

The Structure of Penalties in Environmental Enforcement: An Economic Analysis

KATHLEEN SEGERSON

*Department of Economics, Box U-63, Room 330, The University of Connecticut, Storrs,
Connecticut 06269*

AND

TOM TIETENBERG*

Department of Economics, Miller Library 234, Colby College, Waterville, Maine 04901

Received August 1, 1991; revised November 4, 1991

I. INTRODUCTION

The enforcement of environmental laws has undergone some rather dramatic changes over the last decade. Civil and criminal monetary penalties have been imposed more frequently on violators and the size of the penalties has increased.¹ Individuals, both corporate employees and officers, have become more common targets for penalties.² Finally, incarceration has not only become a more commonly imposed sanction,³ but the resulting jail sentences are longer.⁴

Despite this trend toward increased use of individual sanctions in the form of fines or incarceration, the efficiency implications of this trend have received little attention. The related work that has been published falls into two categories. The first considers the relative impacts or merits of holding firms liable for the actions of their employees, a principle known as "vicarious liability" [31, 32, 11, 29, 15]. Arguments in favor of vicarious liability presume that the firm can sanction employee malfeasance internally through compensation schemes that are tied to

*Tom Tietenberg's participation in this research project was funded by sabbatical research grants from Colby College and Resources for the Future, Inc., and a senior research fellowship from Woods Hole Oceanographic Institution. This financial support is gratefully acknowledged. The authors also thank Tom Miceli, Mark Cohen, and three anonymous reviewers for their contributions to the paper.

¹From its creation through fiscal year 1989 the United States Environmental Protection Agency has imposed \$185.9 million in civil penalties. In FY 1989, \$34.9 million in civil penalties were assessed, \$21.3 million in civil judicial penalties (the second highest in agency history) and \$13.6 million in administrative penalties (an all-time record). See [17].

²For an examination of how courts have begun to expand individual liability for corporate environmental violations, see [5].

³The number of criminal cases referred to the Department of Justice by EPA rose from 0 in fiscal year 1981 to 60 in fiscal year 1989 [17].

⁴For a discussion of this and related issues see [30].

workers' actions. However, in many cases involving environmental risks, internal sanctioning may not be possible. In addition, whenever those risks result from a combination of decisions made by both employees and executive officers or other managers of the firm, the appropriate allocation of liability among the responsible parties may be difficult to determine.

The second strand of related literature addresses the question of whether individual sanctions, if imposed, should take the form of fines or incarceration. Following the seminal article by Becker [1], this literature concludes that, if possible, fines should be used rather than incarceration, since the social costs associated with incarceration are so much higher. These increased costs include both the actual costs of incarceration (such as the costs of prisons and the opportunity cost of the individual's time) and increased transactions costs associated with the use of incarceration as a sanction.⁵ However, when an individual's assets limit the incentive effects of a fine or fines are too low (due, for example, to binding statutory limits), then incarceration may be a second best alternative. If incarceration is to be used as an alternative to fines when fines provide insufficient incentives, what form should the incarceration rule take? Should all individual actions that result in environmental damages be subject to incarceration or should the incarceration option be reserved for certain types of actions?

In this paper we consider the design of efficient sanctions in the context of actions by employees and firms that lead to environmental risks. We address first the choice between imposing fines on the firm itself or imposing them directly on the employees whose decisions can affect those risks. Having derived conditions under which these two distinct approaches will yield the same outcome, we then consider the conditions under which one approach or the other may dominate or when some combination of the two should be used. We then turn to the question of whether incarceration is a possible instrument for inducing efficient behavior when fines provide inadequate incentives, and, if so, what form the incarceration rule should take.

In our examination of both fines and incarceration, we focus on the contractual relationship between an employer (the firm) and its employee (e.g., an environmental risk manager) and the role that this relationship plays in determining the design of efficient sanctions. In particular, we use a principal-agent model to examine the effect of corporate and individual sanctions on compliance incentives in the environmental context, including for individuals the effects of both monetary sanctions and incarceration. Our conclusions suggest that, while the traditional use of corporate penalties for enforcement is still justified in many cases, there are circumstances under which the use of individual fines or incarceration may improve efficiency.

⁵For example, incarceration is normally possible only with a criminal conviction, and criminal convictions require a more stringent standard of proof than do civil convictions [16]. In some cases, however, environmental enforcers have discovered ways of lowering this prosecutorial burden. For example, the Chicago office of the EPA has moved toward using consent decrees signed between a Federal judge and the polluter (typically local governments in this instance). The effect of such action is that after the consent decree is in force non-compliance is considered a violation of the consent decree which results in a contempt of court citation, substantially increasing the likelihood of incarceration without increasing the burden of proof. We are indebted to an anonymous referee for pointing this out to us.

II. THE DESIGN OF EFFICIENT FINES

The efficiency of alternative liability rules designed to induce individuals to take efficient precautions in preventing injuries to others is the subject of a large and growing body of literature.⁶ In cases of “unilateral” care, where the injured party (victim) is unable to affect the probability or magnitude of damages, this literature suggests that the injurer will face efficient incentives when subject to a strict liability rule which holds him accountable for any damages resulting from his actions.⁷ The same efficient incentives would be created by an equivalent fine that is paid to a third party (such as the government) rather than to the victim.⁸ These results suggest that, in order to ensure efficient incentives, liability or a fine should be imposed on each party in an amount equal to the damages that result from its actions. If the actions of more than one party combine to determine damages, then each party should be liable for its incremental contribution to damages [33, 14].

Environmental risks are often the product of combined actions of multiple parties. For example, an environmental risk manager makes day-to-day decisions about the operations of plant and equipment, while higher level managers or corporate executives make overriding safety-related policy decisions, as well as decisions affecting the scale of operations. Both types of decisions can affect the probability and/or magnitude of environmental damages. Thus, environmental risks are often the result of multiple decisions by multiple parties.

The standard results (discussed above) that suggest that each party should be liable for his/her incremental contribution do not necessarily hold in this context. Unlike standard “accidents” where the parties involved are unrelated, here the firm and its employees (in particular the environmental risk manager) have a contractual relationship. The wage contract provides a potential mechanism through which one party can influence the incentives faced by the other party.

In essence the incentive problem associated with managing environmental risk can be characterized as a type of principal-agent problem.⁹ Different remedies trigger different consequences for the firm (principal), and as a result of both the nature of the remedy and the response of upper-level managers, different conse-

⁶See [29] for an excellent overview.

⁷Efficiency can also result from the use of a negligence standard, under which the injurer is liable only if he failed to exercise “due care.” See [29] for a discussion of this result and [18] for conditions under which it would not hold.

⁸Of course the goal of victim compensation is not ensured if fines are used instead of direct liability to the injured party. However, fines are often used in cases where the victim is not a single identifiable individual (such as an accident victim) but rather a group of (possibly unidentified) individuals (such as “society at large”). This is particularly likely in the context of environmental damages, where the specific individuals suffering from environmental damage (e.g., decreased air quality) are not easily identified. Fines can also dominate direct liability in cases of bilateral care where victim compensation can reduce a potential victim’s incentives to avoid damages.

⁹The safety impacts of the agency relationship between firms and employees have been examined by others, including [11, 31, 32, 15]. Both [11] and [15] present formal (mathematical) principal-agent models. However, they both focus on the case where contingent contracts are possible, i.e., where the wage depends on whether or not an accident occurs. In addition, in [15] expected damages depend only on the actions of the worker. While [11] considers the case where the employer’s actions can also affect damages, his use of a contingent contract implies that wages can be made to depend on the employer’s choice of care. We consider situations in which this is not possible. Our analysis also differs in its focus on efficiency analysis.

quences for the environmental manager (agent). Under what conditions can each of the remedies achieve efficiency? Does any one remedy dominate?

Characterizing an Efficient Outcome under a System of Fines

Imagine that an environmental manager (worker) is responsible for some particular risk (such as waste disposal). Assume for simplicity that the probability of an accident is fixed, but if an accident occurs, the level of damages depends on the amount of "care" exercised by the worker and the decisions (safety expenditures and/or output level) of the firm.¹⁰ All parties are assumed to be risk neutral¹¹ and all costs (including the various psychic costs associated with incarceration) can be converted into monetary equivalents.¹² Finally, we assume that enforcement through fines (although not through incarceration) is costless.¹³

These assumptions and their implications can be formalized as follows:

Let a = the actions (level of care) taken by the employee,

x = the firm's production level,¹⁴

$B(x)$ = benefits (revenue) from production, $B'(x) > 0$,

$C(x)$ = the total cost of production,¹⁵ $C'(x) > 0$,

W = direct payment from the firm to the worker (wage or salary),

$D(x, a)$ = environmental damage if an accident occurs,¹⁶ $D_x > 0$, $D_a < 0$,

¹⁰In reality, both the probability of an accident and the level of damages could depend on care and output. It would be conceptually easy, but expositionally cumbersome, to include both effects in the model. Since our results do not depend on this distinction, we opt here for simplicity. Note that simplicity could also have been attained by assuming that only the probability of an accident varies with behavior, i.e., that damages are fixed.

¹¹For related models that incorporate risk aversion, see [19] and [15]. In general penalty structures will affect not only the incentives faced by two parties, but also the allocation of risk between them. When one of the parties is risk averse, then efficient risk-sharing implies that all of the risk should be placed on the risk-neutral party. While we do not consider risk-aversion (and thus risk-sharing) here, some of its implications for our analysis are immediate. For example, even when the incentive effects of corporate and individual fines are identical (see below), they would not necessarily be perfect substitutes. In particular, if firms are risk-neutral and workers are risk-averse, corporate fines would be preferred because they yield a more efficient allocation of risk. Likewise, when the incentive effects of the two types of fines are not identical, a trade-off between efficient incentives and efficient risk-sharing might exist. Such a trade-off would affect the design of an efficient policy.

¹²This is a standard assumption in economic models of agency theory (e.g. [8, 24, 15]) and crime (e.g. [20, 28]).

¹³In particular, in order to focus on the incentives of the firm and the worker, we ignore the cost of detecting violations and imposing fines, thereby implicitly assuming that the probability of detection is one. (The implications of relaxing this assumption are discussed below.) Thus, we abstract from issues relating to the optimal tradeoff between the probability of detection and the magnitude of the fine that arise in the context of costly enforcement. See Polinsky and Shavell [19] for a discussion of these issues.

¹⁴Alternatively, x could represent the firm's investment in safety (e.g., expenditures on safety equipment, training, etc.). Our results would apply under this interpretation of x as well. (Note that in this case, $B(x) = 0$, $C(x) = x$, $D_x \leq 0$ (see definitions below) and (1) reduces to a cost-minimization problem.)

¹⁵For simplicity we assume that production costs do not depend on care, an assumption that is also made by [15]. Our results would also apply, however, to the case where C depends on a , i.e., $C = C(x, a)$.

¹⁶We measure these damages net of any private internal damages that might result from an accident. Thus, D should be interpreted as third party damages. In addition, while we use the term

p = probability of an accident,

$V(a)$ = the dollar equivalent to the employee of the disutility of exercising care,¹⁷ $V'(a) > 0$, and

U^0 = the employee's opportunity cost of working for the firm.

(Subscripts on functions denote partial derivatives.) Note that we allow damages to depend on both a and x , since, for example, the size of the potential injury caused by an operation could depend on both the size of the operation (x) and how carefully it is conducted (a).

The socially efficient choice involves selecting those values of x and a (denoted x^* and a^*) that maximize¹⁸

$$B(x) - C(x) - V(a) - pD(x, a). \quad (1)$$

The corresponding first order conditions are

$$B'(x^*) - C'(x^*) - pD_x(x^*, a^*) = 0 \quad (2)$$

$$V'(a^*) + pD_a(x^*, a^*) = 0. \quad (3)$$

These conditions imply that the production level should be increased until the marginal benefit equals the marginal cost (including any increase in expected damage due to the expanded production) and the level of precautionary actions should be increased until the marginal cost of additional precaution equals the marginal reduction in expected damage.¹⁹

Our interest is in determining whether a system of penalties can be designed to ensure that the private choice of x and a are socially efficient. We focus initially on the case where the penalties take the form of fines (rather than incarceration) and consider two types of fines: (i) a fine, F , that could be levied against the corporation in the event of an accident, and (ii) a fine, G , that could be levied against the environmental risk manager. Since in practice fines on firms and individuals can be used individually or in combination, we allow for both types of fines and let the model determine the efficient combination of the two. In addition we allow the magnitudes of the fines to vary with damages, i.e., $F = F(D)$ and $G = G(D)$, which assumes that damages can be measured *ex post* with certainty.

"accident," this should be broadly interpreted to include both sudden releases or exposure where the damages are immediately apparent and gradual or long term releases or exposure where damages are not apparent for some time.

¹⁷This assumes that the worker's utility function is separable in money and care, an assumption that is standard in agency models [e.g. 8, 24, 15].

¹⁸In the literature on the economics of crime, there has been a debate about whether the benefits of a criminal activity to the criminal himself should be included in the measure of social benefits. A useful survey of the issues is provided in Lewin and Trumbull [13]. They argue that the benefits of truly criminal activities (such as murder, rape and theft) should not be included while the benefits of activities that more closely resemble civil violations (such as regulatory offenses) should be included. In our case, since the purpose of the worker's actions is not to inflict harm and thus we would not generally view the worker's disutility from undertaking care as "reprehensible," it seems appropriate to include it in the measure of social benefits.

¹⁹If x is interpreted as the firm's investment in safety (see Footnote 14), then this latter condition applies to x as well.

Allowing the fines to vary with D does not, however, rule out *a priori* the possibility that the efficient fine will be unrelated to damages, i.e., that under the efficient set of fines $F'(D)$ and/or $G'(D)$ will be zero. Thus, whether the efficient fine is related to damages or not will be determined by the model.

Consistent with previous models of optimal fines e.g., [19, 20], our model does not explicitly distinguish between criminal and civil penalties *per se*. There are two ways that one might try to distinguish between the two types of penalties in a formal economic model.²⁰ First, one could easily distinguish between the probability of conviction under the two types of penalties, reflecting the higher standard of proof required for criminal fines. However, since this would complicate the model without adding additional insight, we have chosen to abstract from the enforcement/conviction aspect of the problem and assume that the probability of conviction is one. (The implications of relaxing this are noted below.) Secondly, one could try to distinguish between criminal and civil penalties by equating them with intentional and unintentional violations, where an intentional violation would correspond to “knowingly” inflicting damages and thus be subject to criminal fines. Landes and Posner [12, Chap. 6] develop an economic model of “intentional” torts, which they define as torts that yield benefits for the injurer only when damages are inflicted on the victim (the injurer deliberately inflicts a wrongful injury). This approach seems inappropriate in our context, since, even if the worker knew that his actions were endangering someone, the main purpose of those actions is not to inflict injury.²¹

Since damages depend on the actions of two parties, the worker (who chooses a) and the firm (or its production manager, who chooses x), our problem is similar to the case of “joint torts” considered in [33] and [14]. However, unlike the joint tort cases in which the actions of two independent parties combine to cause damages, the parties whose actions jointly affect D in our analysis are governed by contractual relationships. As noted by [25] and [22], the existence of a contractual relationship provides a mechanism by which one party can shift some of its liability or costs onto the other party to the contract. However, in order to shift costs efficiently, the relationship between the two parties must provide a mechanism for fully internalizing any costs that one party’s actions can impose on the other party. The existence of such a mechanism depends on the ability to observe or monitor certain actions.

In our context, the firm and the employee have a contractual relationship through the employment contract, and the wage rate provides a possible mechanism for shifting costs forward (from worker to firm) or backwards (from firm to worker). For simplicity, we assume that the employee has a fixed opportunity cost

²⁰To the extent that the distinction between the two rests primarily on moral bases (see, for example, [13]), then little can be done in a formal economic model to capture this distinction.

²¹Thus, his actions would fall instead under Landes and Posner’s [12] category 1 (undertaking an activity with a high probability of causing an injury but not desiring or benefiting from the occurrence of the injury), which they argue is an inappropriate economic definition of “intentional” tort. If intentional is simply defined as knowing that an injury is possible (or even likely), then the actions in our model are all intentional since we assume that all parties make choices voluntarily with full information about all of the relevant functions and parameters. Under this definition, modeling unintentional behavior would require incorporating imperfect information about the relevant relationships, which would cause the perceived distribution of outcomes to differ from the actual distribution. We are unaware of any previous models of criminal activity that have explicitly distinguished between intentional and unintentional effects in this way.

of working for the firm (denoted U^0) and that the firm captures all of the potential rent.²² Thus, the firm chooses the wage (or, equivalently, salary) of the employee, subject to the constraint that the employee's monetary return from this employment just equals his opportunity cost.²³ In the following sections we ask whether corporate or individual penalties can be designed to ensure that efficient levels of x and a are chosen.

A First-Best Solution with F and G Perfectly Substitutable: $W = W(a, x)$

Consider first the case where the actions of the employee are observable or can be monitored with perfect certainty, so that wages can vary with the choice of a . Assume, furthermore, that wages can be made contingent on the performance of the firm ($B(x) - C(x)$), and thereby contingent on x . Thus, the employment contract takes the form $W = W(a, x)$. Note that, while the level of the wage depends on the levels of a and x , it does not depend on whether an accident occurs. Thus, it is not a contingent contract in the sense considered by Newman and Wright [15]. Faced with a fine F for any accident, the firm chooses x and W to maximize its net benefits:

$$B(x) - C(x) - W(a, x) - pF(D(x, a)) \quad (4)$$

subject to a being chosen by the employee to maximize

$$W(a, x) - V(a) - pG(D(x, a)), \quad \text{and} \quad (5)$$

$$W(a, x) - V(a) - pG(D(x, a)) = U^0. \quad (6)$$

The constraint in (6) ensures that the worker's expected return from the job, given his choice of a and his wage contract, just equals his opportunity cost.²⁴

Consider first the firm's choice of W . Clearly, for any level of x , the firm wants to minimize $W(a, x) + pF(D(x, a))$, i.e., it wants the worker to choose a level of a that minimizes the wages it pays plus its expected fine. Adding $pF(D(x, a))$ to both sides of (6) and rearranging terms yields²⁵

$$W(a, x) + pF(D(x, a)) = U^0 + V(a) + pG(D(x, a)) + pF(D(x, a)). \quad (7)$$

Minimizing this expression yields the first-order condition for the level of a that minimizes the firm's costs, i.e., the level that the firm would like to see the

²²This is a standard assumption in agency models (e.g. [8, 24, 15]). The implications of relaxing it are noted below.

²³Since we are interested in incentives regarding environmental risks rather than labor supply decisions, we abstract from the employee's choice of the number of hours worked and instead assume that, if he works for the firm, he works for a fixed number of hours.

²⁴Note that, if all of the rent from the wage contract is not captured by the firm, then (6) would be specified as an inequality rather than an equality constraint. Furthermore, this specification of the firm's choice problem presumes that there exists a combination of x , a and $W(a, x)$ satisfying (5) and (6) that yields net benefits in excess of $\text{Max}\{B(x) - C(x) - pF(D(x, 0))\}$, i.e., that it is "worth it" for the firm to hire the environmental risk manager provided a is chosen to satisfy (5).

²⁵Note that such a substitution would not be possible if the worker were risk-averse, since (6) would depend on expected utility rather than expected wages, as in [15]. In the case of risk-neutrality, however, (6) implies that the firm is actually indifferent to the actual wage contract as long as it induces the level of a that minimizes overall costs for the firm.

employee choose.²⁶

$$V'(a) + p\{G'(D)D_a(x, a) + F'(D)D_a(x, a)\} = 0. \quad (8)$$

Given $W(a, x)$, the employee will in fact choose a to maximize his return from employment, $W(a, x) - V(a) - pG(D(x, a))$. The corresponding first-order condition is

$$W_a(a, x) - V'(a) - p\{G'(D)D_a(x, a)\} = 0. \quad (9)$$

From (8) and (9), it is clear that, if the firm wants the employee to choose the level of a that minimizes (7), it can choose the wage function such that

$$W_a(a, x) = -pF'(D)D_a(x, a).$$

This implies that the wage function takes the form

$$W(a, x) = k(x) - pF(D(x, a)), \quad (10)$$

where $k(x)$ is independent of the worker's choice of a .²⁷ Faced with this wage function the worker's choice of a to maximize his return would simultaneously minimize the firm's costs for any given x . Furthermore, the equilibrium level of k can be determined from (6) and (10), which imply that

$$k(x) = U^0 + V(a_e) + p(G(D(x, a_e)) + F(D(x, a_e))), \quad (11)$$

where a_e is the equilibrium choice of a .

Consider next the firm's choice of x . Given (10) and (11), the firm's choice problem can be rewritten as:

$$\text{maximize } B(x) - C(x) - U^0 - V(a_e) - p(G(D(x, a_e)) + F(D(x, a_e))) \quad (12)$$

subject to (5) and (6).²⁸ The corresponding first-order condition for x is

$$B'(x) - C'(x) - p[G'(D)D_x(x, a_e) + F'(D)D_x(x, a_e)] = 0. \quad (13)$$

Note that, in choosing x , the firm considers the impact of that choice on the wages that it must pay. In other words, higher levels of x *ceteris paribus* imply higher levels of damages, which in turn imply higher possible fines on both the firm and the employee. As compensation for his higher fines, the employee requires a

²⁶Here and throughout the paper we assume that the solutions to first-order conditions are unique and that the corresponding sufficient second-order conditions are met. For a discussion of how to formulate a principal-agent problem without these assumptions, see [6].

²⁷This wage function is not unique. For example, if utility were normalized so that $V(a) = a$ and $G = 0$, then the firm can induce any given level of a (say a') through a wage function of the form $W(a, x) = k(x) + pF(D(a, x))/F'(D)D_a(a', x)$. It could thus induce its most preferred level by setting a' equal to that level. This is similar to the wage function used in [15].

²⁸Note that, given (10) and (11), the firm's objective function is independent of the actual choice of a .

higher wage to induce him to work for the firm. Thus, through the wage contract, fines imposed on the worker can affect the net return of the firm.

Likewise, the wage contract provides a mechanism for shifting the firm's fines onto the worker. Given (10), the first-order condition for (5) becomes

$$V'(a) + p[G'(D)D_a(x, a) + F'(D)D_a(x, a)] = 0. \quad (14)$$

Thus, the worker considers the effect of his choice of a on the firm's expected fine. The higher the expected fine the firm has to pay, the lower the wage it will pay to its worker.

Having defined the conditions for the equilibrium choices of x and a , we can now determine whether corporate or individual fines can ensure that those choices are efficient. Comparing (13) and (14) to (2) and (3), it is clear that the private outcome will be socially efficient if

$$F = D(x, a) \quad \text{and} \quad G = 0 \quad \text{or} \quad (15)$$

$$F = 0 \quad \text{and} \quad G = D(x, a). \quad (16)$$

This result suggests that efficiency can be achieved with either (i) a monetary penalty placed on the corporation alone, without any penalty on the worker, or (ii) a penalty on the responsible individual, without any corresponding penalty on the firm.²⁹ In other words, with $W = W(a, x)$ given by (10) and (11), F and G (appropriately set) are perfect substitutes.

The efficiency of using F alone is consistent with the standard results of principal-agent models, e.g. [24, 8], where under risk neutrality efficient incentives result from a payoff scheme that transfers the full payoff of the principal to the agent, plus or minus a constant take for the principal (as in (10)).

The second case (a penalty on the worker only) has no analogy in the standard principal-agent models, since those models assume that the payoff/penalty is only realized directly by the principal. Here, the ability to penalize the worker directly through individual fines provides an extra mechanism for inducing efficient incentives. It is interesting to note that the use of individual penalties alone leads not only to the efficient level of a , but also to the efficient level of x . Forward shifting of the costs borne by the worker through demands for higher wages provides the link. Since increases in x lead to higher expected damages, which in turn imply higher penalties for the worker, the worker will demand (through (6)) a compensating increase in wages. Thus, the worker is able to shift his penalty onto the firm, thereby internalizing the effect of x on damages and inducing the firm to choose an efficient level of x .³⁰

²⁹Other combinations could lead to efficiency as well. In particular, any fine structure of the form $F(D) = sD + k_F$ and $G(D) = (1 - s)D + k_G$ would be efficient as well, where $0 \leq s \leq 1$ and k_F and k_G are constants. With $0 < s < 1$, a total fine equal to damages would be apportioned between the worker and the firm. With k_F and $k_G > 0$, each would in addition pay a fixed penalty every time an accident occurred. Note, however, that such "lump sum" penalties would have no effect on our results. The lump-sum component for the wage rate (as determined by (6)) would simply adjust to reflect $k_G > 0$. In addition, $k_F > 0$ would not affect the firm's marginal incentives regarding the choice of x . While it could affect a firm's entry/exit decision (i.e., the decision about whether to produce at all), such non-marginal incentives are not considered here. (See [14]). For these reasons, we assume here and throughout the paper that arbitrary lump-sum fines are not used (i.e., that $k_F = k_G = 0$).

³⁰In standard principal-agent models, the principal cannot influence the outcome directly, and thus there is no need to ensure efficient incentives for the principal.

Finally, the above result is consistent with previous work showing that, when two parties have a contractual relationship, then the efficient outcome can be achieved by placing liability on either of the two, since through forward or backward shifting one party can always transfer costs to the other e.g. [25, 22]. Note, though, that it would not be efficient to fine *both* the worker and the firm for the full amount of the damages. Unlike the cases of joint torts considered in [33] and [14], here requiring both parties to be fully liable for damages (even at the margin) would lead to excessive levels of precaution.

While we have developed the substitutability between corporate and individual penalties under the assumption that the worker's actions are observable (so that wages can vary with a), it should be clear from (10) that the results would continue to hold if the firm instead based wages directly on the extent of the fines that were actually incurred by the firm. For example, the firm could base much of the worker's salary on end-of-year bonuses that were dependent on the firm's performance and its safety or environmental record. In this case, (10) becomes

$$W(x, F) = k(x) - pF. \quad (10)'$$

Clearly, either (15) or (16) yield efficiency in this case as well.³¹

While the case where wages can vary directly with both x and a may be very optimistic (in terms of the amount of flexibility available), considering it performs two functions. First it suggests that when the necessary conditions are present, either corporate or individual monetary penalties are sufficient to produce the appropriate incentives. This is a powerful conclusion which supports both the traditional economic model of corporate harm and a good deal of current practice.³²

Perhaps more importantly, however, the model allows us to isolate the conditions that would ensure that corporate and individual penalties are perfect substitutes. In particular, perfect substitutability is ensured if: (1) in the absence of contingent contracts, the actions of the employee are observable so that wage structures can be used to penalize inefficient behavior and reward efficient behavior;³³ (2) wages can vary with the performance of the firm so that firms will be induced to make efficient decisions; (3) principals have the ability to manipulate wages sufficiently that the proper correspondence with actions can be established;³⁴ (4) the penalty or fine is set at the correct level; and (5) both the corporation and the worker have sufficient assets to be able to pay the penalties that are imposed on them.

Violating any one of these conditions can destroy the perfect substitutability of individual and corporate sanctions. We believe that a violation of each of these circumstances is present to some degree in various types of environmental problems. To trace out the implications we consider the effect of relaxing each of these conditions.

³¹In fact, this is similar to the contingent contract that would be implied by the standard principal-agent model under risk neutrality, where outputs are observable but inputs are not. In the standard model, however, since payoffs do not depend on x , k would not have to depend on the decisions of the firm.

³²For firms receiving criminal fines between 1984 and September 1990 for environmental crimes, slightly less than one half involved penalties levied only against the organization. See [4].

³³For evidence that compensation systems can be (and have been) devised to encourage managers to act in accordance with the long range interests of their employer, see [7].

³⁴Corporations are clearly attempting to do this. See the discussion of DuPont's strategy in [9].

Imperfect shifting. The above model assumes that costs can be shifted perfectly between the worker and the firm through the wage contract. Perfect shifting requires that (i) wages can be based on inputs or outputs (a or F) and the firm's decisions (x), and (ii) the wage is determined by the opportunity cost of the worker's employment with the firm (Eq. (6)). However, in many situations posing environmental risk these conditions are unlikely to hold. For example, in many cases it will not be possible for upper-level managers (those who determine wages) to monitor continually the day-to-day operations of the firm; workers whose actions can affect environmental risks can engage in behavior that avoids detection. Likewise, it will very often not be possible to base wages on the environmental performance of the firm. For many environmental risks the consequences of employee actions may not become apparent until some time in the future (if ever), possibly after the worker has left the firm. Thus, even if a can be inferred *ex post* (from the observed values of D and x), it may not be possible to adjust the worker's wage accordingly. In addition, it may not always be possible to trace the cause of an accident (such as a spill) back to the actions of a specific employee.³⁵ Alternatively wages may be governed by contracts such as union or civil service contracts. For all of these reasons firms may not be able to vary wages either directly or indirectly with the actions of the employee or the size and frequency of the penalties it faces. Thus, backward shifting may not always be possible.

The possibilities for forward shifting may similarly be limited. As long as wages are set by contract, at least in the short run workers will not be able to secure increased compensation for firm-level decisions that expose them to higher penalties.³⁶ Likewise, workers may lack the individual bargaining power necessary to secure demands for compensating wage increases, due, for example, to limited mobility. In such cases, they will not be able to shift the incidence of the penalties forward to the firm. In this subsection, we consider the implications of imperfect shifting for the design of efficient fines.

Suppose first that wages cannot be based upon either the worker's input (a) or the outcome (F), i.e., W is independent of both a and F . It is thus determined solely by (6)

$$W(x) = k(x) = U^0 - V(a_e) - pG(D(x, a_e)), \quad (17)$$

where a_e is the new equilibrium level of a that maximizes the worker's full return from employment. Given that the worker views W as constant, he chooses a to satisfy

$$V'(a) + p[G'(D)D_a(x, a)] = 0. \quad (18)$$

Given the worker's choice of a , the firm in turn chooses the level of x (denoted x_e) to

$$\text{maximize } B(x) - C(x) - U^0 - V(a_e) - p[G(D(x, a_e)) + F(D(x, a_e))] \quad (19)$$

³⁵In some cases it may not even be possible to trace the cause to a particular firm. See [26] for a discussion of issues related to uncertainty over causation.

³⁶In the long run, as contracts are renegotiated, the exposure of individuals to penalties may be a bargaining issue that affects subsequent wages.

or, equivalently the level of x that satisfies

$$B'(x) - C'(x) - p[G'(D)D_x(x, a_e) + F'(D)D_x(x, a_e)] = 0. \quad (20)$$

Comparing (18) and (20) to (2) and (3), it is clear that the use of individual penalties alone in this case would lead to efficiency (i.e., $a_e = a^*$ and $x_e = x^*$), while penalties on the corporation alone would not. In other words, $F = 0$ and $G = D$ yields an efficient outcome, while $F = D$ and $G = 0$ does not. Imposing a penalty on the individual induces the worker to take an efficient amount of care. In addition, since the worker's wage varies with x , through forward shifting he can induce the firm to choose an optimal level of x . Thus, efficiency of both a and x are ensured with individual penalties. The same is not true, however, with corporate penalties. Since the wage is independent of the worker's actual actions, the firm has no means of shifting the cost of penalties back to the worker. While corporate penalties can be used to induce an efficient choice of x (given a), they will not induce the worker to choose a efficiently. Thus, when wages cannot be based on the actions of the worker or the resulting fines, then the use of individual penalties is preferred to the use of corporate penalties.

Efficiency of individual penalties in the case where a is not observable requires that the wage in (17) be allowed to vary with the firm's choice of x , i.e., that the worker be able to demand higher wages from firms with more output (or lower safety expenditures) and therefore higher expected environmental damages. It should be clear that if wages were allowed to vary with a but not with x , then the opposite results would be obtained. In other words, with a wage function $W(a)$ that depends on a but is independent of x , corporate penalties lead to efficiency while individual penalties do not. Here, basing wages on the worker's actions allows the firm to shift its expected penalties back to the worker so that the effects of the worker's actions will be internalized. However, if wages are independent of x , the worker has no means of shifting the marginal external effects of x back to the firm. Thus, the wage contract can be a substitute for direct penalties on the worker but not for direct penalties on the firm.

Consider finally the case where the wage is independent of both a and x , i.e.,

$$W = U^0 + V(a_e) + pG(D(x_e, a_e)), \quad (21)$$

where a_e and x_e are the equilibrium choices of a and x under a fixed wage. Since the worker views W as fixed, its choice of a (a_e) satisfies

$$V'(a) + [pG'(D)D_a(x, a)] = 0. \quad (22)$$

Likewise, the firm's choice of x (x_e) satisfies

$$B'(x) - C'(x) - pF'(D)D_x(x, a) = 0. \quad (23)$$

Comparing (22) and (23) to (3), it is clear that with a fixed wage efficiency (i.e., $a_e = a^*$ and $x_e = x^*$) requires that *both* corporate and individual penalties be

used simultaneously. Neither can achieve an efficient choice of both x and a by itself. With W independent of both x and a , neither forward nor backward shifting will internalize costs. Thus, the contractual relationship between the worker and the firm is unrelated to the externality and the problem reduces to the standard case of a joint tort where no contractual relationship exists. For such cases, efficiency requires that each party must bear its costs directly at the margin, e.g. [33, 14]. In our context, this implies that corporate and individual penalties should be used *simultaneously*.³⁷ Note that this result would continue to hold if wages were determined by bargaining rather than by (6) (so that workers and employers share the employment rent), as long as the bargained wage is independent of the actual choices of x and a .

Other limitations. To achieve efficiency with the use of individual or corporate penalties alone, our model suggests that the penalty (either F or G) should be set equal to the actual level of damages. This result assumed that enforcement is perfect, i.e., that the probability of detecting damages and imposing the associated fine is one. In practice, damages can go undetected and, even when damages are detected, fines may not always be imposed. In this case, the expected fine will equal pqF (or pqG), where q is the conditional probability that the fine will be imposed, conditional on damages having occurred.³⁸ Clearly, with $q < 1$, the efficient fine will equal D/q , which exceeds D . In addition, the lower the probability of enforcement, the higher the efficient fine will be.

Cohen [3] compared the level of monetary sanctions with the amount of harm caused and found that the sanctions were on average about equal to the harm for large harms and greater than the harm for small harms.³⁹ This implies that the monetary sanctions are too low to ensure efficient incentives unless the probability of a sanction being imposed for every violation is 1.0 or private settlements, reputational losses and other associated costs are sufficiently large to offset the difference.⁴⁰ Inefficiently low penalties can result either from penalty determination procedures (a reluctance of courts to impose penalties that exceed damages and could trigger bankruptcy) or from statutory limits.⁴¹ In addition, problems in measuring the full amount of social damages and assigning the liability to responsible parties can lead to inefficiently low penalties as well.

A second limitation to the use of either individual or corporate fines is the problem of insufficient assets to pay the imposed fines. Though long recognized as a source of inefficiency in tort law remedies, e.g. [27, 28], insufficient assets can

³⁷In [14] it is demonstrated that to achieve efficient non-marginal (entry/exit) incentives, each party's liability must be adjusted by a constant to ensure that each pays only his incremental contribution to damages. Here, such an adjustment could be achieved through a change in the fixed (constant) wage.

³⁸Note that q could depend on the level of damages, with an increase in the level of damages leading to an increase in the probability of detection.

³⁹In [3] total sanctions are defined as including all government imposed sanctions including federal fines, restitution, administrative penalties, state fines, voluntary restitution, and court-ordered payments to victims. This definition does not include nonmonetary sanctions, losses to reputation or private settlements.

⁴⁰In general the total sanction should be a multiple of the harm where the multiple is equal to the inverse of the likelihood that a violation will result in a monetary sanction. See [19].

⁴¹See [23] for a more detailed discussion of this point.

undermine the efficiency of the use of financial penalties in enforcement as well. Furthermore, introducing individual penalties provides a new source of concern about the problem of insufficient assets since individuals typically have fewer assets than the organizations that employ them. Insufficient assets change the nature of taking environmental risks by curtailing the size of the penalty that would have to be paid.⁴² When firms have insufficient assets to pay the fine, their downside risk is truncated.⁴³ Thus, even if the courts were routinely imposing an efficient level of monetary sanctions, the polluter with insufficient assets to pay the mandated fines would face insufficient incentives.⁴⁴

III. THE CRIMINAL SANCTIONS ALTERNATIVE

When a firm's or an individual's assets are insufficient to cover the efficient penalties that would be levied on them or when the actual penalties are below the efficient levels, then incarceration might be used either as a substitute penalty or as an addition to the monetary penalty to promote efficiency.⁴⁵ With the use of incarceration, two extra social costs are introduced: (i) the dollar equivalent of the direct disutility (or lost income) from a jail sentence for the worker (J), which is a monotonically increasing function of the number of years spent in jail, and (ii) the social costs of incarceration (S), including the costs of providing capital (prisons), labor (guards and support personnel) and raw materials and services (uniforms), which also increase with the length of the sentence. Neither of these costs exist with a fine. Therefore as long as any specific level of deterrence achieved by incarceration could also be achieved by a fine, fines dominate [20]. However, in a second-best world (where efficient fines cannot be levied for one reason or another) incarceration may have some role to play. We consider here the limiting case where fines cannot be used at all and therefore the only instrument available is incarceration. While we focus on incarceration of the worker only, the corresponding results for incarceration of corporate executives or managers (i.e., those individuals who chose x) should be apparent.

In previous economic analyses of incarceration [20, 28], the "criminal" activity is modeled as a discrete choice; the individual either engages in the activity or does not. In our context, however, the activity (care) is a continuous choice. Having a continuous choice variable forces us to confront the issue of how to define the

⁴²In some ways the problem of insufficient assets might be even more severe in enforcement than in tort law. Most tort law damages paid to individuals and costs incurred in cleaning up sites are insurable, providing another source of assets to be tapped. Fines, however, are typically not insurable, particularly criminal fines which fulfill the *scienter* requirement.

⁴³This could create an incentive to set up underfunded subsidiaries for the sole purpose of keeping environmental risks away from the parent corporation. This has apparently happened in the United States according to evidence presented in [21].

⁴⁴This problem could be reduced somewhat by using a "negligence" approach to fines, i.e., imposing the fine only if the party failed to comply with a due standard of conduct. See [2] for a discussion of this approach.

⁴⁵An alternative to incarceration is community service requirements. As with incarceration, community service imposes time costs on the individual as a substitute for monetary costs. However, those time costs generate positive benefits for society in a way that they might not in the case of incarceration. In addition, community service avoids the social costs of incarceration (see below).

point at which the possibility of incarceration should be triggered. Should an individual who undertakes a great deal of care still be subject to incarceration if an accident occurs? Or should the possibility of incarceration be reserved for cases where the level of care was "criminally" low? This distinction is similar to the distinction in tort law between the use of a strict liability rule (liability regardless of the level of care) and a negligence rule (liability only for failure to exercise "due" care). While such a distinction has not been made in previous economic analyses of incarceration, it is crucial in analyzing the role of incarceration in ensuring efficient incentives. In particular, we show below that an incarceration rule based on strict liability principles will not generally yield efficient care, while a rule based upon negligence principles can.⁴⁶

Since it is most closely linked to the analyses of fines above, we consider first an incarceration rule based on strict liability principles. In other words, we assume that if an accident occurs, a jail sentence will be imposed with the length of the jail sentence determined by the level of damages. Since both the disutility of the jail sentence and the cost of incarceration are monotonically increasing in the number of years spent in jail, we can specify the model in terms of J rather than the number of years. Thus, $J = J(D)$ and $S = S(J)$. In addition, we allow for the possibility that the jail sentence will result in reputational losses that exceed the direct disutility of jail. Thus, the total cost of a jail sentence to the worker is $(1 + \phi)J$, where $\phi J \geq 0$ equals reputational costs, which are assumed to be proportional to the severity of the jail sentence.

With incarceration the socially optimal choices of a and x will solve

$$\begin{aligned} & \text{maximize } B(x) - C(x) - V(a) \\ & - p\{D(x, a) + J(D(x, a)) + S(J(D(x, a)))\}, \end{aligned} \quad (24)$$

which yields the following first-order conditions:

$$B'(x) - C'(x) - p\{1 + J' + S'J'\}D_x(x, a) = 0, \quad \text{and} \quad (25)$$

$$V'(a) + p\{1 + J' + S'J'\}D_a(x, a) = 0. \quad (26)$$

Note that in this formulation we assume that possible reputational losses to the worker from incarceration are not viewed as social losses.⁴⁷ While this violates a strict utilitarian approach to defining social welfare, it is consistent with the belief by some that certain private costs resulting from criminal activity should not be

⁴⁶This is in direct contrast to the standard results from the economic analysis of torts. It is well-known that in unilateral care cases efficient care by an individual can be induced through either strict liability or negligence. Even in the standard analysis of bilateral care, strict liability imposed on one party will lead to efficient care by that party, conditional on the care levels of others (which may not be efficient). Thus, while the overall outcome may not be efficient, the incentives of the party facing strict liability will be correct. (See, for example, [25]). This result does not hold in the case considered here, however.

⁴⁷Reputational losses may result in social losses if they prevent the worker from being employed to his fullest potential. We ignore this possibility here.

included in measures of social costs (e.g. [13]). The implications of relaxing this are noted below.

Likewise, with incarceration of individuals but no fines, the firm chooses x to

$$\text{maximize } B(x) - C(x) - W \quad (27)$$

$$\text{subject to the employee choosing } a \text{ so as to} \quad (28)$$

$$\text{maximize } W - V(a) - p(1 + \phi)J(D) \text{ and}$$

$$W - V(a) - p(1 + \phi)J(D) = U^0. \quad (29)$$

Note that this problem has the same form as (4)–(6), with $F = 0$ and $G = (1 + \phi)J$. Thus, the analysis of the private decisions in the case of fines can be used to infer behavior when incarceration is used instead.

Consider the case where wages can vary with both x and a , i.e., $W = W(a, x)$.⁴⁸ Letting $F = 0$ in (10), the firm chooses a wage contract of the form

$$W(a, x) = k(x) = U^0 + V(a_e) + p(1 + \phi)J(D(x, a_e)), \quad (30)$$

where a_e is the new equilibrium choice of a . Note that this wage function is independent of the actual choice of a . Given this function, the worker chooses a to satisfy the following first-order condition:

$$V'(a) + p(1 + \phi)J'D_a(x, a) = 0. \quad (31)$$

Likewise, the firm's choice of x is given by

$$B'(x) - C'(x) - p(1 + \phi)J'D_x(x, a_e) = 0. \quad (32)$$

Comparing (25) and (26) to (31) and (32) implies that, to ensure efficient choices of x and a through incarceration, the following condition must hold:

$$p(1 + \phi)J' = p(1 + J' + J'S'), \quad (33)$$

or, equivalently,

$$J' = 1/(\phi - S'). \quad (33')$$

From (31) it is clear that, in order to achieve efficiency, J' must be strictly positive.⁴⁹ Clearly this requires that $\phi > S'$, i.e., that marginal reputational costs exceed the marginal costs of incarceration. If the marginal reputational costs are small relative to incarceration costs, then efficiency cannot be achieved through incarceration. The inability to ensure efficiency through incarceration when repu-

⁴⁸The results for the other forms of W are similar. For example, if W is independent of x , then the same issues arise in ensuring an efficient choice of a . The only difference is that, since backward shifting is not possible, there is no way to ensure an efficient choice of x .

⁴⁹Since $V' > 0$ and $D_a < 0$, if $J' \leq 0$, then the agent would choose a equal to zero, which is assumed to be inefficient.

tation costs are “small” can be explained as follows. Suppose reputational costs are zero. Then the private costs of incarceration are only the opportunity costs of the worker, J . The social costs, on the other hand, are the opportunity costs, plus the damages, plus the costs of incarceration, $D + J + S$. Thus, regardless of the form of J , private costs are less than social costs and incarceration cannot induce efficiency.⁵⁰ Contrary to what has been thought, however, the inefficiency of incarceration here does not stem from the mere existence of incarceration costs. It should be clear from the above model that, even with $S = 0$, the problem remains. As long as the opportunity cost of the time spent in jail is included in the measure of social losses, as is typical in models of incarceration [20, 28], the private costs (J) will still be less than the social costs ($D + J$) and J cannot be set so that private and social costs are equal even at the margin.⁵¹ As a result, the worker will choose a level of a that is less than the efficient level. A similar result holds for the firm’s choice of x . This inefficiency can only be offset by sufficiently high reputational costs.

Note, also, that the potential inefficiency of incarceration cannot be eliminated by coupling incarceration with fines. If the worker’s assets are insufficient or the efficient fine would exceed a statutory limit, then in the relevant range the worker will view the fine as fixed at its maximum level (equal to his assets or the legal limit).⁵² The existence of the fine will then have no effect on the worker’s marginal incentives. Thus, in the absence of sufficient reputational costs, marginal social costs will continue to exceed marginal private costs and an inefficient amount of care will be chosen.

These conclusions regarding the inefficiency of incarceration hinge on the strict liability form of the incarceration rule that is considered above. In contrast, it is possible to base an incarceration rule on negligence principles, where a jail sentence would only be imposed if the level of care were less than some standard. We consider a negligence-based approach next and show that such a policy could induce efficient care.

Suppose, for example, that a jail sentence with disutility J is imposed if an accident occurs and care was less than the due standard. (No sentence is imposed if the worker took sufficient care.) Assume that the due standard of care is set equal to the socially efficient level of a (a^*) which solves (1).⁵³ For simplicity, we consider only the case where W is independent of x and a .⁵⁴ In this case, the

⁵⁰A similar conclusion would hold if reputation costs are positive but are included in the measure of social cost.

⁵¹Note that this inefficiency does not arise in the models used in [20] and [28] since they consider only discrete choices problems (whether to undertake the activity or not) where marginal incentives are not an issue.

⁵²If the level of damages at the efficient choice of a is below the worker’s assets, then efficient care can be induced through the use of a fine alone (see analysis of fines above) and incarceration is unnecessary.

⁵³This is a standard assumption in economic analyses of negligence rules (e.g. [25]). In practice, however, the actual standard of care applied may differ from the efficient level. See [10] for a discussion of related issues.

⁵⁴This assumption simplifies the analysis by allowing the decisions of the worker and the firm to be considered separately. It implies, however, that penalties imposed on the worker cannot be shifted to the firm, so that, in the absence of direct penalties on the firm, an efficient level of x will not be chosen. Here our interest is simply in the effects of incarceration on incentives of the individual on whom it is imposed rather than the possibility for shifting costs between parties.

worker will choose to maximize his net return, given by

$$W - V(a) - pJ(1 + \phi) \quad \text{if } a < a^*, \quad \text{and} \quad (\text{i})$$

$$W - V(a) \quad \text{if } a \geq a^*. \quad (\text{ii})$$

Let a' denote the level of a that maximizes (i).⁵⁵ Since $V' > 0$, the level of a that maximizes (ii) (given $a \geq a^*$) is clearly a^* . Thus, the worker will choose a^* if and only if

$$W - V(a') - pJ(1 + \phi) \leq W - V(a^*). \quad (34)$$

(We assume that when the two returns are equal, the worker chooses a^* .) Thus, if the social planner wants to ensure an efficient choice of a , he should set the jail sentence to satisfy

$$J \geq \{V(a^*) - V(a')\}/p(1 + \phi) > 0. \quad (35)$$

In other words, if a jail sentence with disutility exceeding $\{V(a^*) - V(a')\}/p(1 + \phi)$ is imposed when an accident occurs if and only if the worker's choice of a was less than a^* , the worker will be induced to choose the first best level of care a^* . Note that, in this case, no jail sentences would ever actually be imposed (if the rule works perfectly) so that the costs of incarceration would never be incurred.

Condition (35) suggests that the efficient penalty is independent of the level of damages. Instead, it is determined by the size of the penalty necessary to deter inappropriate behavior (defined to be $a < a^*$). Clearly, the greater the private benefit of inappropriate behavior ($V(a^*) - V(a')$), the larger the jail sentence must be to deter it. Likewise, the smaller the probability that it will be imposed, the larger the necessary penalty.⁵⁶ Positive reputational costs also reduce the required sentence.

Finally, consider the impacts of fines in this context. If the maximum fine is imposed when an accident occurs regardless of the level of a , then condition (35) will still determine the efficient jail sentence since the worker would incur the penalty regardless of his choice of a , i.e., the penalty will enter both sides of (34). However, if the fine is imposed along with the jail sentence only when a is less than the standard, then clearly the level of J necessary to induce deterrence will be smaller. Note that, to the extent that in practice the system is imperfect and jail sentences are actually imposed, smaller jail sentences would result in lower social costs. Thus, coupling incarceration with a fine imposed only if the level of care falls below the standard would be preferable to the use of incarceration alone.

Although a negligence standard for criminal penalties no doubt sounds like an unrealistic proposal to those familiar with the use of criminal penalties in nonenvironmental settings, this concept has recently been applied.⁵⁷ According to Section 1319(c)(1) of Title 33 of the U.S. Code, negligent violations of Clean Water Act

⁵⁵Note that, if J is independent of damages, then $a' = 0$.

⁵⁶These results are consistent with the results obtained in [28].

⁵⁷Negligence standards are not entirely uncommon in other areas of criminal law. For example, individuals who have engaged in reckless driving could well be charged with negligent homicide. We are indebted to a reviewer for noting this parallel.

provisions can now trigger criminal penalties. Fines of up to \$25,000 per day and 1 year in prison are authorized for a first offense with higher penalties authorized for repeat offenders. Our model suggests that this novel approach to enforcement may be filling an important niche.

IV. SUMMARY AND CONCLUSIONS

The traditional penalty structure for managing environmental risks involves administrative or civil financial penalties imposed on organizations rather than penalties or jail sentences imposed on individuals.⁵⁸ Targeting firms rather than individuals involves lower transactions costs and is less likely to create serious financial disruptions.⁵⁹ The perceived effectiveness of this approach in creating appropriate incentives for workers hinges on the firm's ability to sanction its employees internally for actions that contribute to environmental risks.

In practice, the efficient management of environmental risks is often more complicated than implied by this simple rationale. Risks can result from a combination of actions by workers and management. While workers may control day-to-day decisions that affect the level of risk, managers set overall safety policies and make decisions about safety investments and expansion of the firm's operations, which may also affect the magnitude of the associated risks. In such cases the enforcement approach must be structured to ensure that both parties face efficient incentives to reduce these risks.

In this paper we have examined the efficiency of alternative penalty structures, including both corporate and individual fines as well as incarceration, in the context of this more realistic setting. In particular, we compare alternative approaches, recognizing the role that both workers and management play in determining the ultimate level of environmental risks and the contractual relationships that allow workers and management to shift the ultimate incidence of any imposed penalty.

Using a simple principal-agent model, we have derived conditions for the traditional approach of using corporate penalties to be justified on efficiency grounds. Corporate penalties can be fully efficient if the firm can adequately monitor the employee's activities and base his/her wages on the environmental consequences of those activities. With full observability and completely flexible compensation schemes, the firm is able to shift the corporate penalty onto the worker through a compensating wage reduction. This shifting induces efficient care by the workers, even if they face no direct penalties themselves. If, in addition, the worker has an exogenous reservation level of utility that dictates an "acceptable" wage level and the wage demand can vary with the risk-related decisions of the firm, then any wage reduction will, in turn, be shifted back to the firm. Our model

⁵⁸See [17] for a breakdown of the EPA's enforcement actions. This report shows that about 90% of federal enforcement actions take the form of administrative proceedings.

⁵⁹The difference in transactions costs between individual and corporate penalties, however, seems to be getting smaller over time. Officers and directors of corporations have traditionally been protected from liability for wrongs normally attributable to the organization. Increasingly these traditional protections are being eroded in both the common law and the statutory law, thereby reducing the traditional burden of proof faced by government enforcers. See [5] for a discussion of this trend in environmental law.

suggests that under these conditions efficient incentives could also have been achieved with a penalty on the individual worker without any penalty on the firm.

Under the full observability and wage flexibility assumptions the combined effect of forward and backward shifting of corporate penalties assures efficient incentives for both the worker and the firm; corporate and individual sanctions are perfect substitutes in terms of their incentive effects. Such a result is consistent with previous work showing that, when two parties have a contractual relationship, the efficient outcome can be achieved by placing the penalty on either party. Our analysis also makes clear, however, that it would not be efficient to fine *both* the worker and the firm for the full amount of damages. Given their lower transaction costs, corporate penalties would be preferred.

In practice fulfilling these conditions for perfect substitutability between corporate and individual penalties may not be possible. Long lag times in discovering the environmental effects of a worker's actions may prevent shifting the incidence of corporate penalties to responsible employees. Whenever wages are dictated by inflexible contracts established by bargaining that focuses on other issues, the necessary backward and forward shifting cannot occur. Finally corporate penalties that are large in comparison to the earnings of environmental risk managers cannot, as a practical matter, be recovered from compensating reductions in wages.

Without such backward and forward shifting the two penalty types are not necessarily perfect substitutes. One penalty or the other (or even some combination of the two) may be preferred. According to our analysis individual penalties may be preferred if wages cannot be based on the worker's day-to-day decisions, but workers can demand higher compensation from firms forcing them to face higher potential penalties. Likewise if wages are unrelated to the safety decisions of either the worker or the firm, then the problem reduces to one of "joint torts" and penalties on both the worker and the firm would be necessary to induce efficient incentives.

Neither type of financial penalty, either separately or in combination, can produce efficient behavior when the penalties are set incorrectly (as a result, for example, of institutional or statutory constraints) or whenever the parties involved have insufficient assets to pay the penalty. In such cases incarceration may promote efficiency when used either as a substitute for or in combination with financial penalties.

To consider the potential role for incarceration in this "second-best" world, we examined the incentive effects of two alternative sentencing or incarceration rules, one based upon strict liability principles and the other based upon negligence principles. Our results suggest that sentencing practices based upon strict liability principles can produce an efficient outcome if and only if reputational costs are large enough. Marginal reputational costs must exceed the marginal costs of incarceration. If the marginal reputational costs are small relative to incarceration costs, efficiency cannot be achieved through a strict liability incarceration rule. If reputational costs are zero, regardless of the form of J , private costs would be less than social costs and incarceration could not induce efficiency.

When a jail sentence is based on negligence principles (imposed only when employee actions fall below some established threshold), our analysis indicates that efficient incentives can be preserved even in the absence of reputational costs. In this case the efficient penalty would be independent of the level of damages. Instead, it would be determined by its effectiveness in deterring inappropriate

behavior. The greater the private benefit of inappropriate behavior, the larger the jail sentence must be to deter it. Likewise, the smaller the probability that the sentence would be imposed, the larger the necessary sentence. Positive reputational costs also reduce the required sentence. To the extent that in practice the system is imperfect (implying that fines by themselves are insufficient), smaller jail sentences coupled with fines imposed only if the level of care falls below the efficient level would be preferable to the use of incarceration alone.

REFERENCES

1. G. Becker, Crime and punishment: An economic approach, *J. Polit. Econom.* **78**, 169–217 (1968).
2. M. A. Cohen, Optimal enforcement strategy to prevent oil spills: An application of a principal–agent model with moral hazard, *J. Law Econom.* **30**, 23–51 (1987).
3. M. A. Cohen, Corporate crime and punishment: An update on sentencing practice in the federal courts, 1988–1990. *Boston Univ. Law Rev.* **71**, 247–280 (1991).
4. M. A. Cohen, Criminal penalties, in “Innovation in Environmental Policy: Economic and Legal Aspects of Recent Developments in Liability and Enforcement” (T. H. Tietenberg, Ed.), Edward Elgar Publishing, Cheltenham, England, (1992).
5. D. M. Cordiano and D. J. Blood, Individual liability for environmental law violations, *Conn. Bar J.* **64**, 214–236 (1990).
6. S. J. Grossman and O. D. Hart, An analysis of the principal–agent problem, *Econometrica* **51**, 7–45 (1983).
7. K. Hagerty, A. Ofer, and D. Siegel, “Managerial Compensation and Incentives to Engage in Far-Sighted Behavior,” delivered at the Allied Social Science Convention, Washington, DC (December 29, 1990).
8. B. Holmstrom, Moral hazard and observability, *Bell J. Econom.* **10**, 74–91 (1979).
9. D. Kirkpatrick, Environmentalism: The new crusade, *Fortune* **121**, 44–50, 54–55 (February 12, 1990).
10. C. D. Kolstad, T. S. Ulen, and G. V. Johnson, *Ex post* liability for harm vs. *ex ante* safety regulation: Substitutes or complements?, *Amer. Econom. Rev.* **80**, 888–901 (1990).
11. L. A. Kornhauser, An economic analysis of the choice between enterprise and personal liability for accidents, *Cal. Law Rev.* **70**, 1345–1392 (1982).
12. W. M. Landes and R. A. Posner, “The Economic Structure of Tort Law,” Harvard Univ. Press, Cambridge, MA (1987).
13. J. L. Lewin and W. N. Trumbull, The social value of crime, *Internat. Rev. Law. Econom.* **10**, 271–284 (1990).
14. T. Miceli, and K. Segerson, Joint liability in torts: Marginal and infra-marginal efficiency, *Internat. Rev. Law Econom.* **11**, 235–249 (1991).
15. H. A. Newman and D. W. Wright, Strict liability in a principal–agent model, *Internat. Rev. Law Econom.* **10**, 219–231 (1990).
16. Office of Enforcement and Compliance Monitoring, United States Environmental Protection Agency, “Fundamentals of Environmental Compliance Inspections,” Government Institutes, Inc., Rockville, MD (1989).
17. Office of Enforcement, “Enforcement Accomplishments: FY 1989,” United States Environmental Protection Agency, Washington, DC (1990).
18. A. M. Polinsky, Strict liability vs. negligence in a market setting, *Amer. Econom. Rev.* **70**, 363–367 (1980).
19. A. M. Polinsky and S. Shavell, The optimal tradeoff between the probability and magnitude of fines, *Amer. Econom. Rev.* **69**, 880–891 (1979).
20. A. M. Polinsky and S. Shavell, The optimal use of fines and imprisonment, *J. Public Econom.* **24**, 89–99 (1984).
21. A. H. Ringleb and S. N. Wiggins, Liability and large-scale, long-term hazards, *J. Polit. Econom.* **98**, 574–595 (1990).
22. K. Segerson, Liability for groundwater contamination from pesticides, *J. Environ. Econom. Management* **19**, 227–243 (1990).

23. K. Segerson and T. Tietenberg, Defining Efficient Sanctions, in "Innovation in Environmental Policy: Economic and Legal Aspects of Recent Developments in Liability and Enforcement" (T. H. Tietenberg, Ed.), Edward Elgar Publishing, Cheltenham, England, (1992).
24. S. Shavell, Risk-sharing and incentives in the principal-agent relationship, *Bell J. Econom.* **10**, 55-73 (1979).
25. S. Shavell, Strict Liability versus negligence, *J. Legal Stud.* **9**, 1-25 (1980).
26. S. Shavell, Uncertainty over causation and the determination of civil liability, *J. Law Econom.* **28**, 587-609 (1985).
27. S. Shavell, The judgment proof problem, *Internat. Rev. Law Econom.* **6**, 45-58 (1986).
28. S. Shavell, The optimal use of nonmonetary sanctions as a deterrent, *Amer. Econom. Rev.* **77**, 584-592 (1987).
29. S. Shavell, "Economic Analysis of Accident Law," Harvard Univ. Press, Cambridge, MA (1987).
30. J. W. Starr and T. J. Kelly Jr., Environmental crimes and the sentencing guidelines: The time has come . . . and it is hard time, *Environ. Law Rep.* **20**, 10096-10104 (1990).
31. A. O. Sykes, An efficiency analysis of vicarious liability under the law of agency, *Yale Law J.* **91**, 169-206 (1981).
32. A. O. Sykes, The economics of vicarious liability, *Yale Law J.* **93**, 1231-1280 (1984).
33. T. H. Tietenberg, Indivisible toxic torts: The economics of joint and several liability, *Land Econom.* **65**, 305-319 (1989).