Centre de recherche sur l'emploi et les fluctuations économiques (CREFÉ)
Center for Research on Economic Fluctuations and Employment (CREFE)

Université du Québec à Montréal

Cahier de recherche/Working Paper No. 100

Costly Sanctions and the Maximum Penalty Principle *

Dominique Demougin
Otto-von-Guericke Universität, Mageburg, and CREFE

Claude Fluet
Department of Economics, UQAM, and CREFE

December 1999

Demougin: Faculty of Management and Economics Universittsplatz 2 Postfach 41 20, 39016 Magdeburg, Germany, email `demougin@ww.uni-magdeburg.de`
Fluet: CREFE, P.O. Box 8888, Downtown Station, Montreal H3C 3P8, Canada, email: `fluet.claude-denys@uqam.ca`

* The authors are thankful to FCAR and SSHRC for financial supports and to the participants of the "Mechanism Design and Incentive" session at the 1999 EEA meeting in Santiago de Compostela for useful comments

Résumé:
Cet article analyse le problème de la dissuasion des comportemenmts indésirables dans un contexte d'aléa moral avec aversion au risque, information imparfaite et coûts de sanction. Nous montrons que, si les sanctions imposées aux individus sont une pure perte sociale, la politique utilitariste optimale consiste à utiliser un mécanisme de sanction dichotomique satisfaisant le principe de la sanction maximale. Si les sanctions sont pécuniaires mais qu'imposer des sanctions implique un coût en ressource suffisamment élevé, la sanction maximale permise devrait également être imposée avec une probabilité positive. Comme justification possible de sanctions limitées, nous analysons la politique de dissuasion optimale avec une fonction de bien-être rawlsienne. Le sanction maximale est dans ce cas inférieure à celle d'une politique utilitariste, mais elle est imposée plus fréquemment.

Abstract:
We study the problem of deterring undesirable behavior in a moral hazard framework with risk averse individuals, noisy information and costly sanctions. We find that, if sanctions are a pure loss, a utilitarian society should use a bang-bang penalty scheme satisfying the maximum penalty principle. If sanctions are monetary but imposing sanctions involves a sufficiently large resource cost, the maximum feasible sanction should also be imposed with positive probability. As a possible justification for endogenously limiting sanctions, we derive the optimal penalty scheme under a Rawlsian welfare function. The maximum sanction actually imposed is then smaller than in the utilitarian case, but it is imposed more frequently.

Keywords:
Deterrence, optimal enforcement, moral hazard, maximal penalty, Rawl's criterion
JEL classification: D8, K1

Costly Sanctions and the Maximum Penalty Principle

Dominique Demougin* and Claude Fluet†‡

December 1999

Abstract. We study the problem of deterring undesirable behavior in a moral hazard framework with risk averse individuals, noisy information and costly sanctions. We find that, if sanctions are a pure loss, a utilitarian society should use a bang-bang penalty scheme satisfying the maximum penalty principle. If sanctions are monetary but imposing sanctions involves a sufficiently large resource cost, the maximum feasible sanction should also be imposed with positive probability. As a possible justification for endogenously limiting sanctions, we derive the optimal penalty scheme under a Ralwsian welfare function. The maximum sanction actually imposed is then smaller than in the utilitarian case, but it is imposed more frequently. [JEL D8, K1]

Keywords. Deterrence, optimal enforcement, moral hazard, maximal penalty, Rawls' criterion.

1. Introduction

This paper studies the problem of deterring undesirable behavior in a framework characterized by risk aversion, noisy information and costly sanctions. Due to imperfect information, any sanction scheme involves risk-bearing costs. Imposing sanctions also involves a resource cost, assumed to be increasing in the magnitude of sanctions. For a given information structure, the issue is therefore to determine how sanctions should vary with respect to the likelihood of undesirable behavior, taking into account risk-bearing costs and the resource costs of imposing sanctions. We find that, if the sanctions incurred by individuals are a pure social loss, a utilitarian society

*Otto-von-Guericke University (demougin@ww.uni-magdeburg.de)

†Université du Québec à Montréal, CREFE and CRT (fluet.claude-denys@uqam.ca)

‡The authors are thankful to FCAR and SSHRC for financial support and to the participants of the "Mechanism Design and Incentive" session at the 1999 EEA meeting in Santiago de Compostela for useful comments.

should use a bang-bang penalty scheme irrespective of the individuals' attitude towards risk. Specifically, when penalties are imposed, they are set as high as possible; furthermore, they are used as seldom as possible — ie, only when the likelihood of undesirable behavior is particularly high. When sanctions are not a pure loss but involve a sufficiently large resource cost, the optimal sanction scheme is quasi bang-bang and the maximum feasible penalty is imposed with positive probability. In either case, the maximum feasible sanction is a binding constraint and it follows that allowing harsher maximum penalties would be welfare improving.

Our analysis puts in a new light Becker's (1968) well known maximum penalty principle. The original argument assumes risk-neutral individuals, monetary sanctions and stochastic monitoring as defined by a probability of detecting delinquents. Due to the risk-neutrality assumption, deterrence depends only on the expected fine. Varying the probability of detection is socially costly because it requires resources, whereas fines are costless in a utilitarian perspective as they are pure transfers. To minimize social costs subject to maintaining compliance, the probability of punishment should therefore be reduced as much as possible and the fine raised to the maximum feasible level. Under risk-neutrality, this result can be generalized to costly penalties (e.g. prison rather than fines) if social costs are roughly proportional to the sanctions born by individuals, and to the case where innocents can be mistakenly punished.[1] It can also be generalized to risk aversion, provided complete deterrence is achieved and ex post information is perfect. The argument still holds because sanctions are then never actually imposed. However, the standard view is that the argument brakes down when there is risk aversion combined with incomplete compliance or with a possibility of erroneous penalization. In such cases, there is a tradeoff between detection costs and risk-bearing costs with the result that it may not be optimal to impose the maximum feasible sanction.[2]

The main objective of this paper is to show that this line of reasoning is not always correct. For that purpose, we start with a model that is, in some sense, the dual of Becker's original framework. First, we take the information structure as given; hence, information costs are ignored. Secondly, information is noisy; that is, even if an individual conforms to the prescribed standard of behavior, there is some probability that he will be penalized. Thirdly, agents are risk averse so that any incentive compatible penalty scheme (given that information is noisy) generates risk-bearing costs for the individual. Finally, imposing sanctions implies a real resource cost. If the information structure is sufficiently rich, deterrence can be achieved by many different sanction schemes in terms of the likelihood of undesirable behavior.

[1] See Shavell (1987), Kaplow (1990), Polinsky and Shavell (1992) and Shavell and Kaplow (1994).

[2] See Polinsky and Shavell (1979) and Kaplow (1992). There are many other explanations for non-maximal sanctions, but these ususally rely on more complex enforcement games than considered here. A useful survey is Garoupa (1997).

For instance, deterrence may be possible by punishing harshly but very seldom, or by imposing light sanctions more frequently. This suggests that there could be a tradeoff between risk-bearing costs and the resource costs associated with sanctions. However, we show that these two sources of costs interact in a peculiar way. In essence, risk aversion is relevant as a possible explanation for non-maximal sanctions only if the sanction imposed on an individual provides a sufficiently large payback to the rest of society. When the payback is negligible, the best utilitarian deterrence scheme is to penalize infrequently and harshly, irrespective of the degree of risk aversion and of the possibility of penalizing the innocent.

To illustrate, suppose a barely monetized peasant economy. Abstracting from corporal punishment, the only feasible sanctions take the form of destroying a peasant's wealth, say by burning a part of his crops: because means of transport are primitive, it is not possible to seize an individual's crop so as to redistribute it to the rest of society. The magnitude of a sanction is the amount of crop destroyed and peasants are risk averse in the usual sense: they prefer a certain crop to a random one with the same expected value. Suppose also that the act of destroying crops is itself costless; ie, the social resource cost of a sanction is simply the loss born by the individual in terms of the amount of crop destroyed. Then the best enforcement policy, from the point of view of maximizing the expected utility of the peasants' themselves, is to punish infrequently but, if one is to be punished, to destroy one's crops to the fullest extent allowable. If in addition the act of destroying crops requires resources, it could then be preferable to impose less than maximal sanctions. However, this would arise not because of the peasants' risk aversion, but only because enforcement costs could be increasing disproportionately fast with the magnitude of sanctions. By contrast, in a modern monetary economy where "crops" can be seized at modest collection costs, the loss incurred by the penalized individual is compensated by a payback to the rest of society. It is only in this case that risk aversion matters in the design of the optimal sanction scheme. Nevertheless, if collection costs are sufficiently large, the best utilitarian scheme still requires imposing the maximum feasible penalty with positive probability, even though it may also impose less than maximal sanctions with some probability.

The tale we have just told may contain an element of sociological or historical truth. As a crude generalization, it may be that sanctions tend to be extreme in societies where means of transferring wealth are limited. Still, even in such societies, it is safe to say that one does not always observe enforcement systems where maximum feasible sanctions (whatever that may mean) are actually imposed. One possibility is that maximum *allowable* penalties refer to non-utilitarian considerations. As emphasized in Ehrlich (1982), there may be other objectives than utilitarian efficiency, for instance justice as avoidance of legal error or justice as ex post equality under the law. In the situation considered in the present paper, avoiding legal error *per se* means less frequent penalties and therefore harsher sanctions if deterrence is to be maintained. Thus, it does not provide an endogenous upper bound on sanctions.[3] But such an upper bound may be explained by a concern for the ex post distributional consequences of legal error. As an alternative, in order to endogenize the maximum penalty actually imposed, we analyze the optimal penalty scheme under a Rawlsian welfare function. Such a scheme seeks to maximize the ex post utility of the most penalized individual, subject to providing the required deterrence. When sanctions are a pure loss, the optimal minimax penalty scheme is bang-bang, as in the utilitarian case, but the sanction imposed is now the smallest one compatible with deterrence. Of course, the sanction must now be imposed more frequently so that, in a sense, there is a greater probability of legal error. Interestingly, an individual is then punished whenever it is more likely than not that he has misbehaved. When sanctions are not a pure loss, the "standard of proof" is lower still and individuals face an even greater probability of sanction.

Most of the literature on optimal enforcement, with or without costly sanctions, has dealt with dichotomous penalty schemes; that is, either the individual is not penalized or he is penalized (possibly erroneously) and the issue is to determine the magnitude of the sanction, given the tradeoff with detection costs. By contrast, we abstract from monitoring costs but assume that more complex information is obtained with respect to the individual's behavior. The problem is then to determine a penalty schedule in terms of what is observable about the individual. Our framework is therefore similar to the standard principal-agent setting with moral hazard. One difference is that there is no participation constraint with respect to the agent. By the mere fact of being a member of society, one must partake in the "penalty-game". Also, since the optimal penalty scheme is designed so as to maximize the utility of the agent or representative individual, a sanction scheme can be interpreted as self-imposed for the sole purpose of inducing optimal behavior.

The remainder of the paper is organized as follows. In the next section, we present the model. In section 3, the optimal penalty scheme is derived for the case where sanctions are a pure loss. In section 4 we assume a positive return on sanctions. Section 5 derives the optimal maximin sanction scheme. In section 6 we discuss several extensions and conclude.

2. The Model

Society is composed of a large number of identical individuals, each of whom can undertake one of two possible actions. By choosing the action a_0 rather than a_1, an individual generates an external benefit b but incurs a private cost c. When everyone does the same, each individual's utility is $u(w) - c + b$ where $u(w)$ is the utility of

[3] This point is also made in Shavell and Kaplow (1994).

wealth function, assumed strictly concave, and where the cost and per capita benefit from the action are expressed in utility terms. Assuming $b > c$, the action a_0 is therefore the socially desirable one, although everyone has an incentive to free-ride by choosing a_1. This saves the private cost and has negligible effect on the per capita benefit because of the large number of individuals in society.

Behavior is unobservable, but for each individual it is possible to observe a signal with positive density functions $f_i(x)$ on the interval $[0, \overline{x}]$, where i is 0 or 1 depending on the individual's choice of action. The signal satisfies the monotone likelihood ratio property (hereafter MLRP):

$$\frac{d}{dx}\left(\frac{f_1(x)}{f_0(x)}\right) > 0, \qquad x \in (0, \overline{x}). \tag{1}$$

Because the densities have common support, the signal never reveals perfectly an individual's behavior, but the socially undesirable action appears relatively more likely when larger values of x are observed.

Assuming MLRP is without loss of generality. Suppose the actual information structure consists of a multidimensional signal. Then x can be defined as any increasing function of the likelihood ratio of a_1 relative to a_0 in terms of the multidimensional signal. Such an x is a sufficient scalar statistic of the available information; if density functions exist, they satisfy (1). Also, the underlying signal need not consist only of pure information. For instance, it may be that both actions produce an observable random level of external harm, so that the external benefit b can be interpreted as the difference between the expected harm under a_1 and under a_0. The underlying signal would then include the actual harm caused by the individual, as well as other possible evidence.

Society's problem is how to enforce the socially desirable action (or deter the more harmful one) by penalizing individuals on the basis of what is observable, and to decide whether the first-best action is in fact worth implementing in this second-best world. Let the penalty schedule be denoted by $s(x)$, this being understood as a monetary equivalent from the point of view of the individual. When the sanction is a fine, T denotes the average per capita amount levied from sanctions and which can be returned to individuals as a lump-sum payment. A utilitarian society wishing to implement a_0 looks for the sanction scheme that maximizes the expected utility of the representative individual. This problem is written as

$$\max_{s(x),T} \int_0^{\overline{x}} u(w + T - s(x)) f_0(x)\, dx \quad \text{subject to}$$

$$\int_0^{\overline{x}} u(w + T - s(x)) f_0(x)\, dx - c \geqslant \int_0^{\overline{x}} u(w + T - s(x)) f_1(x)\, dx\,, \tag{2}$$

$$T \leq \int_0^{\overline{x}} h(s(x)) f_0(x)\, dx, \tag{3}$$

$$0 \leq s(x) \leq w - \underline{w} + T\,. \tag{4}$$

The first condition is the incentive compatibility constraint. It states that one is better off undertaking a_0 rather than a_1. In the next constraint, the lump-sum amount T that can be paid back depends on a function h describing the return on sanctions. We assume $h(0) = 0$ and marginal return $h'(s) \in [0, 1)$ for all s. When the return on sanctions is identically zero, penalties can be interpreted as nonmonetary. Even when the marginal return on sanctions is positive, there is always a dead-weight loss from imposing sanctions and perfectly costless fines are specifically excluded. The condition $h' < 1$ reflects collection expenses, prosecution costs or any other "transaction costs" in the sense that an increased disutility to penalized individuals is not entirely compensated by a pay-back to the rest of society.[4] For the main part we assume $h' \geq 0$, although one could also argue that there are cases where the return is actually negative — the total marginal resource cost is then greater than the loss born by the individual. This possibility is discussed briefly in section 6.

The third constraint states that sanctions are nonnegative and bounded above. The maximum feasible penalty consists in driving an individual's utility of wealth down to $u(\underline{w})$, where the minimum allowable wealth $\underline{w}$ or wealth equivalent is taken as exogenous. Since individuals have initial wealth w and are paid back T, this translates into the upper bound on sanctions described in the constraint. We assume that $u(\underline{w})$ and $u'(\underline{w})$ are finite. For the time being, this is simply interpreted as describing a situation where allowable sanctions are limited.

To simplify notation, write the likelihood ratio as $R(x) = f_1(x)/f_0(x)$. A likelihood ratio ranges from less than to greater than unity and we denote by x^* the realization of the signal satisfying $R(x^*) = 1$. Writing the respective cdf's as F_0 and F_1, we assume

$$[F_0(x^*) - F_1(x^*)][u(w) - u(\underline{w})] > c. \tag{5}$$

As will become obvious in the next section, this condition guarantees that allowable penalties are sufficient to implement the required action, given the informativeness of the signal.

Let $\overline{u}$ denote the expected utility of wealth under the optimal sanction scheme. Because information is imperfect and individuals are risk averse, $\overline{u} < u(w)$. From a utilitarian point of view, the first-best action is worth implementing only if the net benefits of the action are greater than the welfare loss due to the deterrence system;

[4] In Polinsky and Shavell (1992), such costs are taken as proportional to the sanction and are referred to as variable enforcement costs as opposed to the fixed costs pertaining to the monitoring system.

that is, only if $\overline{u}-c+b > u(w)$. The value of $\overline{u}$ clearly depends on the cost of sanctions, through the return function $h(s)$, and it may also depend on the maximum feasible sanction as defined by the lower bound $\underline{w}$.[5] A basic result of the present paper is that these two considerations are related: when the dead-weight loss of sanctions is large (h' sufficiently small), the maximum feasible sanction constraint is binding. As a consequence, an exogenous decrease in $\underline{w}$ increases $\overline{u}$ and may make the first-best action worth implementing when otherwise it would not be.

3. Nonmonetary Sanctions

We first examine the case where sanctions are a pure loss. As suggested above, the real issue here is not, strictly speaking, whether sanctions are pecuniary as opposed to other forms (exile, imprisonment, etc.), but whether the imposition of sanctions generates a net revenue that can be redistributed. In the present case, sanctions are equivalent to destroying wealth and the return T is zero irrespective of the level of sanctions. Our first proposition states that the optimal scheme is bang-bang and that it involves the maximum feasible sanction (all proofs are in the appendix).

Proposition 1. *When $h(\cdot) \equiv 0$, the optimal sanction satisfies $s(x) = 0$ if $x \leq x_c$ and $s(x) = w - \underline{w}$ if $x > x_c$, where x_c is defined by*

$$[F_0(x_c) - F_1(x_c)][u(w) - u(\underline{w})] = c, \quad x_c \in (x^*, \overline{x}). \tag{6}$$

The result is represented in figure 1. MLRP implies first-order stochastic dominance, which means that $F_0 - F_1$ is positive over the interval $(0, \overline{x})$; it also implies that this expression is quasiconcave so that it reaches a maximum at x^*. Equation (6) is the incentive compatibility condition for a bang-bang scheme using the maximum feasible penalty. Because of the quasiconcavity, the equation has two solutions but only $x_c > x^*$ is relevant since it implements the required action with a smaller probability of sanction.Enforcing the prescribed action is also possible with less extreme bang-bang schemes, such as the one with the threshold $\widehat{x}$ and the sanction $\widehat{s}$ in the figure. As should be obvious, the optimal scheme described in proposition 1 is the one with the smallest probability of sanction among all such binary schemes. This statement can be generalized to all feasible sanction schedules, whether bang-bang or not.

Corollary 1. *When $h(\cdot) \equiv 0$, the solution to the optimal sanction problem minimizes $\Pr[s(x) > 0 \mid a_0]$ over all feasible functions $s(x)$ implementing a_0.*

[5] More generally, it will of course also depend on the degree of risk aversion and on the informativeness of the signal.

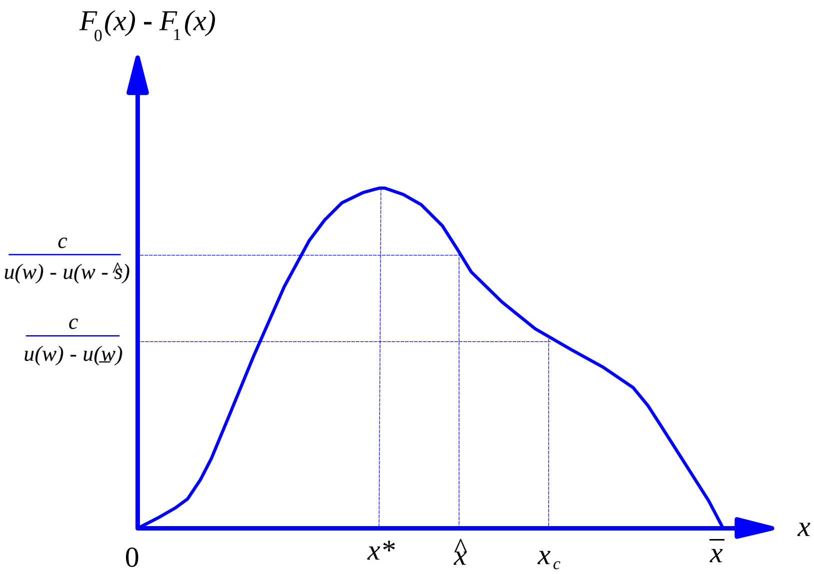

Figure 1

Both propositions taken together amount to Becker's maximum penalty principle: sanctions are imposed as seldom as possible but, when they are, individuals are punished as hard as possible. For instance, if x represents the external harm caused by the individual, no sanction should be imposed if the harm is smaller than some critical level, but the individual should be penalized to the fullest extent possible whenever the harm is greater than the threshold. From a statistical point of view, the optimal scheme is like a Neyman-Pearson efficient test (minimizing the probability of "rejecting" the hypothesis a_0 when in fact a_0 is true), except that the degree of confidence constraint of the test has been replaced by an incentive compatibility condition.

The surprising part of these results is that, irrespective of the individuals' degree of risk aversion, the sanction is either maximal or zero. In a principal-agent setting with a risk-neutral agent facing a liability limit, the optimal contract is a bonus scheme penalizing the agent to the extent of his liability limit when information is unfavorable.[6] Why is the optimal scheme similar to what would obtain under risk neutrality? The explanation lies in the nature of sanctions. Because they are pure losses, sanctions can be of no gain to anyone; since redistribution is impossible, there are therefore no opportunities for risk-sharing. Formally, the result follows from the

[6]To be precise, this is true whenever the quality of information and the liability limit are such that the first-best is unattainable (e.g., Kim, 1995, and Demougin and Fluet, 1998).

fact that the optimization problem is linear in the disutility imposed on individuals. To see this, write the disutility of sanctions as $\delta(x) = u(w) - u(w - s(x))$. The original optimization program can then be reformulated as:

$$\min_{\delta(x)} \int_0^{\overline{x}} \delta(x) f_0(x)\, dx \quad \text{subject to}$$

$$\int_0^{\overline{x}} \delta(x) f_0(x)\, dx + c \leq \int_0^{\overline{x}} \delta(x) f_1(x)\, dx, \tag{7}$$

$$0 \leq \delta(x) \leq u(w) - u(\underline{w}). \tag{8}$$

This problem being linear in the decision variable, its solution is bang-bang. From this, it is straightforward to obtain the optimal $s(x)$ and this will of course also be bang-bang.

Another similarity with the standard Beckerian result is that the optimal scheme minimizes the expected resource cost of providing deterrence. Equivalently, it maximizes the expected wealth of the representative individual.

Corollary 2. *When $h(\cdot) \equiv 0$, the optimal scheme minimizes $\int_0^{\overline{x}} s(x) f_0(x)\, dx$ over all feasible functions $s(x)$ implementing a_0.*

The intuition for the latter result is as follows. The optimal scheme uses the most extreme outcomes: the sanction is either zero or maximal; furthermore, it involves the smallest probability of imposing the maximal sanction, which also implies the largest probability of no sanction. Therefore, all other schemes satisfying the incentive compatibility condition as an equality must be less risky. By a stochastic dominance argument, since these other schemes are not preferred, it must be that the benefits of a smaller risk are not enough to make up for a smaller expected wealth.

Clearly, the upper bound on sanctions is a binding constraint in the solution to the optimal sanction problem. It follows that relaxing this constraint (an exogenous decrease in $\underline{w}$) will increase the expected utility of the representative individual. This can be checked directly by noting that

$$\overline{u} = u(w) - [u(w) - u(\underline{w})][1 - F_0(x_c)]. \tag{9}$$

A smaller $\underline{w}$ increases the loss when the individual is penalized, but it is easily shown that this is more than compensated by the decrease in the probability of imposing the sanction.[7]

[7] As noted earlier, $\overline{u}$ will in general also depend on the informativemess of the signal and on the degree of risk aversion. A more informative signal is essentially equivalent to an upward shift in the $F_0 - F_1$ curve of figure 1 (see Demougin and Fluet, 1999); this allows a smaller probability of sanction and therefore increases $\overline{u}$. On the other hand, changes in risk aversion are irrelevant here since what matters is the utility difference between $u(w)$ and $u(\underline{w})$ compared to the disutility cost c. Any utility of wealth function can be normalized so as to equal $u(w)$ and $u(\underline{w})$ at w and $\underline{w}$ respectively.

4. Costly Monetary Sanctions

With a positive return on sanctions, an individual's wealth for a realization x of the signal is $w + T - s(x)$, where T is the overall per capita amount raised. The revenue generated by sanctions introduces a trade-off that did not exist when sanctions were pure losses. Modifying $s(x)$ over some interval now changes T, which in turn affects the individual's utility in all other states of the world. By contrast, when sanctions were pure losses, changes in $s(x)$ over some states could only affect utility in those states. Risk aversion is relevant to this trade-off and the consequence, in general, is a more gradual optimal sanction schedule.

The fact that sanctions generate revenue implies that the required action can be implemented by punishing less harshly. To see this, consider bang-bang schemes even though these may not be optimal here. With no sanction if $x \leq \widehat{x}$ and a sanction equal to $\widehat{s} \leq w - \underline{w} + T$ if $x > \widehat{x}$, the resource and incentive compatibility constraints are

$$T = (1 - F_0(\widehat{x}))h(\widehat{s}). \tag{10}$$

$$[F_0(\widehat{x}) - F_1(\widehat{x})][u(w+T) - u(w+T-\widehat{s})] = c, \quad \widehat{x} \in (x^*, \overline{x}). \tag{11}$$

Equation (10) is the total return on sanctions. Suppose the maximum feasible sanction $\widehat{s} = w - \underline{w} + T$. In equation (11), the difference in utility between the sanction and no-sanction states is then equal to $u(w+T) - u(\underline{w})$. Because $T > 0$, the wedge is greater than in proposition 1. From figure 1 it is easily seen that this implies $\widehat{x} > x_c$, which means that the individual is punished less often. Expected utility therefore improves in two ways: although the sanction is the same in the sense that utility is brought down to the allowable minimum, the individual is punished less often; furthermore, when he is not punished, he is better off due to $T > 0$. Obviously, one could reduce $\widehat{s}$ slightly from the maximum feasible level in such a way that the individual is punished both less often and less harshly than when sanctions are a pure loss.

Sanctions remain costly in the present situation since the marginal return on sanctions is less than unity. We impose no conditions on the curvature of the return function $h(s)$ because none seems particularly convincing. A convex return function amounts to economies of scale in the imposition of sanctions, which is not unreasonable in some situations. With sufficient convexity, the optimal scheme could be bang-bang with the maximum feasible penalty as in proposition 1. This follows because large sanctions are then relatively less costly, abstracting from risk aversion considerations. On the other hand, a concave return function is also not unreasonable. For instance, it may be that small sanctions are monetary while larger ones tend to be nonmonetary or involve a larger marginal dead-weight loss. Concavity in the return function tends to reinforce the effect of risk aversion. The formulation of the next proposition takes both possibilities into account.

Proposition 2. *When $h' > 0$, the optimal sanction $s(x)$ is nondecreasing and $s(x) = 0$ with positive probability; if h' is sufficiently small, $s(x) = w - \underline{w} + T$ with positive probability for some $x > x^*$. In the range, if one exists, where $s(x) \in (0, w - \underline{w} + T)$,*

$$\frac{ds}{dx} = -\left(\frac{\lambda R'}{1 + \lambda(1 - R)}\right)\left(\frac{u''}{u'} + \frac{h''}{h'}\right)^{-1} > 0 \tag{12}$$

where λ is a positive constant such that $1 + \lambda(1 - R) > 0$ in that range.

Individuals are never penalized for small values of the signal. This feature follows from the assumption that $h' < 1$ and it is similar to the positive deductible provision in optimal insurance contracts when there are administrative costs.[8] At interior solutions, the sanction is increasing in the value of the signal, a standard result given that the likelihood ratio $R(x)$ is strictly increasing. Thus, if the signal amounts to the external harm caused by an individual, there should be no penalty for small levels of harm but otherwise the penalty should be increasing in the harm done. The inequality in (12) requires that at interior solutions

$$\frac{u''(w + T - s(x))}{u'(w + T - s(x))} + \frac{h''(s(x))}{h'(s(x))} < 0. \tag{13}$$

This condition is a concavity requirement for an interior solution to obtain in some range. If there are "economies of scale" in the imposition of sanctions (ie, $h'' > 0$) and if these are always large relative to risk aversion, the solution is bang-bang with $s(x)$ equal to zero or to the maximum feasible sanction.

The main point of the proposition is the result concerning the possibility that the maximum penalty constraint be binding. As just seen, the optimal scheme is always bang-bang and the constraint is therefore binding if the concavity requirements do not hold. Abstracting from this case, the proposition states that the maximum feasible penalty is imposed with positive probability whenever sanctions are sufficiently costly. The optimal scheme is then quasi bang-bang, with no sanction for small values of x, increasing sanctions over an intermediate range and maximum feasible penalty for large values. From the principal-agent literature (see Mirrlees, 1974), it is well known that "limited liability" constraints may be binding when likelihood ratios are unbounded. The point made here is that, when transaction costs are sufficiently large, the constraint is necessarily binding even with a bounded likelihood ratio. As in the case where sanctions are a pure loss, individuals could then be made better-off ex ante if the maximum feasible penalty constraint were relaxed.[9]

[8] See Raviv (1979) and Huberman, Mayers and Smith (1983). These papers do not consider ex ante moral hazard.

[9] If infinite penalties were allowed, it would be possible to implement the action with an arbitrarily

5. Maximin Penalty Schemes

As noted earlier, there are many explanations for the fact that maximal feasible sanctions are seldom observed, but most rely on a more complex enforcement framework than the one described here. For instance, limiting the magnitude of sanctions may reduce avoidance activities on the part of delinquents (Malik, 1990). When enforcement is delegated, non-maximal sanctions may prevent overenforcement when enforcers are concerned with cost-efficiency (Bose, 1995). Also, limited sanctions are useful in preventing the corruption of enforcers given that higher sanctions are likely to induce greater bribes (Becker and Stigler, 1974).[10] In a repeated game context, when sanctions are costly and society cannot commit, non-maximal sanctions can be interpreted in terms of the tradeoff between reputation building for deterrence purposes and the ex post cost of imposing sanctions (Boadway et al., 1996).

Abstracting from these extensions, non-maximal sanctions are usually attributed to risk-bearing costs and to the resource costs associated with sanctions. From the preceding sections, these arguments clearly have limited explanatory power in the present context. Our individuals are risk averse, they are erroneously penalized and sanctions are costly. Nevertheless, the maximal sanction is imposed with positive probability whenever sanctions are sufficiently costly. The fact that observed sanctions are limited could also be interpreted in terms of a social "disutility" for the magnitude of sanctions, but then such considerations must necessarily be non-utilitarian. If the maximum allowable sanction corresponds to a social norm — say, a moral reluctance to impose harsher sanctions — that norm will have an opportunity cost from a utilitarian point of view. A strong reluctance to impose harsh sanctions implies a lower expected utility $\overline{u}$; furthermore, since enforcement costs are increased, it means that society will in some cases prefer to forego the benefits from implementing the desirable action, even though the representative individual would be better-off if the norm were relaxed and the action implemented. One could also analyze the norm itself as the result of a tradeoff between non-welfarist considerations and utilitarian considerations pertaining to $\overline{u}$.

This section examines the case where society's concern for the harshness of sanc-

small probability of sanction. However, this may not be optimal. From (10) and (11) it can be verified that, for any return function h, as $\widehat{s} \to w - \underline{w} + T$ where $\underline{w}$ is finite but $u(\underline{w}) = -\infty$,

$$\overline{u} \to u(w) + \frac{c}{\lim\limits_{x \to \overline{x}} R(x) - 1}$$

The limit is the first-best outcome only if the likelihood ratio is unbounded above. A best scheme is then pure threat and sanctions are (almost) never imposed. Otherwise, a more gradual scheme may do better even though infinite sanctions are allowed.

[10] In hierarchical agencies, limiting the stakes is a standard collusion-proofness device (Tirole, 1992).

tions — or the fact that they are erroneously imposed — is formalized through a welfare function satisfying Rawl's criterion. Ex post, the worst-off individual is the one who is the most penalized. A Rawlsian enforcement policy therefore seeks to minimize the maximum sanction imposed on individuals, subject to the incentive compatibility and resource constraints; equivalently, it maximizes the ex post utility of the most penalized individual. Formally, the optimal maximin scheme solves:

$$\max_{s(x),T} \min_{x} u(w + T - s(x))$$

subject to (2), (3) and (4). The constraints are the same as before, with the lower bound $\underline{w}$ interpreted as a physically feasible minimum; given the assumption on the informativeness of the signal, the constraint $s(x) \leq w - \underline{w} + T$ will in fact never be binding.

Proposition 3. *Let $s(x)$ be the optimal Rawlsian sanction and define $s_m = \max_x s(x)$. When $h' > 0$, the sanction is nondecreasing and $s(x) = 0$ with positive probability; furthermore, there exists $x_c \in (0, x^*)$ such that $s(x) = s_m$ for all $x > x_c$. In the range, if one exists, where $s(x) \in (0, w - w_m + T)$,*

$$\frac{ds}{dx} = -\frac{R'}{(1-R)} \left(\frac{u''}{u'} + \frac{h''}{h'} \right)^{-1} > 0. \tag{14}$$

The overall form of the sanction schedule is similar to the one obtained with a utilitarian welfare function and the same remarks apply with respect to the concavity requirements for non bang-bang solutions. The difference is that the maximum sanction actually imposed, driving one's utility down to $u(w + T - s_m)$, is now endogenous. The Rawlsian scheme also differs from the utilitarian one in the fact that the maximal sanction is imposed more frequently. In the utilitarian policy, the lower bound $\underline{w}$ is reached (if at all) only for some $x > x^*$. In the Rawlsian policy, the minimum $w + T - s_m$ is reached for all values of the signal greater than some x_c, where $x_c < x^*$. The difference between the two schemes is seen most clearly when there is zero return on sanctions.

Proposition 4. *When $h(\cdot) \equiv 0$, the optimal Rawlsian sanction satisfies $s(x) = 0$ if $x \leq x^*$ and $s(x) = s_m$ if $x > x^*$, where s_m is defined by*

$$[F_0(x^*) - F_1(x^*)][u(w) - u(w - s_m)] = c. \tag{15}$$

When sanctions are a pure loss, the optimal Rawlsian sanction is bang-bang, as in the utilitarian case, but it now involves the lightest sanction consistent with enforcing the desired behavior (see figure 2). Because individuals are penalized less than in the utilitarian scheme, they must of course be punished more frequently. The sanction is now imposed whenever it is more likely than not that the undesirable action has been chosen.

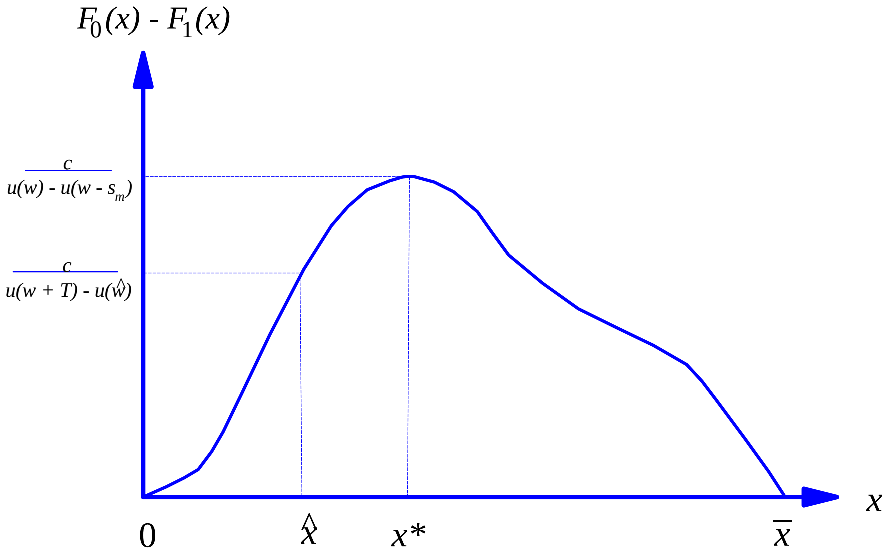

Figure 2

The figure is also useful in understanding the effect of a positive return on sanctions. Consider the problem of determining the optimal Rawlsian sanction, subject to the restriction that the scheme be bang-bang. Assume the sanction is positive only for $x > \widehat{x}$ and denote by $\widehat{w}$ the net wealth of a penalized individual, where these satisfy

$$T = (1 - F_0(\widehat{x}))h(w - \widehat{w} + T), \tag{16}$$

$$[F_0(\widehat{x}) - F_1(\widehat{x})][u(w + T) - u(\widehat{w})] = c. \tag{17}$$

Suppose $\widehat{x} = x^*$. Comparing with the preceding proposition, we must have $\widehat{w} > w - s_m$ because $T > 0$. Now consider a small variation in $\widehat{x}$ around x^*. This has only a second-order effect on $F_0 - F_1$ because that expression reaches a maximum at x^*. However, from (16), an increase in $\widehat{x}$ reduces T. To maintain incentives, $\widehat{w}$ must therefore decrease and the penalized individual is made worse off. Conversely, decreasing $\widehat{x}$ penalizes more often and increases T. From (17), this allows $\widehat{w}$ to increase so that the penalized individual is better off. The argument can be extended to non marginal changes in $\widehat{x}$. Thus, when there is a return on sanctions, the optimal critical $\widehat{x}$ for a bang-bang scheme under the Rawlsian criterion should satisfy $\widehat{x} < x^*$ as depicted in figure 2, in contrast to what was obtained with a utilitarian welfare function. The reason is that levying more fines makes it possible to maintain a sufficient wedge between sanction and no-sanction states, without requiring a harsh penalty. A non bang-bang scheme may be preferable still because it may allow $\widehat{w}$ to be increased further.

Consider now whether or not the first-best action should be implemented. With a Rawlsian welfare function, implementing the action is worthwhile only if no one is made worse off; that is, only if $u(w - s_m) - c + b \geq u(w)$. When this condition holds, even the most penalized individual is ex post at least as well off as without implementation. Recalling that $\overline{u}$ is the expected utility of wealth under the best utilitarian enforcement policy, it is easily seen that $u(w - s_m) < \overline{u}$, otherwise utilitarians would in fact prefer the Rawlsian scheme.[11] Thus, a Rawlsian society is less likely to implement the first-best action than a utilitarian one.

6. Concluding Remarks

So far we have assumed $h' \geq 0$. As noted earlier, there are cases where the marginal return on sanctions is negative, at least in some range. Thus, fines may involve only limited collection expenses while sanctions greater than the individual's financial wealth (imprisonment, etc.) imply substantial costs. Obviously, imposing the maximum feasible sanction may not be optimal if the marginal return is positive at low sanction levels but negative at high levels. However, suppose as in Shavell (1987) that all sanctions are nonmonetary and that imposing sanctions is itself costly, in addition to the loss born by the individual. Let this additional cost be denoted by $k(s)$, with $k(0) = 0$ and $k' > 0$. Suppose also that the cost of imposing sanctions is financed by a per capita levy

$$t = \int_0^{\overline{x}} k[s(x)] f_0(x)\, dx, \tag{18}$$

so that at x an individual's wealth equivalent is $w - t - s(x)$. Finally, suppose that the utility of wealth function has non-increasing absolute risk aversion. It can then be shown that a sufficient condition for the results in proposition 1 to hold is

$$\frac{k''(s)}{k'(s)} \leq -\frac{u''(w)}{u'(w)} \qquad \text{all } s > 0. \tag{19}$$

That is, the optimal scheme is then bang-bang with a sanction either equal to zero or to the maximal feasible level $w - \underline{w} - t$. The optimal scheme also minimizes the

[11] Since the Rawlsian $s(x)$ is nondecreasing, $\int_0^{\overline{x}} u(w + T - s(x)) f_0(x)\, dx > u(w - s_m)$. Hence, $u(w - s_m) \geq \overline{u}$ would contradict the fact that $\overline{u}$ is defined with respect to the best utilitarian policy.

probability of imposing a sanction, subject to the incentive compatibility condition; furthermore, it minimizes the per capita expected resource cost

$$\int_0^{\overline{x}} \{s(x) + k[s(x)]\} f_0(x)\, dx.$$

The condition (19) is satisfied if the marginal cost of imposing sanctions does not increase too fast with the level of sanctions, compared with the representative individual's risk aversion.

The analysis can be extended in several directions. For instance, adverse selection can be introduced by assuming more than one type of individual, where types differ in the private cost of undertaking the more beneficial action or in the expected external benefits that this action generates. As in a some of the literature on optimal enforcement, one can then examine the case where it is not be desirable to implement action a_0 for some types. With noisy information, since types and choice of action are not be observable, sanctions are erroneously imposed for two reasons: upon observing the signal x, it is not known whether an individual undertook a_0 rather than a_1 nor in fact whether he should have undertaken a_0 rather than a_1. All the previous results hold in this set-up. A somewhat more sophisticated deterrence policy for such a situation would be to allow individuals the opportunity to self-report their choice of action. In the optimal separating scheme, those who report a_1 are imposed a non-random sanction. Those who do not (and have in fact chosen a_0) face a random sanction $s(x)$ in terms of the likelihood of having undertaken a_1. Again, the results of the present paper hold with respect to the optimal penalty schedule $s(x)$ designed for those who should be deterred.

Appendix

Proof of proposition 1: Set $T = 0$ and ignore constraint (3) since it is irrelevant here. Omitting the argument x for brevity, the Lagrangean of the principal's problem is

$$\int_0^{\overline{x}} u(w-s) f_0\, dx + \lambda \int_0^{\overline{x}} u(w-s)(f_0 - f_1)\, dx \\ + \int_0^{\overline{x}} [\eta s + \xi(w - \underline{w} - s)] f_0\, dx. \tag{20}$$

where λ, $\eta(x)$ and $\xi(x)$ are nonnegative multipliers and where for all x

$$\eta(x)s(x) = \xi(x)[w - \underline{w} - s(x)] = 0. \tag{21}$$

Along the optimal path,

$$-u'(w - s(x))\varphi(x) + \eta(x) - \xi(x) = 0 \tag{22}$$

where

$$\varphi(x) = 1 + \lambda(1 - R(x)). \tag{23}$$

λ cannot be such that φ is everywhere positive, otherwise $\eta > 0$ and therefore $s = 0$ everywhere, which is inconsistent with (2). Since φ cannot be everywhere negative either, there exists x_c such that $\varphi(x_c) = 0$, which requires $R(x_c) > 1$ so that $x_c \in (x^*, \overline{x})$. Given MLRP, $\varphi' < 0$ and x_c is unique. The rest of the proof follows using (21), (22) and MLRP.

Proof of corollary 1: Allowing randomized penalties, let $\pi(x)$ be the conditional probability of sanction. Defining $\overline{s} = w - \underline{w}$, we minimize the probability of sanctions by solving

$$\min_{\substack{\pi(x)\in[0,1] \\ s(x)\in[0,\overline{s}]}} \int_0^{\overline{x}} \pi f_0 \, dx$$

subject to (2) now written as

$$\int_0^{\overline{x}} [\pi u(w - s) + (1 - \pi)u(w)](f_0 - f_1)\, dx \geq c. \tag{24}$$

Write $\lambda \geq 0$ for the multiplier of (24). At all x, π and s minimize

$$\psi = \pi f_0 - \lambda[\pi u(w - s) + (1 - \pi)u(w)](f_0 - f_1). \tag{25}$$

subject to the boundary constraints. Differentiating,

$$\frac{\partial \psi}{\partial s} = \lambda \pi u'(w - s)(1 - R) f_0, \tag{26}$$

$$\frac{\partial \psi}{\partial \pi} = \{1 + \lambda[u(w) - u(w - s)](1 - R)\} f_0. \tag{27}$$

Given that $R(x)$ is strictly increasing, $\pi(x) > 0$ in (26) leads to either $s(x) = 0$ or $s(x) = \overline{s}$ except possibly at an isolated point. Similarly, (27) implies that either $\pi(x) = 0$ or $\pi(x) = 1$. Combining the two results and using MLRP, the solution is easily seen to be as in proposition 1.

Proof of corollary 2: Let $\lambda \geq 0$ be the multiplier of the incentive compatibility constraint in the minimum expected sanction problem. Along the optimal path, the first and second order necessary conditions for an interior solution are

$$1 + \lambda[1 - R(x)]u'(w - s(x)) = 0, \tag{28}$$

$$-\lambda[1 - R(x)]u''(w - s(x)) \geq 0. \tag{29}$$

The condition (28) can hold only when $1 - R(x) < 0$, but then this contradicts (29). Hence, the solution is bang-bang and is easily seen to be as in proposition 1.

Proof of proposition 2: Compared with the proof of proposition 1, the Lagrangean is now augmented to

$$\int_0^{\overline{x}} u(w + T - s)f_0\,dx + \lambda \int_0^{\overline{x}} u(w + T - s)(f_0 - f_1)\,dx$$
$$+ \int_0^{\overline{x}} [\eta s + \xi(w - \underline{w} + T - s)]f_0\,dx + \mu[\int_0^{\overline{x}} h(s)f_0\,dx - T]. \tag{30}$$

where μ is a nonnegative multiplier. At all x,

$$\eta(x)s(x) = \xi(x)[w - \underline{w} - s(x)] = 0, \tag{31}$$

$$-u'(w - s(x) + T)\varphi(x) + \mu h'(s(x)) + \eta(x) - \xi(x) = 0, \tag{32}$$

where $\varphi(x)$ is defined as in (23). The first-order condition with respect to T is

$$\int_0^{\overline{x}} [u'(w - s + T)\varphi + \xi]f_0\,dx - \mu = 0. \tag{33}$$

To prove that the sanction is nondecreasing, let s_1 and s_2 be the optimal sanctions at x_1 and x_2, where $x_1 < x_2$. Because the optimal sanction maximizes $u\varphi + \mu h$,

$$u(w + T - s_2)\varphi(x_2) + \mu h(s_2) \geq u(w + T - s_1)\varphi(x_2) + \mu h(s_1),$$

$$u(w + T - s_1)\varphi(x_1) + \mu h(s_1) \geq u(w + T - s_2)\varphi(x_1) + \mu h(s_2),$$

leading to

$$[u(w + T - s_2) - u(w + T - s_1)][\varphi(x_2) - \varphi(x_1)] \geq 0. \tag{34}$$

Due to MLRP, $\varphi(x_2) < \varphi(x_1)$ so that (34) can hold only if $s_2 \geq s_1$.

To prove that the sanction is zero with positive probability, substitute from (32) in (33) so as to obtain

$$\int_0^{\overline{x}} \eta(x)f_0(x)\,dx = \mu \int_0^{\overline{x}} [1 - h'(s(x))]f_0(x)\,dx. \tag{35}$$

If $\mu = 0$, the solution is bang-bang as in proposition 1. But then the LHS of (35) is zero, which is incompatible with a bang-bang solution. Therefore, $\mu > 0$. Since

$h' < 1$, the LHS is then positive and this implies a positive probability of zero sanctions.

To obtain the expression for $s'(x)$ at interior solutions, set $\eta = \xi = 0$ in (32) and differentiate totally with respect to x so as to get

$$\frac{ds(x)}{dx} = \frac{u'\varphi'}{u''\varphi + \mu h''} > 0. \tag{36}$$

The sign follows from the second-order condition for a maximum and from $\varphi' < 0$. The expression in the proposition follows by substituting for μ from (32). From the latter it is also easily seen that $\varphi(x) = 1 + \lambda(1 - R(x)) > 0$ at interior solutions.

Finally, suppose that $s(x) < w - \underline{w} + T$ for all x. We show that this leads to a contradiction if $h' \leq m$ for m sufficiently small. Let $x' \in (x^*, \overline{x})$ be such that $s(x') > 0$. Such an x' necessarily exists, otherwise the sanction would be zero everywhere. From (32), given $\eta(x') = \xi(x') = 0$,

$$\varphi(x') = \frac{\mu h'(s(x'))}{u'(w + T - s(x'))} \leq \frac{\mu m}{u'(w + T - s(x'))} \tag{37}$$

From (33), given that by assumption $\xi = 0$ everywhere,

$$\mu = \int_0^{\overline{x}} u'(w + T - s(x))\varphi(x) f_0(x)\, dx \leq u'(\underline{w})\varphi(0). \tag{38}$$

where the inequality follows from $\varphi' < 0$ and $u'' < 0$. Furthermore

$$T = \int_0^{\overline{x}} h(s(x)) f_0(x)\, dx \leq h(w - \underline{w} + T) \leq m(w - \underline{w} + T). \tag{39}$$

so that

$$w + T \leq \frac{w - m\underline{w}}{1 - m} \tag{40}$$

and therefore

$$u'(w + T - s(x')) > u'\left(\frac{w - m\underline{w}}{1 - m}\right). \tag{41}$$

Combining (37), (38) and (41), we get

$$\varphi(x') < \varepsilon(m)\varphi(0) \tag{42}$$

or equivalently

$$1 + \lambda(1 - R(x')) < \varepsilon(m)[1 + \lambda(1 - R(0))], \tag{43}$$

where

$$\varepsilon(m) = \frac{mu'(\underline{w})}{u'\left(\frac{w-m\underline{w}}{1-m}\right)} \tag{44}$$

is strictly increasing, with $\varepsilon(0) = 0$. Now, for any $x'' \in (x', \overline{x})$,

$$\begin{aligned} \varphi(x'') &= 1 + \lambda(1 - R(x'')) \\ &< 1 + \left[\frac{\varepsilon(m) - 1}{1 - R(x') - \varepsilon(m)(1 - R(0))}\right](1 - R(x'')) \\ &= \frac{R(x'') - R(x') + \varepsilon(m)[R(0) - R(x'')]}{1 - R(x') - \varepsilon(m)(1 - R(0))} \end{aligned} \tag{45}$$

The inequality follows from (43), given that $R(x'') > R(x') > 1 > R(0)$. For m small, $\varepsilon(m)$ is small and the last expression in (45) is negative, implying $\varphi(x'') < 0$. But then $\varphi(x) < 0$ for all $x \geq x''$. Therefore, for all such x and all $s \in [0,\, w - \underline{w} + T]$ we have

$$-u'(w + T - s))\varphi(x) + \mu h'(s)) > 0. \tag{46}$$

This implies $s(x) = w - \underline{w} + T$, thus leading to a contradiction.

Proof of proposition 3: The optimization problem can be rewritten as

$$\max_{T,\widehat{s},s(x)\in[0,\widehat{s}]} T - \widehat{s}$$

subject to (2) and (3). The Lagrangean is

$$\begin{aligned} &T - \widehat{s} + \lambda \int_0^{\overline{x}} u(w + T - s)(f_0 - f_1)\, dx \\ &+ \int_0^{\overline{x}} [\eta s + \xi(\widehat{s} - s)] f_0\, dx + \mu[\int_0^{\overline{x}} h(s) f_0\, dx - T]. \end{aligned} \tag{47}$$

where λ, $\eta(x)$ and $\xi(x)$ are nonnegative. Along the optimal path,

$$\eta(x)s(x) = \xi(x)[\widehat{s} - s(x)] = 0, \tag{48}$$

$$-\lambda u'(w + T - s(x))(1 - R(x)) + \mu h'(s(x) + \eta(x) - \xi(x) = 0. \tag{49}$$

The first-order conditions with respect to T and $\widehat{s}$ are

$$1 + \lambda \int_0^{\overline{x}} u'(w + T - s)(f_0 - f_1)\, dx = 0, \tag{50}$$

$$\int_0^{\overline{x}} \xi f_0 \, dx - 1 = 0. \tag{51}$$

As in proposition 2, adding (50) and (51) and substituting from (49), one shows that $\mu > 0$ and that $s(x) = 0$ with positive probability. The expression for $s'(x)$ is also derived as in proposition 2. Finally, from (48) and (49), $s(x) < \widehat{s}$ is possible only if $R(x) < 1$ or equivalently $x < x^*$.

Proof of proposition 4: Solving the same problem as in proposition 3 but with (2) replaced by $T = 0$, along the optimal path

$$-\lambda u'(w - s(x))(1 - R(x)) + \eta(x) - \xi(x) = 0. \tag{52}$$

Given MLRP and the definition of x^*, the result is then straightforward.

References

Becker, G. S. (1968), "Crime and punishment: An economic approach", *Journal of Political Economy* 76, 169-217.

Becker, G. S. and G. J. Stigler (1974), "Law enforcement, malfeasance and compensation of enforcers", *Journal of Legal Studies* 3, 1-18.

Boadway, R., N. Marceau and M. Marchand (1996), "Time-consistent criminal sanctions", *Public Finance* 51, 149-165.

Bose, P. (1995), "Regulatory error, optimal fines and the level of compliance", *Journal of Public Economics* 56, 475-484.

Demougin, D. and C. Fluet (1998), "Monitoring vs incentives", *European Economic Review*, to appear.

Demougin, D. and C. Fluet (1999), "Ranking of information systems in agency models: An integral condition", WP no. 29, Otto-von-Guericke University.

Ehrlich, I. (1982), "The optimum enforcement of laws and the concept of justice: a positive analysis", *International Review of Law and Economics* 2, 3-27.

Garoupa, N. (1997), "The theory of optimal law enforcement", *Journal of Economic Surveys* 11, 267-295.

Huberman, G., Mayers, D. and C. W. Smith (1983, "Optimal insurance policy indemnity schedules", *Bell Journal of Economics* 14, 415-426.

Kaplow, L. (1990), "A note on the optimal use of nonmonetary sanctions", *Journal of Public Economics* 42, 245-247.

Kaplow, L. (1992), "The optimal probability and magnitude of fines for acts that are definitely undesirable", *International Review of Law and Economics* 12, 3-11.

Kaplow, L. and Shavell (1994), "Accuracy in the determination of liability", *Journal of Law and Economics* 37, 1-15.

Kim, S. K. (1995), "Limited liability and bonus contracts", *Journal of Economic Management and Strategy* 6, 899-913.

Malik, A. S. (1990), "Avoidance, screening and optimum enforcement", *Rand Journal of Economics* 21, 341-353.

Mirrlees, J. A. (1974), "Notes on welfare economics, information and uncertainty", in M. Balch, D. McFadden, and S. Wu, eds., *Essays in Economic Behavior Under Uncertainty*, North-Holland, Amsterdam.

Polinsky, A. M. and S. Shavell (1979), "The optimal tradeoff between the probability and magnitude of fines", *American Economic Review* 69, 880-891.

Polinsky, A. M. and S. Shavell (1992), "Enforcement costs and the optimal magnitude and probability of fines", *Journal of Law and Economics* 35, 133-148.

Raviv, A. (1979), "The design of an optimal insurance contract", *American Economic Review* 69, 84-96.

Shavell, S. (1987), "The optimal use of nonmonetary sanctions", *American Economic Review* 77, 584-592.

Tirole, J. (1992), "Collusion and the theory of organizations", in J. J. Laffont, ed., *Advances in Economic Theory*, vol. 2, Cambridge University Press, Cambridge.