence on them — that is, whether $s^* > 0$ — differentiate social welfare (7) with respect to s_0 assuming $p = 1$. A sufficient condition for the derivative to be positive at $s_0 = 0$, regardless of q , is

$$w_L + \{(1 + c)[1 - R(w_L)]/r(w_L)\} < h. \quad (15)$$

It is evident that (15) can be satisfied. For instance, if $R(w_L)$ is sufficiently close to one, (15) will hold because $w_L < h$, in which case $s^* > 0$. Clearly, other examples in which $s^* = 0$ and $s^* > 0$ could be constructed similar to this one even if it is costly to raise the probability of detection and to audit wealth levels.▪

That this case can occur should not be surprising. Because high-wealth individuals can be deterred by fines from committing the offense to a greater extent than can low-wealth individuals, the value of an imprisonment sentence in terms of generating better deterrence is

¹⁶ Part (c) of Proposition 4 implies that the converse case — in which an imprisonment sentence is imposed only if an individual pays the fine for the offense — cannot be optimal.

lower for them. Thus, given the social cost of imprisonment, it might not be worthwhile to impose an imprisonment sentence on high-wealth offenders who pay the fine for the offense, yet it may be desirable to impose an imprisonment sentence on low-wealth offenders who cannot pay the fine for the offense.

As noted above, the optimal enforcement system in the present case might involve a positive audit probability, as in the case characterized by Proposition 3. Specifically, the optimal imprisonment sentence s^* imposed on individuals who do not pay the fine for the offense might not be high enough, together with a payment equal to the wealth level of low-wealth individuals, to cause high-wealth individuals to pay the fine for the offense. In other words, it might be that $w_L + s^* < f^*$, in which case auditing is required to induce high-wealth individuals to pay f^* .

An alternative outcome also is possible, in which it is optimal to not audit at all, as in the case described by Proposition 4. This would occur if the optimal imprisonment sentence s^* imposed on individuals who do not pay the fine for the offense is such that $w_L + s^* > f^*$. High-wealth individuals then will pay the fine for the offense even if there is no risk of being audited.

One theme that emerges from the analysis in this section is that, if the enforcement authority cannot observe offenders' wealth levels without costly audits, auditing and imprisonment are alternative means of deterring offenders from misrepresenting their wealth levels. When imprisonment sentences are not used at all to control the offense — the case addressed by Proposition 3 — auditing is socially valuable and should be employed if it is not too costly. When imprisonment sentences are used to a limited extent, just for offenders who do not pay the fine for the offense — the case addressed by Proposition 5 — auditing might or might not be socially valuable. When imprisonment sentences are used most extensively, both for offenders who do not pay the fine for the offense as well as for those who do — the case addressed by Proposition 4 — auditing is of no value because the optimal structure of imprisonment sentences makes auditing unnecessary. Thus, in a rough sense, the greater the use of imprisonment sentences as a sanction for the offense, the lesser the need for auditing.

IV. Prior Literature

As observed in the introduction, the conventional analysis of sanctions presumes that individuals' wealth levels can be observed at no cost. The simplest version of this analysis

assumes that all individuals have the same level of wealth. Then, if individuals are risk neutral and the sanction is a fine, it is well known that the optimal fine is maximal, equal to the individuals' wealth level, and that the optimal probability of detection is such that there is some underdeterrence. If an imprisonment sentence also is considered, it should not be used unless the fine is maximal; if an imprisonment sentence is employed, its optimal level might lead to overdeterrence or underdeterrence.

Of greater relevance to the present article are versions of the preceding analysis in which individuals have different levels of wealth, and their wealth levels can be observed at no cost. Polinsky and Shavell (1984, pp. 95-98) briefly consider this case when there are two levels of wealth. They show that if fines are used alone, the optimal fine for the offense exceeds the wealth level of the low-wealth individuals, but may be less than the wealth level of the high-wealth individuals. Low-wealth individuals are underdeterred, but, if the optimal fine is less than the wealth level of the high-wealth individuals, high-wealth individuals are ideally deterred.¹⁷ This latter result differs from my result here, where both low-wealth and high-wealth individuals are underdeterred (see part (d) of Proposition 3). The reason is that, in my analysis, because wealth is not costlessly observable, the threat of an audit is required to induce high-wealth individuals to pay a higher fine; and because audits are costly, it is optimal to set the audit probability such that some underdeterrence is tolerated.

When Polinsky and Shavell consider imprisonment sentences, they assume that the private cost of imprisonment is greater for high-wealth individuals because such individuals have a higher opportunity cost of time. As a result, in their model the optimal imprisonment sentence could be higher for high-wealth individuals than for low-wealth individuals (because the social cost of imprisonment per unit of deterrence is lower for high-wealth individuals). Here, as seen, if imprisonment sentences are used, the optimal sentence for the offense is lower than the optimal sentence imposed on individuals who do not pay the fine for the offense (that is, $s^* < s^*$). Thus, in my analysis, high-wealth individuals always bear a lower imprisonment sentence than low-wealth individuals. This difference stems from my assumption that the private cost of

¹⁷ Polinsky and Shavell (1991) show that this conclusion generalizes to a continuous distribution of wealth levels, and Garoupa (1998, pp. 484-87) demonstrates that it continues to hold if the enforcement authority costlessly observes an underestimate of individuals' wealth levels.

imprisonment is the same for both types of individuals (this assumption is discussed further in the next section).¹⁸

As noted in the introduction, Levitt (1997) analyzes optimal sanctions when the enforcement authority is assumed to be unable to observe individuals' wealth levels at any cost. Although he also considers two types of individuals, described as "rich" and "poor," his model differs in significant respects from mine. In Levitt's framework, the 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. Another significant difference is that Levitt just considers (in his formal analysis) the choice between an imprisonment sentence and a fine, not the optimal combination of the two, as I do here.

Levitt first derives the optimal use of imprisonment sentences when imprisonment is the only possible sanction. In his model, the optimal imprisonment sentence is maximal and the probability of detection, if positive, is either high enough to just deter both groups from committing the offense or high enough to just deter the group that obtains the lower benefit. He then considers whether social welfare can be raised by offering offenders the choice between a fine and an imprisonment sentence. 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 may not result in higher social welfare. Consequently, 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

¹⁸ In an earlier article, Friedman (1981) examined optimal sanctions when sanctions are costly and the cost of imposing the sanction varies with wealth (he referred to the sanction as a fine, though his results also apply to imprisonment). In his analysis, the optimal sanction could be higher for high-wealth individuals or low-wealth individuals.

enforcement authority cannot observe offenders' wealth levels. Their analysis is, in a sense, the converse of Levitt's — they first derive the optimal use of fines when fines are the only possible sanction, and then ask whether social welfare can be improved by introducing imprisonment sentences. Their main point is that, because of 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. Thus, like Levitt, they emphasize the desirability of imprisonment sentences relative to fines.

Although I did not analyze explicitly the case in which the enforcement authority cannot observe offenders' wealth levels, this case corresponds in my framework to the situation when the cost of auditing k is so high as to make auditing undesirable. It is clear from my analysis that the optimal enforcement system then would be characterized in one of three ways. One possibility is that fines alone are used to control the offense and the optimal fine f^* is equal to the wealth level of low-wealth individuals w_L . Another possibility is that f^* exceeds w_L and high-wealth individuals pay f^* because they would serve a sufficiently long imprisonment sentence s^* otherwise; in other words, $w_L + s^* > f^*$. The last possibility is similar, but with the offense being sanctioned both by an imprisonment sentence s^* and a maximal fine; this case is characterized by $w_L + s^* > w_H + s^*$. Thus, if the enforcement authority can employ both fines and imprisonment sentences together, rather than just one or the other (as considered by Levitt), it often will be optimal to use them both. Moreover, imprisonment sentences should not be used unless the party subject to the imprisonment sentence has paid a fine equal to his wealth level (in contrast to Chu and Jiang's result).¹⁹

In passing, it is worth noting that there is a large theoretical literature on the auditing of wealth levels in the context of controlling tax evasion, which has been ably summarized by Mookherjee (1997, pp. 207-31) and Andreoni, Erard, and Feinstein (1998, pp. 823-35). Although this body of research might appear to be quite relevant to my analysis, it is only tangentially so, mainly because its focus is on controlling offenses that affect the distribution of

¹⁹ Also of some relevance to the subject of this 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 their defense reduces the overdeterrence that otherwise would occur. For further discussion of this point, see Garoupa and Gravelle (2000).

income rather than that cause harm. It also ignores, to my knowledge, consideration of imprisonment as a sanction. The main point of similarity between my analysis and the tax evasion literature is the common finding that the optimal fine for misrepresenting one's wealth level is maximal so that the probability of an audit can be as low as possible.

V. Concluding Remarks

One of the surprising results in my analysis is that the difficulty of observing offenders' wealth levels might not hinder the design of the optimal enforcement system. As observed in Section III, auditing is not necessary if high-wealth individuals are induced to pay the fine for the offense because they would bear more burdensome sanctions if they claimed to be low-wealth individuals. Thus, if the optimal enforcement system has this character, the inability to audit wealth levels will not result in a reduction of social welfare. However, as also seen in Section III, there are many other circumstances — including when the cost of imprisonment makes its use undesirable — in which auditing is beneficial and should be employed if it is not too expensive. The general contribution of this article has been to demonstrate these two points and to describe the circumstances in which they are applicable.

I made several simplifying assumptions in order to keep the analysis tractable. The most notable ones are that there are only two levels of wealth; that individuals bear the same disutility from time in prison regardless of their wealth level; and that individuals are risk-neutral with respect to both fines and imprisonment. I do not believe that any of these assumptions are crucial to the results. 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. As here, it seems clear that the optimal menu could have the property that the sum of the sanctions declines with the fine paid, which would induce individuals to pay the highest fine of which they are capable. If this condition holds, auditing of wealth levels would be unnecessary. But if this condition does not hold, as it could not if fines are used alone, auditing would be beneficial because some auditing would be required to induce higher-wealth individuals to pay higher fines.

If individuals differed in terms of their disutility from time in prison, with, say, high-wealth individuals suffering greater disutility than low-wealth individuals (as the other authors

discussed in Section IV assumed), my results also do not appear to be qualitatively affected. While the imprisonment sentences that would be optimal obviously would be different, it still might — or might not — be optimal to choose the menu of sanctions in such a way that the disutility of the sanctions declines with the fine paid.²⁰ Thus, both of the main points I made would continue to apply, though the circumstances in which they are relevant would be different.

If individuals are not risk neutral with respect to fines or imprisonment, the optimal sanctions would change, but, again, the sanctions still could be such that high-wealth individuals will prefer to pay the fine for the offense in order to avoid the alternative sanctions, in which case auditing is not socially valuable. Or it could be the case that auditing is needed to induce higher-wealth individuals to pay higher fines.

Thus, while these generalizations and modifications of the model certainly would affect the circumstances in which auditing is or is not beneficial, they do not appear to change the basic points made.

²⁰ To see that this condition could still hold, consider the following slight modification of my model. Suppose that the disutility borne by a low-wealth individual from an imprisonment sentence of length s remains equal to s , but that the disutility borne by a high-wealth individual is λs , where $\lambda > 1$. It is clear from the proof that $b_L > b_H$ in part (d) of Proposition 4 that $w_L + \lambda s \delta^* > w_H + \lambda s \delta^*$ will hold if λ is sufficiently close to one. In other words, high-wealth individuals will still prefer to pay the fine $f \delta^* = w_H$ rather than misrepresent their wealth level, even if there is no risk of an audit.

References

- Andreoni, James, Brian Erard, and Jonathan Feinstein, "Tax Compliance," *Journal of Economic Literature*, Vol. 36, No. 2 (June 1998), pp. 818-60.
- Becker, Gary S., "Crime and Punishment: An Economic Approach," *Journal of Political Economy*, Vol. 76 (1968), pp. 169-217.
- 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, "Does Efficient Deterrence Require That the Wealthy Should Be Able to Buy Justice?," Discussion Paper No. 2000/7, Department of Economics and Related Studies, University of York, United Kingdom, February 2000.
- 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.
- Mookherjee, Dilip, "The Economics of Enforcement," in Amitava Bose, Mihir Rakshit and Anup Sinha (Editors), *Issues in Economic Theory and Policy, Essays in Honor of Tapas Majumdar* (New Delhi: Oxford University Press, 1997), pp. 202-49.
- 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.