

INCENTIVE COMPATIBILITY CONSTRAINTS
AS AN EXPLANATION FOR THE
USE OF PRISON SENTENCES
INSTEAD OF FINES

Steven D. Levitt*

Discussion Paper No. 187

6/96

Harvard Law School
Cambridge, MA 02138

The Center for Law, Economics, and Business is supported by
a grant from the John M. Olin Foundation.

*Junior Fellow, Harvard Society of Fellows.

Incentive Compatibility Constraints as an Explanation for the Use of Prison Sentences Instead of Fines

Steven D. Levitt*

* Harvard Society of Fellows and the National Bureau of Economic Research, 78 Mount Auburn Street, Cambridge, MA 02138. E-mail address: slevitt@fas.harvard.edu. I would like to thank Peter Diamond, Christine Jolls, Louis Kaplow, Mitch Polinsky, Matthew Rabin, two anonymous referees and participants in the Harvard Law School Research Seminar in Law and Economics for helpful discussions on the topic and comments on earlier drafts of the paper. All remaining errors are my own. Financial support of the National Science Foundation is gratefully acknowledged.

Abstract

In standard models of optimal deterrence, which assume perfect information on the part of the social planner, fines dominate jail sentences as an instrument for punishing crime. In the real world, however, punishment using prisons is quite common. This paper provides an explanation for that empirical observation. If criminals either have private information about their own wealth, or have a substantial portion of their wealth in the form of human capital, the social planner cannot simply impose a fine on a criminal, who can always claim to have insufficient wealth to pay the fine. Rather, because of the need for incentive compatibility, it is as if the social planner offers the criminal a choice: either pay the fine or go to jail. That additional constraint on the social planner dramatically reduces the effectiveness of fines vis-a-vis a perfect information world. Mandatory sentencing guidelines exacerbate the problems associated with incentive compatibility by restricting a judge's ability to tailor penalties to individual criminals.

A sizable literature has arisen in recent years concerning the optimal punishment of criminal acts.¹ Relatively little attention, however, has been directed to the question of whether criminals should be punished using fines or jail sentences. Becker (1968) and Posner (1977) argue informally for the use of fines where possible to avoid the social costs of maintaining prisons, the private disutility of incarceration, and the social waste of idle human capital in jails.² The analysis of Polinsky and Shavell (1984) confirms that intuition formally under the assumption that the social planner has perfect knowledge of each individual's wealth, has a costless means of enforcing fines once the criminal is caught, and can costlessly tailor punishments to each individual. In that scenario, the social planner's optimal approach is to fully exploit fines, only using jail terms if monetary sanctions are insufficient to deter criminal activities. In practice, however, reliance on incarceration as punishment is extensive. At year-end 1994, the combined U.S. jail and prison population exceeded one million. In contrast, fines are rarely used exclusively except for minor crimes such as traffic violations and some federal crimes.³ Even when fines are used, the dollar amounts levied are typically small.

Three common arguments are put forth as explanations for the limited use of fines as

¹ Included in this recent literature are Andreoni (1991), Kaplow (1990,1992), Kaplow and Shavell (1994), Lott (1987), Polinsky and Rubinfeld (1991), and Polinsky and Shavell (1984, 1991, 1992). The classic works in this area are Becker (1968), and Stigler (1970).

² The distinction between prisons and jails is unimportant to the arguments that follow. Consequently, those two terms are used interchangeably in the paper.

³ The use of fines is more extensive in many other countries (see Gillespie 1981). There are also some indications that interest in the use of fines is on the increase in the United States (NIJ 1992). Waldfogel (1995) examines the use of fines in federal fraud cases. Weisburd et al. (1991, 150-157) consider the use of fines for a range of white collar crimes.

punishment in spite of the theoretical predictions. A first argument is that fines are "unjust" because the rich are able to pay fines while the poor serve jail sentences. If, however, the social planner knew the precise monetary equivalent of a jail sentence for a particular individual, and imposed a fine of that magnitude or greater, redistributing the revenue to either the victim of the crime or the poor, it is hard to see how such a fine could be deemed unjust.

A second argument against the use of fines is that the typical criminal has extremely low wealth. As a result, feasible fines might be so small as to be inconsequential as deterrents. Undermining this argument, however, is the fact that even if a criminal does not have substantial *physical* assets, he or she typically does have human capital that is capable of generating a flow of future income with which to pay a fine. There is no legal restriction on the government's ability to collect fines in installments to take advantage of future earnings. Nonetheless, while monthly payments are standard for child support awards, they are rare for criminal punishments in the United States (Hillsman et al. 1982).

A final argument for the use of prisons rather than fines is the value of incapacitation of repeat offenders. According to prisoner surveys, the median prisoner commits 12-15 serious non-drug crimes per year when not behind bars (Visser 1986, Piehl and DiIulio 1995). While the incapacitation argument is compelling, the puzzle of why prison terms are not combined with wealth-exhausting fines remains.

This paper offers an alternative explanation for the limited use of fines. The fundamental premise of the paper is that the government cannot enforce fines except under

the threat of further punishment, in particular prison sentences.⁴ When imposing fines, the social planner is therefore implicitly offering the criminal a choice: either pay the fine, or if you refuse/are unable, you will be required to spend time in jail.⁵ Put into standard economic jargon, only fines that are incentive compatible will be paid. Adopting that assumption, while maintaining the other basic assumptions of the previous literature, it is demonstrated that the usefulness of fines is greatly reduced relative to the earlier papers that assume perfect information and enforceability of fines.

There are at least three reasons why governments may have difficulty collecting fines even if agents have sufficient wealth, perhaps in the form of human capital. One reason is the existence of private information on the part of individuals concerning their own wealth (either physical assets or human capital). With private information, agents will have an incentive to report wealth levels below the true value, simply refusing to pay a given fine due to (supposed) inability. Secondly, even if the government can accurately assess an individual's wealth, there may be no direct means to compel payment of a fine. That is certainly the case with respect to future returns on human capital, the most valuable asset for most criminals. While a criminal may be capable of generating a steady flow of income that the government could take as a fine, there is no way for the government to force the criminal

⁴ As a concrete example, traffic violations are typically punished using fines. The government has only limited success in collecting such fines, and therefore resorts to more serious forms of punishment including confiscation of vehicles, revocation of driving privileges, and traffic school.

⁵ This ignores any difficulties that the social planner might have in enforcing jail terms. Enforcing jail terms is likely to be much less of an issue because the criminal presumably is present at his or her sentencing.

to earn that income because the individual's effort is unobserved (as well as inalienable) and is therefore outside the government's control. While physical assets are certainly easier to confiscate, for many criminals with sizable holdings, e.g. drug traffickers and participants in organized crime, it is likely to be extremely difficult to locate those assets. Finally, while the social planner may be able to confiscate assets, the costs of doing so may be prohibitive relative to the revenue generated, especially for small fines.

To understand why incentive compatibility limits the effectiveness of fines, it is useful to consider the two purposes served by fines in a world with perfect information and costless enforcement of fines. The first purpose of fines is to provide a greater level of deterrence for a given likelihood of detection. If the social planner can impose both a fine equal to wealth *and* the maximum jail sentence, deterrence is greater than utilizing the maximum jail sentence alone. If incentive compatibility is required, however, the combined punishment using both fines and jail cannot be any greater than the punishment using the maximum jail sentence alone. If the social planner attempts to impose a stiffer penalty, the criminal can simply refuse to pay the fine, instead accepting the maximum jail sentence. Thus, in a world with incentive compatibility constraints, introducing fines (weakly) lowers rather than raises the deterrent effect of a given level of detection.

The second use of fines in a perfect information world is to minimize the social costs of punishing offenders. Given that a crime occurs, it is socially more efficient to use fines than prisons for punishment, due both to the costs of maintaining prisons and the wasted human capital associated with incarceration. While this justification of fines remains relevant, its practical usefulness is reduced if the social planner cannot tailor the punishment

to the individual's type (either because the individual's type is private information and therefore not known to the social planner, or because of mandatory sentencing guidelines that rule out such tailoring). Since the typical criminal is likely to have a low disutility of prison (in dollar terms) relative to the population as a whole, the fine must be low relative to the jail sentence in order to be incentive compatible. The availability of a low fine as a potential punishment, however, will make crime more attractive to the rest of the population. That reduction in deterrence can easily outweigh the social benefits of avoiding incarceration costs.

Despite the above limitations on implementing fines, the model presented in this paper shows that there are some situations where fines may be effective despite the presence of incentive compatibility constraints. Fines are especially useful for crimes where the private benefits to crime (subtracting off the costs of committing the crime) are greater for the rich than the poor, such as tax evasion. For those crimes, large fines will be preferred to jail by the typical tax-evader, but will not be attractive to most non-offenders.⁶ Fines are also effective when there is a big gap between the private benefits to crime for criminals and non-criminals. When that is the case, any reduced deterrence associated with the availability of fines is unlikely to affect the non-criminal's decision. A crime such as child molestation, which provide large private benefits to a few members of society, but negative private benefits to most people, is one example. In this case, however, the optimal fine is likely to greatly exceed an individual's wealth. There may also be strong incapacitation arguments for

⁶ In addition, the social costs of idle human capital associated with jail terms is presumably greater when the foregone earnings of the criminal is high. The decline in future earnings associated with imprisonment is greater for high-income offenders (Lott 1992, Waldfogel 1994).

the use of prisons rather than fines in punishing child molesters.

Fines are more useful when the jail/fine tradeoff is similar for criminals and non-criminals since the incentive-compatibility constraint binds less tightly in such cases. The value of fines also tends to be greatest when the social externality to crime is small relative to the costs of detection and incarceration. The reliance on fines for traffic offenses and other minor crimes is consistent with this last point.

The structure of the paper is as follows. Section I presents a simple model in which individuals have private information about their own types, and the social planner has access only to jail sentences as an instrument for punishing criminals. (In the presence of private information, there cannot be only fines as punishment because agents would report low wealth and refuse to pay.) In Section II, fines are introduced as an additional instrument for the social planner. Because of private information, however, any punishment involving fines must be incentive compatible. Section III explores conditions under which fines are more or less effective and derives general propositions concerning the usefulness of fines in the model. The model employed in this paper is intentionally sparse. The benefit of such an approach is that the adverse effects of incentive compatibility constraints are highlighted. The cost, however, is that some measure of realism and generality is sacrificed. Section IV therefore considers the effect of generalizing the model in numerous ways. Section V offers a brief set of conclusions.

Section I: Optimal Punishment in a Jail-Only World

Agent i , assumed to be risk neutral, maximizes the following expected utility function

$$E(U_i) = \begin{cases} 0 & \text{if no crime is committed} \\ \beta_i - \alpha J \theta_i & \text{if a crime is committed} \end{cases} \quad (1)$$

where β_i is the private benefit to agent i of committing the crime, α is the likelihood of detection given that the crime is committed, J is the length of the jail term if detected, and θ_i is a measure of the disutility of jail to agent i . For algebraic simplicity, it is assumed that the individual derives all benefits from the crime before the detection occurs; the substantive results of the analysis are unchanged when this assumption is altered.

The social planner commits to a level of α and J prior to the agent's decision about whether or not to commit the crime, so as to maximize social welfare W given by

$$\begin{aligned} \underset{\alpha, J}{\text{Max}} \quad W &= \sum_{i=1}^N E[U_i^*(\beta_i, \theta_i, \alpha, J)] - c(\alpha, J) * [L + \alpha J \psi] - e(\alpha) \\ \text{s.t.} \quad E[U_i^*] &\in \text{argmax } E[U_i] \quad \forall_i \end{aligned} \quad (2)$$

where U_i^* reflects agent i 's utility maximizing choice of behavior, c is the number of agents committing the crime (which depends on the social planner's choice of policies), L is the social externality associated with the crime, ψ is the external cost borne by society of incarcerating criminals that are detected (assumed to increase linearly with the number of prisoner-years served), and e is the cost of enforcement associated with a given probability α

of detection.⁷ The social planner maximizes W subject to the incentive-compatibility constraint for each agent with respect to committing the criminal act. Private benefits to crime are included in the social welfare function; excluding such benefits does not alter the paper's conclusions.

The following assumptions are made:

- | | |
|--|---|
| (A1) $0 < \beta_i < L \quad \forall_i$ | (A5) $i \in \{P, R\}$ |
| (A2) $J \in [0, J_{\max}]$ | (A6) $\theta_R > \theta_P$ |
| (A3) $\psi > 0$ | (A7) An agent's type i is private information |
| (A4) $e'(\alpha) > 0, e''(\alpha) < 0$ | |

A1 says that the social costs of crime outweigh the private benefits to crime for all individuals; if deterrence were costless, the social planner would want no crimes committed. The basic conclusions of the model are unchanged if this assumption is relaxed, allowing some crimes to be beneficial from a social perspective. A2 imposes a maximum jail sentence. This assumption is standard in analyses such as this one to avoid the Becker (1968) result of infinite punishments, and can be motivated by constitutional restrictions, the need for marginal deterrence, societal desires for horizontal equity, or simply the fact that people have finite lives. A3 says that there are costs borne by society of maintaining prisons, and these costs outweigh any possible benefits of incapacitating potential repeat offenders. In other words, if prison terms did not have deterrence value, the social planner

⁷ Using the terminology of Polinsky and Shavell (1992), enforcement costs are modeled in this paper as entirely "fixed." The results are not substantively altered by allowing some portion of the enforcement costs be "variable."

would not choose to incarcerate criminals.⁸ A4 implies that enforcement is costly, and that there are decreasing returns to expenditures on enforcement. A5 limits the analysis to the case of two types, hereafter described as "rich" and "poor." For simplicity, it is assumed there is an equal number of each of the two types. The rationale for calling the two types rich and poor is given by A6: the disutility of going to jail is assumed to be greater for the rich than the poor. This is presumably due to the greater foregone consumption of the rich while incarcerated and a larger decline in post-conviction earnings for the rich (Lott 1992), but is perhaps also a function of relevant psychological costs to jail that are greater for the rich than the poor. A7 implies that the social planner cannot make punishments directly conditional on the individual's type (i.e., $J_i = J \forall_i$) so must consider each individual's incentive-compatibility constraint when choosing optimal policies.⁹ As mentioned previously, there are other reasons beside private information that punishments would not be type-dependent, including mandatory sentencing guidelines (Champion 1989) and prohibitive costs of tailoring punishments to individuals. Although I will refer to A7 as the private

⁸ A3 seems to be a reasonable assumption empirically. Studies of the relationship between prison populations and aggregate crime rates find that the marginal social costs of incarceration are approximately equal to the marginal social benefit due to crime reductions (Marvell and Moody 1994, Donohue and Siegelman 1994, Levitt 1995a). To the extent that a substantial portion of the crime-reducing effect of higher prison populations is due to increased deterrence, incapacitation effects alone would not justify current levels of incarceration. Ehrlich (1981) and Levitt (1995b) provide evidence that deterrence is at least as strong a factor in reducing crime as incapacitation.

⁹ Even with private information it may be possible in a dynamic context to encourage payment of fines if the social planner is able to vary the amount of enforcement resources directed at a particular criminal. The social planner can threaten to increase future enforcement more if criminal elects to go to jail rather than pay the fine. This line of argument is similar to that of Polinsky and Rubinfeld (1991).

information assumption hereafter, any of these other explanations are equally compatible with the formal model. While an individual's type is private information, the social planner is assumed to know the values of β_P , β_R , θ_P , and θ_R .

The solution to the individual's maximization problem is straightforward; the crime is committed only if the expected punishment is less than the private benefit to crime:

$$E[U_i^*] = \begin{cases} 0 & \text{if } \alpha J \geq \beta_i / \theta_i \\ \beta_i - \alpha J \theta_i & \text{if } \alpha J < \beta_i / \theta_i \end{cases} \quad (3)$$

Given equation (3), the social planner chooses α and J to maximize social welfare.

Lemma 1 follows directly from the set-up of the problem:

Lemma 1:

- (i) The social planner always sets $J = J_{\max}$.
- (ii) If both agents commit the crime at the social optimum, the level of enforcement α must be equal to zero.
- (iii) The socially optimal choice of $\alpha \in \{0, \beta_P / (J_{\max} \theta_P), \beta_R / (J_{\max} \theta_R)\}$.

proof: (i) Note first that the level of deterrence in equation (3) depends on α and J only through their product. Holding αJ fixed in equation (2), $U_i^*(\cdot)$, $c(\cdot)$, and the social costs of incarceration are unaffected as the ratio of α to J changes. Enforcement costs, however, are strictly increasing in α , but are not a function of J . Therefore, for a given level of deterrence, social welfare is maximized by setting J as large as possible.

(ii) Assume that allowing both types to commit the crime is the socially optimal solution to the social planner's problem in equation (3). Since enforcement is costly, and the

only benefit of enforcement is deterrence, if no deterrence occurs at the social optimum, the cost minimizing enforcement level ($\alpha=0$) must be chosen.

(iii) Setting $J=J_{\max}$ by part (i) of this lemma, the social planner then chooses α to maximize social welfare. There are four possible cases to consider: both types commit the crime, one type or the other commits the crime, or neither type commits the crime. Because enforcement is costly, the social planner sets α as low as possible while still deterring one or both types. The cutoff values for α come directly from the agents' maximization problem. If both types commit the crime at the social optimum, then $\alpha=0$ by Lemma 1(ii).

QED

Lemma 1(i) says that the social planner punishes criminals with the highest possible jail sentence.¹⁰ Enforcement is costly, so the cheapest means of obtaining a given level of deterrence is to set $J=J_{\max}$, lowering α , the probability of detection, as far as possible while still deterring those crimes for which it is socially desirable to do so. By Lemma 1(ii), if enforcement is costly enough, the social planner will prefer to set the detection rate α equal to zero and let both types commit the crime. Part (iii) of Lemma 1 identifies the range of enforcement levels that must be considered when searching for a social optimum. Because enforcement is costly, the social planner sets the detection rate at the lowest level consistent with deterrence.

Given Lemma 1, the social planner's problem simplifies to a computation of social welfare at the three potentially optimal levels of α . Which of those choices of α is optimal will depend on the parameter values of the model; for different sets of parameters, it may be optimal to deter both types, exactly one type, or neither type.

¹⁰ This result contrasts with Kaplow (1990) who finds that optimal jail sentences are often less than the maximum allowed because of the deadweight loss associated with jail sentences served by inframarginal criminals. The Kaplow result does not arise in the simple model developed here because there are only two types; therefore, there will never be an inframarginal type serving a jail sentence in this model.

Section II: Introducing Fines as an Additional Instrument in a Jail-Only World

Now, assume that instead of punishing criminals exclusively by sending them to jail, the social planner can offer the criminal the option of paying a fine and foregoing the prison sentence. It is assumed, however, that the social planner cannot mandate that the criminal pay the fine if the criminal refuses. As discussed in the introduction, that assumption can be motivated in three ways. First, since the social planner does not know an agent's wealth, it may be impossible to determine whether the agent's wealth is sufficient to cover the fine. Second, the criminal's wealth may be held in the form of human capital and therefore will be difficult to extract. Finally, the government may simply not have the means of collecting fines when there is no further sanction and the criminal refuses to cooperate; the limited success on the part of the government in enforcing child support payments and parking violations are two examples of that difficulty.¹¹ While no revenue can be collected from a criminal that refuses to pay a fine, it is assumed that the criminal can be sent to jail. In practice, then, the social planner offers an individual who has been caught committing a crime the option of either paying fine F or serving jail term J .¹² As before, an agent's wealth is private information so that punishments cannot be made directly conditional on wealth. It is assumed that all fines that are collected are redistributed as a lump sum across

¹¹ For example, only 54 percent of parents with child support awards regularly receive payment (GAO 1992), but payments have repeatedly been shown to be highly responsive to threats of further punishment.

¹² A more realistic assumption would allow the social planner to offer a more complicated menu of contracts involving combinations of fines and jail terms. To make the underlying intuition as transparent as possible, however, the simplest case is examined here. The more general case is discussed in Section IV, but it is worth noting here that the assumption of either fines or jail is not as restrictive as it first appears.

all individuals.

Imposing the further assumption of quasi-linear preferences, the individual's utility function becomes

$$E[U_i] = \begin{cases} 0 + \alpha f F & \text{if no crime is committed} \\ \beta_i - \alpha * \text{Min}[J\theta_i, F] + \alpha f F & \text{if a crime is committed} \end{cases} \quad (4)$$

where f is the percentage of those agents who commit a crime that elect to pay the fine if detected. The "Min" term in the second line reflects the incentive compatibility requirement; the agent has the option of choosing between the two forms of punishment. The last term in both lines of the utility function is the lump sum redistribution. The assumption of quasi-linearity implies that each agent values an extra dollar of income identically, i.e. utility is transferable. The combination of a utilitarian social welfare function and quasi-linear preferences rules out the use of the criminal justice system for redistributive purposes. It seems reasonable to assume that the social planner has access to a more direct means of achieving redistribution if that is an objective.¹³

The social welfare function is now given by

¹³ Nonetheless, it is worth noting that fines might be a useful tool for achieving redistribution of income from the rich to the poor if the assumption of quasi-linearity is relaxed. If the rich have a lower marginal utility of income, it might be socially beneficial to allow the rich type to commit a crime that otherwise would be optimally deterred, paying a fine that is then distributed as a lump sum.

$$\begin{aligned}
\text{Max}_{\alpha, J, F} \quad W = & \sum_{i=1}^N E[U_i^*(\beta_i, \theta_i, \alpha, J, F)] - c(\alpha, J, F) * L + (1-f(\alpha, J, F)) * \alpha J \psi - e(\alpha) \\
\text{s.t.} \quad & E[U_i^*] \in \text{argmax } E[U_i] \quad \forall_i
\end{aligned} \tag{5}$$

The social planner's problem differs in three ways from the jail-only world. First, there are now three instruments available: α , J , and F . Second, additional incentive compatibility constraints on the part of the agents have been introduced due to the fine option. Finally, the social cost associated with incarcerating convicted criminals is borne only by the fraction of criminals $(1-f)$ who elect to serve a jail sentence if detected, rather than on all detected criminals.

Assume further,

$$(A8) \quad F \in [0, \infty]$$

i.e., there is no exogenously imposed maximum fine. Assumption A8 rules out all situations where the maximum fine is so low that introducing fines has adverse effects due to possible underdeterrence. Of course, the social planner can (and will) set the actual fine below the maximum allowable fine.¹⁴

The outcome of the agent's utility maximization problem is as follows:

¹⁴ In practice, a restriction on the maximum fine is the criminal's ability to pay. In what follows, it is assumed that the maximum fine is such that the criminal can cover the fine, either through current assets or using the flow of future income. For crimes requiring greater punishment than that, fines will be even less effective than the following analysis suggests.

$$\begin{aligned}
E[U_i^*] = & \alpha Ff & \text{if } (\beta_i \leq \alpha J\theta_i \text{ and } \beta_i \leq \alpha F) \\
& \beta_i - \alpha F + \alpha Ff & \text{if } (\theta_i > F/J \text{ and } \beta_i > \alpha F) \\
& \beta_i - \alpha J\theta_i + \alpha Ff & \text{if } (\theta_i \leq F/J \text{ and } \beta_i > \alpha J\theta_i)
\end{aligned} \tag{6}$$

The first line is the utility of those not committing a crime, the second line corresponds to those who commit the crime and choose to pay the fine when caught. The third line reflects those criminals who opt for jail if detected.

Parts (i) and (ii) of Lemma 1 extend in a straightforward manner to a fine-and-jail world. The following lemma helps to further characterize possible social optima in the fine-and-jail world:

Lemma 2:

- (i) If the poor type committing a crime prefers a given fine to a given jail term, the rich type will also prefer the fine to jail when convicted of a crime.
- (ii) If only the rich type commits the crime in a social optimum, that optimum will involve the rich choosing the fine rather than serving the jail sentence when caught.
- (iii) The socially optimal choice of $\alpha \in \{0, \beta_P/\theta_P J_{\max}, \beta_R/\theta_R J_{\max}, \beta_R/\theta_P J_{\max}\}$.

proof:

(i) *This result follows trivially upon inspection of equation (4) given the assumption that $\theta_R > \theta_P$.*

(ii) *This is most easily proved by contradiction. Assume that at the social optimum only the rich type commits the crime, and given the social planner's choice of J and F , elects to go to jail if caught. For the choices of the two types to be incentive compatible, it must be the case that*

$$\theta_R < F/J_{\max} \text{ and } \beta_R > \alpha J_{\max} \theta_R \quad (\text{IC for the rich})$$

$$\theta_p > \beta_p/\alpha J_{\max} \text{ and } \beta_p < \alpha F \quad (\text{IC for the poor})$$

Combining the two conditions of the IC constraint for the poor, $\theta_p > F/J_{\max}$. But that is a contradiction, because $\theta_R < F/J_{\max}$ from the IC constraint of the rich, and $\theta_R > \theta_p$ by assumption A6.

(iii) There are five possible cases: both types commit the crime, only the rich commit the crime and pay the fine if detected, only the poor commit the crime and either pay the fine or go to jail, or neither type commits the crime. Because enforcement is costly, the social planner chooses the lowest α consistent with inducing a given set of behaviors. If both types commit the crime, $\alpha=0$ is optimal by Lemma 1. If only the rich commit the crime, it must be the case that $\beta_R > \beta_p$, so the binding constraint on the poor is the jail sentence, i.e. $\alpha = \beta_p/\theta_p J_{\max}$. If only the poor commit the crime, the optimal α depends upon whether the poor prefer to serve the jail sentence or pay the fine. If the poor prefer jail, the lowest α that will deter the rich is the same as in the jail-only world: $\beta_R/\theta_R J_{\max}$. If the poor prefer the fine, however, manipulation of the IC constraints reveals that α must be higher to deter the rich: $\beta_R/\theta_p J_{\max}$. If neither type commits the crime, one of the above values of α must be binding.

QED

Parts (i) and (ii) of Lemma 2 are a direct consequence of the assumption that the tradeoff between fines and jail is more favorable for the rich. Part (iii) defines the set of enforcement levels that must be checked when seeking a social optimum. These enforcement levels are identical to those in the jail-only world, except that a new possibly optimal enforcement level exists for the case where the poor type commits the crime and pays the fine, but the rich type is deterred. Since the jail/fine tradeoff is more favorable to the rich than the poor, a fine that is attractive to the poor must be accompanied by an increased level of enforcement if the rich are to be deterred.

Section III: Comparing Social Welfare in a Jail-Only World to a Fine-and-Jail World

The preceding sections broadly outlined the possible outcomes in jail-only and fine-

and-jail worlds respectively. A complete accounting of the exact solutions to the two models is both tedious and unrevealing. A more useful exercise is to consider what effect the introduction of fines has on social welfare vis-a-vis a world that previously did not have fines available as instruments. Because the model developed here is quite simple, it is actually the case that the availability of fines has no impact whatsoever on social welfare in many circumstances. In more general settings discussed in Section IV, it is clear that introducing fines must make at least a marginal contribution to social welfare. Nonetheless, examining this simpler model provides insight into both the reasons why fines are less useful in the presence of private information, and the conditions under which fines continue to be effective punishments.

The analysis of this section proceeds in two steps. First, the conditions under which introducing fines does or does not increase social welfare are outlined. Second, consideration is given to how specific parameters of the model (e.g. relative private benefits to crime for the rich and poor types, magnitude of the externality associated with a particular crime) affect the usefulness of fines.

The following proposition addresses the first of these issues:

Proposition 1: Introducing fines leads to an increase in social welfare if and only if the fine-and-jail social optimum involves

- (i) precisely one type committing the crime,
- (ii) that type chooses to pay the fine when detected.

proof: If both types commit the crime in the fine-and-jail world, the optimal α is zero by Lemma 1. This was a feasible strategy in the jail-only world, however, so social welfare cannot have improved.

If both types are deterred in the fine-and-jail world, social welfare cannot have increased since the only effect of fines in such an equilibrium is to introduce two additional incentive compatibility constraints. Since no costly jail sentences are actually implemented in this equilibrium, there is no social cost associated with deterrence using jail sentences.

If exactly one type commits the crime, and that type elects to go to jail rather than pay the fine in a jail-and fine world, then once again, fines cannot have had a positive impact on social welfare since this outcome was feasible in a jail-only world.

To prove sufficiency, it needs only be noted that by assumption A8, any equilibrium of the jail-only world is still obtainable in the fine-and-jail world. Therefore, if the social planner selects the outcome in the proposition, it must be a social improvement.

QED

Proposition 1 demonstrates the substantial limits that the incentive compatibility constraint induced by private information places on the usefulness of fines. The only benefit to fines in the private information world is the avoidance of costly jail terms; the case highlighted in the proposition is the only instance where such a benefit is realized. Proposition 1 is in stark contrast to models with perfect information and perfectly enforceable fines in which introducing fines unambiguously raises welfare, both because of the increased deterrence potential and the ability to minimize jail time through the use of individually tailored fines. The incentive compatibility constraints arising from private information adversely affect both of those channels. With private information, the criminal will never pay a fine or accept a fine-jail combination that is more punitive than the maximum jail-only punishment so deterrence cannot increase with the introduction of fines. Private information also greatly reduces the social planner's ability to tailor punishments to individual criminals since the criminal's type is private information.

While Proposition 1 highlights the weakness of fines as an instrument in the presence

of private information, the following proposition points out that there is one set of circumstances in which the availability of fines unambiguously increases social welfare:

Proposition 2: The introduction of fines unambiguously increase social welfare if the jail-only social optimum involves the rich committing the crime and the poor successfully deterred.

proof: If the rich type commits the crime in the jail-only optimum, but the poor type does not, then it must be true that $\beta_R > \beta_P$. Since $\theta_R > \theta_P$, it is always possible for the social planner to introduce a fine such that the enforcement level α is unchanged, the rich continue to commit the crime, the poor continue to be deterred, and those rich types who are detected elect to pay the fine rather than go to jail. Inspection of equation (5) shows that social welfare must have improved with the fine.

QED

Proposition 2 states that fines unambiguously increase social welfare if the initial jail-only equilibrium has the rich type committing the crime with the poor type successfully deterred. This requires that the private benefits to crime for the rich are substantially greater than the private benefits to crime for the poor since the rich are more sensitive to jail punishments. Tax evasion and other white-collar crimes are possible examples of crimes that may satisfy that condition.¹⁵ Low income individuals are unlikely to have sufficient income or access to reap large benefits from such crimes.

While Propositions 1 and 2 outline general conditions under which fines will or will

¹⁵ This point assumes that not only the financial benefit, but also the marginal utility of the gains from the crime are greater for the rich than the poor, since private benefits here are expressed in utility terms.

not increase social welfare, it is perhaps more interesting to consider how changing the values of the underlying parameters of the model affects the usefulness of fines. In some cases, the logic is transparent. For instance, the greater is the social cost of imprisonment (ψ), the more valuable are fines because they provide an alternative to incarceration. Also, as Proposition 2 points out, higher private benefits to crime for the rich vis-a-vis the poor are also associated with fines being effective.

The relationship between other parameters of the model and the effectiveness of fines is more subtle. For example, extremely high private benefits to crime for the poor vis-a-vis the rich also make fines valuable since the social planner can offer a fine low enough to entice the poor type to pay it rather than go to jail without having to be concerned that the rich will also find the crime attractive now that the fine option exists.

Similar disutilities of jail for rich and poor types are also associated with fines being effective. If the rich are very averse to prison terms, the poor will be the first to commit the crime, even if their private benefit from the crime is lower than that of the rich. When fines are introduced, enforcement will need to be raised in order to deter the poor type. The cost of increased enforcement offsets the benefits of fewer jail terms served.

Fines are also more effective when the externality ($L\beta_i$) associated with a crime is small relative to the costs of imprisonment (ψ) and marginal enforcement costs ($e'(\alpha)$). In such a case, it will often not be worthwhile to punish offenders when the only form of punishment is incarceration since the social costs of preventing crime may outweigh the social costs of allowing it to occur. When fines, a less socially costly means of punishment, are introduced, it now may be socially beneficial to punish the crime. Traffic violations are

a good example of such a situation.

Section IV: Extensions and Generalizations of the Basic Model

The results of the previous sections are striking: under the particular modeling assumptions used, fines dominate jail terms as punishments with perfect information, but in many circumstances are completely useless in the presence of private information. *While the intuition that motivates the reduced efficacy of fines is quite robust, the precise result that fines provide no increment to social welfare is not.* In this section, two generalizations of the model are discussed.

One restriction on the model is the assumption that the social planner punishes criminals using fines *or* jail sentences, but not combinations of the two, although this limitation is not as restrictive as it first appears. With private information, the social planner cannot enforce a punishment that involves both J_{\max} *and* a fine since no further additional jail term can be imposed on a criminal who claims insufficient wealth and refuses to pay the fine.¹⁶ Therefore, the basic incentive-compatibility constraint is still present, even when combinations of punishments are allowed.

Combination punishments can sometimes be useful since they allow the social planner to substitute a fine for some portion of the socially costly jail sentence. The social gains

¹⁶ While J_{\max} is treated as exogenously given in the model, it is also possible that society's maximum acceptable maximum jail sentence changes with the introduction of fines. For instance, if society thinks that those criminals who refuse to pay the fine are actually able to pay the fine but are lying about their wealth, a jail term greater than J_{\max} may be justified. On the other hand, if the rich pay fines and only the poor go to jail, society may conclude that there is injustice in this arrangement and lower the maximum acceptable jail sentence.

from combinations are generally not very large, however, for two reasons. First, only the poor type serves jail-time in the fine-and-jail world by Lemma 2(ii). Any reduction in the jail-sentence of the poor must be counter-balanced by an increased level of enforcement in order for the rich to continue to be deterred. If the marginal cost of deterrence is high relative to the social costs of incarceration, access to combinations of fines and jail may not increase social welfare at all. Secondly in numerical simulations of the model, jail sentences were found to be relatively infrequent in a fine-and-jail world.¹⁷ Consequently, there is relatively little gain to society from a marginal reduction in time served.

A second extension of the model is to allow for a continuum of types. The existence of a continuum makes separating equilibria more likely; for most sets of parameter values some types commit the crime and others do not. A separating equilibrium is a necessary (but not sufficient) condition for fines to increase social welfare by Proposition 1. Not surprisingly, therefore, fines are more effective with a continuum of types. For example, as long as some types commit the crime, while others do not, it is possible to show that the introduction of fines will increase social welfare. The magnitude of that increase, however, turns out to be smaller than one might first suspect, precisely for the reasons put forth in this paper. A fine that makes the marginal criminal indifferent between paying the fine or serving the prison sentence will not be attractive to the inframarginal criminal, and therefore will have only an infinitesimal impact on social welfare. If the social planner wants to obtain a substantial social benefit from fines, a lower fine will need to be imposed so that more

¹⁷ Full discussion of the numerical simulations were included in previous versions of the paper, but have been omitted due to space constraints. Results are available from the author on request.

criminals will choose the fine. The existence of that lower fine, however, will make the crime appear attractive to others who previously were deterred by the possibility of a prison term. The social planner is then faced with either (i) increasing the level of enforcement, which is costly, or (ii) accepting a higher crime rate, which is also socially costly. While the precise social gains or losses depend on the specific case, it is clear that the presence of costs associated with the introduction of fines makes their use less attractive than is the case in a perfect information world in which fines strictly dominate prison terms as punishment.

Section VI: Conclusions

Models that assume perfect information on the part of the social planner predict a strong reliance on fines, a prediction that is inconsistent with real world practice. In this paper, it is demonstrated that private information greatly reduces the usefulness of fines due to the additional incentive-compatibility constraint that binds the social planner. While caution is warranted in drawing positive conclusions from a model as limited as the one presented here, this paper provides a possible theoretical justification for the heavy reliance on jail sentences in the real world.

Even with private information, there are numerous situations where fines are a valuable instrument to the social planner. Fines are more useful for crimes that are committed by the rich, not only because they have more resources to pay fines, but also because of the direction in which the incentive-compatibility constraint binds. Fines are also useful when the social externality of crime is small relative to the enforcement and incarceration costs. Possible examples of such crimes include traffic violations and most

other misdemeanors, both of which are frequently punished using fines in practice. Fines can also be useful when the private benefits to crime are much higher for one segment of the population than the rest since underdeterrence of the latter type is not likely to be an issue.

While the assumption that a criminal's type is completely private information is undoubtedly extreme, the assumption of perfect information on the part of the social planner is likewise unrealistic. The court has some information about an individual's wealth and disutility of jail. If used effectively, that information can certainly be welfare improving. In light of the informational restrictions, however, the relatively limited use of fines in practice may be justified.

The results of this paper have implications for policies such as mandatory sentencing guidelines that limit judicial discretion. While there are other justifications for such policies (Champion 1989), one negative consequence is that if judges cannot tailor punishments to the situation of individual criminals, it is as if the criminal's type is completely private information, exacerbating the problems associated with incentive compatibility.

References

- Andreoni, James, 1991, "Reasonable Doubt and the Optimal Magnitude of Fines: Should the Penalty Fit the Crime?" *RAND Journal of Economics* 22:385-395.
- Becker, Gary, 1968, "Crime and Punishment: An Economic Approach," *Journal of Political Economy* 76:169-217.
- Champion, Dean (editor), 1989, The U.S. Sentencing Guidelines: Implications for Criminal Justice, New York: Praeger.
- Donohue, John, and Peter Siegelman, 1994, "Is the United States at the Optimal Rate of Crime?" Mimeo, American Bar Foundation.
- Ehrlich, Isaac, 1981, "On the Usefulness of Controlling Individuals: An Economic Analysis of Rehabilitation, Incapacitation, and Deterrence," *American Economic Review* 71:307-322.
- Gillespie, Robert, 1981, "Sanctioning Traditional Crimes with Fines: A Comparative Analysis," *International Journal of Comparative and Applied Criminal Justice* 5:197-204.
- Government Accounting Office, 1992, Interstate Child Support, Washington, DC: GAO.
- Hillsman, Sally, Joyce Sichel, and Barry Mahoney, 1982, Fines in Sentencing: A Study of the Use of the Fine as a Criminal Sanction, Washington, DC: U.S. Department of Justice.
- Kaplow, Louis, 1990, "A Note on the Optimal Use of Nonmonetary Sanctions," *Journal of Public Economics* 42:245-47.
- Kaplow, Louis, 1992, "The Optimal Probability and Magnitude of Fines for Acts that Definitely are Undesirable," *International Review of Law and Economics* 12:3-11.
- Kaplow, Louis, and Steven Shavell, 1994, "Optimal Law Enforcement with Self-Reporting of Behavior," *Journal of Political Economy* 102:583-606.
- Levitt, Steven, 1995a, "The Effect of Prison Population Size on Crime Rates: Evidence from Prison Overcrowding Litigation," Forthcoming, *Quarterly Journal of Economics*.
- Levitt, 1995b, "Why Do Increased Arrest Rates Appear to Reduce Crime: Deterrence, Incapacitation, or Measurement Error?" Mimeo, Harvard University.
- Lott, John, 1987, "Should the Wealthy Be Able to 'Buy Justice'," *Journal of Political*

Economy 95:1307-1316.

Lott, John, 1992, "Do We Punish High Income Criminals Too Heavily?" *Economic Inquiry* 30:583-608.

Marvell, Thomas, and Carlisle Moody, 1994, "Prison Population Growth and Crime Reduction," *Journal of Quantitative Criminology* 10:109-140.

National Institute of Justice, 1992, Day Fines in American Courts: The Staten Island and Milwaukee Experiments, McDonald, Douglas C, (ed.). Washington, DC: National Institute of Justice.

Piehl, Anne, and John DiIulio, 1995, Does Prison Pay? Revisited," *The Brookings Review* (winter), 21-25.

Polinsky, A. Mitchell, and Daniel Rubinfeld, 1991, "A Model of Optimal Fines for Repeat Offenders," *Journal of Public Economics* 46:291-306.

Polinsky, A. Mitchell, and Steven Shavell, 1984, "The Optimal Use of Fines and Imprisonment," *Journal of Public Economics* 24:89-99.

Polinsky, A. Mitchell, and Steven Shavell, 1991, "A Note on Optimal Fines When Wealth Varies Among Individuals," *American Economic Review* 81: 618-621.

Polinsky, A. Mitchell, and Steven Shavell, 1992, "Enforcement Costs and the Optimal Magnitude and Probability of Fines," *Journal of Law and Economics* 35:133-148.

Posner, Richard, 1977, *Economic Analysis of Law*, Boston: Little-Brown.

Stigler, George, 1970, "The Optimum Enforcement of Laws," *Journal of Political Economy* 78:526-536.

Visher, C.A., 1986, "The RAND Inmate Survey: A Reanalysis," in Blumstein, A. et al. (eds.), Criminal Careers and "Career Criminal." Volume II, National Academy Press: Washington, D.C.

Waldfoegel, Joel, 1994, "The Effect of Criminal Conviction on Income and the Trust 'Reposed in the Workmen'," *Journal of Human Resources* 29:62-81.

Waldfoegel, Joel, 1995, "Are Fines and Prison Terms Used Efficiently? Evidence on Federal Fraud Offenders," *Journal of Law and Economics* 38:107-139.

Weisburd, David, Stanton Wheeler, Elin Waring, and Nacy Bode, 1991, Crimes of the Middle Classes, New Haven: Yale University Press.