THE SOCIAL DESIRABILITY OF PUNISHMENT AVOIDANCE

Avraham D. Tabbach*

*Tel Aviv University, adtabbac@post.tau.ac.il
This working paper is hosted by The Berkeley Electronic Press (bepress) and may not be commercially reproduced without the permission of the copyright holder.
http://law.bepress.com/taulwps/art101
Copyright ©2009 by the author.
THE SOCIAL DESIRABILITY OF PUNISHMENT AVOIDANCE

Avraham D. Tabbach

Abstract

This paper argues that the law should sometimes encourage offenders to incur costs in order to avoid punishment, or, at the least, should not discourage such efforts. Avoidance efforts, such as concealment of evidence, perjury, or obstruction of justice, are generally deemed socially undesirable because they waste real resources and reduce deterrence. However, since avoidance efforts are also costly to offenders, they may substitute for socially costlier punishments such as imprisonment. If the resulting savings in punishment costs outweigh the social harm associated with less deterrence or the additional enforcement costs required for maintaining the same level of deterrence, then avoidance efforts are socially desirable. This, however, does not imply that punishing avoidance or otherwise increasing its private costs or decreasing its productivity is socially undesirable. Rather, it suggests that such measures should discourage avoidance as little as possible or even encourage it. This way, the social costs of avoidance can be reduced while its social benefits are maintained. This paper also casts doubt on the argument that sanctions for the underlying offense should generally not be maximal if avoidance is present. It shows that this result holds only if fines are the sole form of punishment. Otherwise, if optimal punishment utilizes both fines and imprisonment, fines should be maximal. Similarly, if imprisonment is the sole form of punishment, it should be maximal as well. The latter result is another manifestation of the social desirability of encouraging avoidance efforts.
THE SOCIAL DESIRABILITY OF PUNISHMENT AVOIDANCE

Avraham D. Tabbach **

ABSTRACT

This paper argues that the law should sometimes encourage offenders to incur costs in order to avoid punishment, or, at the least, should not discourage such efforts. Avoidance efforts, such as concealment of evidence, perjury, or obstruction of justice, are generally deemed socially undesirable because they waste real resources and reduce deterrence. However, since avoidance efforts are also costly to offenders, they may substitute for socially costlier punishments such as imprisonment. If the resulting savings in punishment costs outweigh the social harm associated with less deterrence or the additional enforcement costs required for maintaining the same level of deterrence, then avoidance efforts are socially desirable. This, however, does not imply that punishing avoidance or otherwise increasing its private costs or decreasing its productivity is socially undesirable. Rather, it suggests that such measures should discourage avoidance as little as possible or even encourage it. This way, the social costs of avoidance can be reduced while its social benefits are maintained. This paper also casts doubt on the argument that sanctions for the underlying offense should generally not be maximal if avoidance is present. It shows that this result holds only if fines are the sole form of punishment. Otherwise, if optimal punishment utilizes both fines and imprisonment, fines should be maximal. Similarly, if imprisonment is the sole form of punishment, it should be maximal as well. The latter result is another manifestation of the social desirability of encouraging avoidance efforts.
1. INTRODUCTION

Avoidance efforts to escape punishment are a widespread phenomenon and a common feature of crime and law enforcement. They take various forms, such as fleeing the scene of a crime, concealing incriminating evidence, influencing witnesses, lying, hiding, litigating, and so forth. Intuitively, avoidance efforts are socially costly because they waste real resources and reduce expected punishment, and thereby, deterrence. Therefore, it is not surprising that many lawmakers, legal scholars and economists recommend imposing sanctions or taking other measures to discourage such efforts (Sanchirico, 2006; Stanely, 1995).

The law has always attempted to deter certain avoidance efforts. Some avoidance efforts, such as perjury and obstruction of justice, are themselves deemed punishable crimes. Others, such as employing "sophisticated means" while committing a crime, are not deemed crimes themselves, but they may increase the punishment for the principal crime. Recently, after evidentiary fouls such as those at Enron, WorldCom and HealthSouth, lawmakers responded by broadening the definition of obstruction of justice and stiffening the penalties for substantially obstructive acts with a view toward deterring or discouraging such conduct (Sanchirico, 2006, p 1331).

This paper argues that the social (un)desirability of avoidance efforts crucially depends on the form of sanctions. The common intuition as to why avoidance is socially undesirable applies if sanctions are socially costless (i.e., monetary sanctions) or if their social costs are less than the costs imposed on the offenders. However, if sanctions are socially costly over and above the costs incurred by the offenders, as is

---

1 Sophisticated means are generally defined as "especially complex or especially intricate offense conduct pertaining to the execution or concealment of an offense. Conduct such as hiding assets or transactions, or both, through the use of fictitious entities, corporate shells, or offshore financial accounts ordinarily indicates sophisticated means. Their utilization in offenses such as tax evasion or basic economic offenses (theft, larceny and so on) under certain circumstances may increase the offense level by two.

http://law.bepress.com/taulwps/art101
generally the case with imprisonment, then avoidance efforts may confer a social benefit by reducing punishment costs that might outweigh their social costs. Since avoidance efforts are also costly to offenders, they may substitute for socially costlier punishments. If, for example, the net savings in public punishment costs outweigh the social harm associated with less deterrence or the additional enforcement costs required for maintaining the same deterrence level, then avoidance is socially beneficial. To illustrate this, consider the following example:

**Example 1:** Suppose that without avoidance efforts, the punishment for some offense is 10 years in prison with a probability of 10% for an expected punishment of 1 year. Suppose also that each year in prison creates a disutility of 100 to offenders and an additional cost of 50 to society. Finally, suppose that offenders can engage in some avoidance measure that costs the equivalent of 9.9 and can reduce the probability of punishment from 10% to 9% (i.e., by 1%).

Offenders are better off engaging in this kind of avoidance measure, spending 9.9 to save 10 (1% × 10 × 100). Deterrence is negligibly reduced since the actual expected sanctions plus the offenders' avoidance costs fall from 100 (10% × 10 × 100 + 0) to 99.9 (9% × 10 × 100 + 9.9). Enforcement costs are unchanged. However, since the probability of punishment is reduced by 1%, the public costs of punishment are reduced from 50 (10% × 10 × 50) to 45 (9% × 10 × 50). Thus, social welfare rises by 5 (50 – 45).

In view of the possible social benefits, attitudes toward avoidance may change dramatically. Not only should avoidance efforts not necessarily be discouraged, as is usually suggested, but in fact, society may consider adopting schemes that will

---

2 Stanley (1995), Sanchirico (2006), and others make no distinction between fines and imprisonment with regard to the social undesirability of avoidance efforts. Sanchirico (2006, pp. 1363-4), for example, discusses imprisonment explicitly, but completely ignores the benefits of avoidance in saving public punishment costs.
encourage avoidance. If properly designed, these schemes will induce avoidance which will substitute for imprisonment, thereby enhancing social welfare.\(^3\)

The social desirability of avoidance efforts does not imply that punishing avoidance or, if possible, taking other measures to increase its private costs or reduce its productivity is socially undesirable. On the contrary, such measures are generally welfare-enhancing. However, in opposition to commonly held views, such measures should discourage avoidance as little as possible or even encourage it.\(^4\) This idea can be explained roughly as follows: punishing avoidance, increasing its private costs, or decreasing its productivity without completely discouraging it decreases the downside of avoidance in terms of deterrence dilution, while not eliminating (potentially even increasing) the upside in terms of punishment substitution. To illustrate, consider the following example, which is a variation of Example 1:

**Example 2:** Suppose that without avoidance efforts, the punishment for some offense is 10 years in prison with a probability of 10% for an expected sanction of 1 year. Suppose that each year in prison creates a disutility of 100 to offenders and costs society an additional 50. Suppose also that offenders can engage in some avoidance measure that costs the equivalent of 5.5 and reduces the probability of punishment from 10% to 9% (i.e., by 1%). Suppose, finally, that punishment for the underlying offense can be increased if offenders engage in avoidance, so avoidance itself can be punished.

---

\(^3\) To illustrate, modify Example 1 and assume that the private costs of avoidance are 10.1 rather than 9.9. Since the private benefit of avoidance in reducing the expected sanction is only 10, offenders will not engage in avoidance. By subsidizing the private costs of avoidance by 0.2, avoidance can be positively induced and social welfare increased by 5.

\(^4\) But how is it possible to increase the private costs of avoidance or decrease its productivity, for example, without discouraging avoidance? The idea is to change the *total* costs or productivity without affecting the *marginal* costs or productivity of avoidance. Borrowing from public finance jargon, the idea is to tax away any pure private profits (i.e., “avoidance surplus”) associated with avoidance.
If avoidance is punished by more than 6 months, say by 7 months, avoidance will be completely discouraged, because its private benefit 10 (1%X10X100) will be less than its total private costs 10.75 [the sum of the direct costs, 5.5, and the indirect costs stemming from the increased punishment, 5.25 (9%X(7/12)X100)]. However, if avoidance is punished by just less than 6 months, avoidance will not be discouraged, because its private benefit, 10, will be just greater than its total private costs, 9.9 (5.5+9%X0.49X100). Deterrence will be negligibly affected since offenders still face a total cost of approximately 100 (5.5+9%X10.49X100). Enforcement costs are unchanged. However, since the expected punishment is less than one year (9%X10.49), the public costs of imprisonment are reduced from 50 (10%X10X100) to 47.20 (9%X10.49X50). Thus, social welfare rises by 2.8 (50 – 47.2).\(^5\)

Indeed, as is shown in Section 4, punishing avoidance is socially desirable, but the optimal punishment for avoidance should encourage rather than discourage it.

Avoidance efforts of different sorts are not only considered socially undesirable, but may also qualify Becker's (1968) famous maximal sanctions result, according to which sanctions – imprisonment or fines – should be employed to the maximum possible level, while the probability of sanctions should be set appropriately low, so as to save enforcement efforts.\(^6\) As Malik (1990) explains, avoidance makes fines socially costly at the margin since more severe punishment induces more avoidance. If so, fines may not be utilized to the maximum possible level, but should be set so that their marginal benefits and costs are equal.

\(^5\) Alternatively, suppose that the private costs of avoidance can be increased by imposing a unit tax on the avoidance measure. Although avoidance can be completely discouraged by setting the unit tax just above 4.5, the optimal unit tax should be set just below 4.5 (say, 4.49). This would not discourage avoidance because the total private costs of avoidance 9.99 (5.5+4.49) are less than its private benefits, 10. However, it would “tax” away the pure profits associated with avoidance. Social welfare is thereby increased because deterrence is negligibly affected while the public costs of imprisonment are reduced from 50 (10%X10X50) to 45 (9%X10X50).

\(^6\) This result is qualified in the literature on various grounds (see, generally, Polinsky and Shavell, 2007). For our purposes, it suffices that it strictly applies if offenders are risk-neutral in fines/imprisonment. Throughout this paper, risk neutrality is assumed.
This paper shows that Malik's (1990) argument against maximal sanctions applies if, and only if, sanctions are monetary or more generally have no social costs above and beyond the costs incurred by offenders. Otherwise, Becker's result holds; the sanction for the underlying offense should be maximal, and enforcement efforts should be appropriately low, regardless of avoidance. The reason is simple. Without avoidance, increasing imprisonment and reducing its probability so as to leave the expected sanction unaltered saves on enforcement costs without affecting the costs of punishment. With avoidance, increasing imprisonment provides another type of benefit because it induces more avoidance, which substitutes for costlier punishment, thereby saving not only enforcement but also punishment costs.

Furthermore, if optimality requires utilizing both fines and imprisonment, Malik's argument again does not apply, while Becker's result holds. Fines should be maximal regardless of avoidance. The explanation is that avoidance and deterrence depend on the total rather than on the mix of punishment. Therefore, increasing fines to the maximum and appropriately reducing imprisonment while keeping total punishment constant does not alter avoidance or deterrence, yet saves on the punishment costs.

The results in this paper are based on the insight that avoidance efforts can substitute for socially costly sanctions, such as imprisonment. This immediately raises the question of why imprisonment rather than fines or other less costly corporal sanctions is employed. How is the use of imprisonment consistent with the fact that

---

7 Malik (1990) implicitly confined himself to fines. He did not mention how his analysis and results would change if sanctions took the form of imprisonment or imprisonment and fines. However, the logic of his argument seems to apply to all forms of sanctions. The restrictive nature of his results is not generally recognized. For example, Innes (2001, p. 240) presents Malik's argument, referring generally to "sanctions" and not restrictively to "fines" or "monetary sanctions". Similarly, Sanchirico (2006, p. 1350, note 64) points out that the great Italian philosopher Cesare Beccaria, who is among the founders of the classical approach to law enforcement, anticipated Malik's argument supporting mild punishment: "The worse the ill that confronts them, the more men are driven to evade it. The very savagery of a punishment has this effect, and to avoid the penalty for the one crime they have already committed, men commit other crimes". This statement (Beccaria, pp. 43-44) is clearly meant to apply to both fines and imprisonment.

8 See, however, the qualification in Section 5.
offenders engage in or spend on avoidance? Several explanations may be offered. One could postulate a constraint on the choice of punishment instrument. For example, some may argue that for moral or political considerations, corporal punishments (other than imprisonment) are simply unacceptable punishment tools. Others may claim that we simply have a strong aversion to their use (Miceli, 1990; Polinsky and Shavell, 200b). Thus, avoidance may take a non-monetary form, while imprisonment is utilized because offenders are judgment-proof (i.e., wealth-constrained). In addition, some may believe that fines and imprisonment are non-commensurable as punishment instruments, so that only imprisonment should be employed. Others may argue that the government should not benefit from punishing offenders, because otherwise it has an incentive to punish excessively (Friedman, 1999). Thus, fines (and certain forms of corporal punishment) are ruled out or kept to a minimum as punishment instruments. Avoidance can be monetary or non-monetary, while punishment for the principal offense is constrained so as to be socially costly. The above scenarios are also roughly consistent with the legal reality, in which imprisonment is utilized while fines and other corporal punishments are not fully exploited. In any event, the second-best nature of the problem is evident and should be kept in mind.

The results in this paper also depend critically on the notion that avoidance measures do not generate negative externalities beyond those portrayed above. This issue, which determines the scope of this paper, will be discussed in Section 5.2.

This paper is organized as follows. Section 2 introduces the general model. Section 3 reviews the properties of optimal law enforcement with and without avoidance.
showing that avoidance efforts may be socially desirable. Section 4 explores methods for controlling avoidance, such as punishment, and shows that the optimal enforcement scheme involves punishing, but not discouraging avoidance, even though it is feasible and costless. Section 5 extends the analysis discusses several other issues that may affect the conclusions of this paper.

2. The Model

Risk-neutral individuals with respect to imprisonment choose whether to engage in a proscribed activity that causes positive harm \( h \). Different individuals obtain various benefits from the activity \( b \), assumed to be distributed with continuous density \( f(b) \) and cumulative distribution \( F(b) \) on the support \([0, \infty)\).\(^{11}\)

If an individual commits the harmful act, he faces some probability of being caught and imprisoned. The maximum imprisonment term \( s \) is assumed to be bounded so that \( s \leq s_{\text{max}} \). Imprisonment is socially costly, imposing a monetary equivalent of \( \alpha \) on offenders and additional costs of \( \gamma \) on society,\(^{12}\) per unit imprisonment term. The latter represents the public cost of operating the prisons. Thus, the social costs of imprisonment \( \alpha + \gamma \) are greater than the costs incurred by offenders \( \alpha \).

The probability of punishment \( p \) depends on the public enforcement expenditures \( e \) to apprehend and punish offenders and on offenders' avoidance efforts \( z \) to conceal their identity or the commission of a crime or to otherwise reduce the probability of being detected and punished. The avoidance efforts, which are assumed not to create

---

\(^{10}\) The model loosely follows those proposed by Polinsky and Shavell (1984) and Innes (2001).

\(^{11}\) This implies that complete deterrence is generally undesirable. If, instead \( b \) were distributed on \([0, \bar{b}]\) and \( h > \bar{b} \), then it would not be desirable for anyone to engage in the proscribed activity. However, complete deterrence is not necessarily optimal because of high enforcement and punishment costs. In any event, this assumption is not critical for the results of this paper. See Section 5.1.

\(^{12}\) The average cost of imprisoning a convict for a year is about $22,650 (U.S. Department of Justice, 2001). Implicitly, this paper ignores any possible incapacitation or rehabilitation value of imprisonment. Such value may clearly affect the results. If the real social costs of imprisonment are less than the costs incurred by offenders, avoidance will be definitely undesirable, as explained below.
other externalities, may be monetary or non-monetary, but all have a monetary equivalent of \( z \). The probability of punishment thus takes the form \( p(e, z) \) with the following usual properties: \( p_e(e, z) > 0, p_{ee}(e, z) < 0 \) and \( p_z(e, z) < 0, p_{zz}(e, z) > 0 \), reflecting decreasing marginal returns to both \( e \) and \( z \) respectively. For simplicity, it is assumed that \( p_{ee}(e, z) = 0 \), i.e., the effectiveness of enforcement expenditures or avoidance efforts on \( p \) are independent of one another.\(^{13}\)

Individuals engaging in the proscribed activity choose their level of avoidance efforts so as to minimize expected losses, which consist of the expected sanctions plus avoidance costs. Formally,

\[
L = \min \left[ p(e, z) \alpha s + z \right].
\]

The optimal avoidance efforts, denoted \( z^* \) (asterisks are used to denote optimal choices throughout), assuming an interior solution, satisfy the first-order condition

\[
- p_z(e, z) \alpha s = 1.
\]

Evidently, optimal avoidance efforts increase with higher sanctions but are unaffected by enforcement efforts, since by assumption, \( p_{ee}(e, z) = 0 \). Formally,

\[
\frac{dz^*}{ds} > 0 \quad \text{and} \quad \frac{dz^*}{de} = 0.
\]

Individuals will engage in the proscribed activity if and only if the benefits are greater than the minimum expected losses, i.e., iff \( b > L \).\(^{14}\) Therefore, the level of deterrence is determined by:

\[
\tilde{b} = L = \min[p(e, z) \alpha s + z] = p(e, z^*) \alpha s + z^*.
\]

\(^{13}\) This assumption cannot be strictly valid because \( p(0, z) = 0 \). However, as explained in Section 5.5, it does not affect the main qualitative results of this paper.

\(^{14}\) If \( b = z^* + p(e, z^*) \alpha s \), individuals are indifferent as to whether to engage in the proscribed activity or not. By convention, it is assumed that they do not.
Evidently, the level of deterrence $\tilde{b}$ (or minimum expected losses $L$) increases with enforcement and punishment, although with the latter, the actual probability of punishment given adjustment of avoidance efforts, actually falls. Formally, by the envelope theorem,

$$\frac{db}{de} = \frac{\partial \tilde{b}}{\partial e} = p_e > 0, \quad \frac{db}{ds} = \frac{\partial \tilde{b}}{\partial s} = \alpha p(e, z) > 0, \quad \frac{dp(e, z)}{ds} = p_z(e, z) \frac{dz^*}{ds} < 0.$$

Social welfare is assumed to be equal to the aggregate gains to those engaging in the proscribed activity, less the harm done, the costs of avoidance, the private and the public costs of imprisonment for the detected subset, and the enforcement costs. Hence, social welfare can be expressed as:

$$W = \int [b - h - p(e, z)s(\alpha + \gamma)]f(b)db - e.$$

Before delving into the social desirability of avoidance efforts, let us first derive the optimal enforcement schemes under conditions of (1) no avoidance and (2) avoidance.

3. ANALYSIS

3.1. No Avoidance

Assume first a world in which avoidance technology is lacking or avoidance is completely deterred without cost (i.e., a world without avoidance). The social problem reduces to the standard problem of choosing the imprisonment term $s (\leq s_{\text{max}})$ and the enforcement expenditures $e$ so as to maximize social welfare:

$$W = \int [b - h - p(e, 0)s(\alpha + \gamma)]f(\tilde{b})db - e,$$

where $\tilde{b} = p(e, 0)\alpha s$ determines the level of deterrence.

It is well-known that if individuals are risk-neutral, the optimal imprisonment term is maximal (Becker, 1968). The proof is by contradiction. If the imprisonment term
were not maximal, it would be possible to increase it to the maximum and decrease enforcement expenditures so as to achieve the same expected punishment and, therefore, deterrence at lower costs.\footnote{To illustrate, suppose that the optimal solution is characterized by \( s^* < s_{\text{max}} \) and \( e^* > 0 \). Raise \( s \) to \( s_{\text{max}} \) and lower \( e \) to \( \tilde{e} \) such that \( p(\tilde{e},0)\alpha s_{\text{max}} = p(e^*,0)\alpha s^* \). Since the expected punishment is unaltered, offenders' total utility, the level of deterrence, and the public costs of punishment remain the same. However, enforcement costs are saved, so social welfare rises. Thus, \( s^* < s_{\text{max}}, e^* > 0 \) could not have been optimal (see Polinsky and Shavell, 1984).}

Given that \( s^* = s_{\text{max}} \), the optimal enforcement expenditures are determined by maximizing social welfare (7) with respect to \( e \). The resulting first-order condition, assuming an interior solution, may be written as:

\[
(h - p(e,0)\alpha s_{\text{max}} + (\alpha + \gamma) p(e,0)s_{\text{max}})\frac{dF(b)}{de} = 1 + [1 - F(b)](\alpha + \gamma) p_s(e,0)s_{\text{max}}.
\]  

(8)

The RHS of (8) reflects the marginal costs of raising enforcement expenditures, which include the direct costs involved, 1, and the indirect costs stemming from the increase in the private and public costs of imprisonment due to the apprehension of a greater proportion of offenders, \( [1 - F(b)](\alpha + \gamma) p_s(e,0)s_{\text{max}} \). The LHS of (8) reflects the marginal benefits of raising the enforcement expenditures, which are equal to the reduction in the number of offenders, \( \frac{dF(b)}{de} = p_s(e,0)\alpha s_{\text{max}}f(b) \), times the sum of the harm, \( h \), less the benefits to the marginal offender, \( p(e,0)\alpha s_{\text{max}} \), plus the savings in the private and public costs of imprisonment, \( (\alpha + \gamma) p(e,0)s_{\text{max}} \). At the optimum, of course, the marginal costs and benefits should be equal.

Condition (8) reveals, as is well-known, that the first-best behavior is generally non-optimal. Rather, some degree of under-deterrence (the benefits to some offenders are less than the harm) or over-deterrence (the benefits to some deterred individuals are
greater than the harm) is socially desirable (Polinsky and Shavell, 1984). To see this, suppose that \( s = s_{\text{max}} \) and \( e \) is such that expected sanctions are equal to harm, \( p(e,0)\alpha s_{\text{max}} = h \), so that there is first-best deterrence. Consider now a slight reduction in enforcement expenditures, leading to a slight reduction in \( p \). This reduces enforcement costs and saves the private and public costs of punishment for those engaging in the proscribed activity. It also leads the marginal individuals to engage in the proscribed activity. On the one hand, this has no first-order effect on social welfare because these marginal individuals receive benefits equal to the harm caused (since \( p(e,0)\alpha s_{\text{max}} = h \)). On the other hand, the private and public costs of punishment are increased because there is some probability of punishment of additional offenders. If the former effect is dominant, under-deterrence would be optimal; otherwise, over-deterrence might be socially desirable.\(^{16}\)

### 3.2. Avoidance

Alternatively, assume that controlling avoidance is prohibitively costly, such that avoidance is not controlled at all. The social problem is to choose the imprisonment term \( s \) (\( s \leq s_{\text{max}} \)) and the enforcement expenditures \( e \) so as to maximize social welfare:

\[
W = \int_{b}^{z} \left[ b - h - z - p(e, z)s(\alpha + \gamma) \right] \, f(\tilde{b}) \, db - e,
\]

where \( \tilde{b} = p(e, z^*)\alpha s + z^* \) determines the level of deterrence.

The solution to this problem is similar to the previous one (Section 3.1). In particular, the optimal imprisonment term is maximal. Otherwise, it would be possible

---

\(^{16}\) The possibility that the optimal enforcement scheme may be characterized by over-deterrence implies that the fact that avoidance reduces deterrence, at least to some extent, is not a social cost, as is commonly understood, but rather a social benefit. Moreover, the dilution of deterrence implies that there are additional offenders, some of whom will be punished. Therefore, avoidance may actually increase rather than decrease the (total) public costs of punishment.
to increase the term of imprisonment to the maximum and lower enforcement expenditures so as to achieve the same level of deterrence with lower enforcement and public punishment costs. To illustrate this, suppose that \( s^* < s_{\text{max}} \) and \( e^* > 0 \).

Raise \( s \) to \( s_{\text{max}} \) and change \( e \) to \( \tilde{e} \) such that the expected punishment plus optimal avoidance efforts are unaffected, i.e., 
\[
p(e^*, z^*)\alpha s_{\text{max}} + z^* = \tilde{p}(\tilde{e}, \tilde{z})\alpha s_{\text{max}} + \tilde{z}
\]
where \( z^* \) and \( \tilde{z} \) are the optimal avoidance efforts before and after the change in the enforcement scheme, respectively. By construction, deterrence and offenders' total utility are unaffected. Since punishment is increased but deterrence remains the same, enforcement expenditures are necessarily lower, \( \tilde{e} < e^* \). In addition, since punishment is increased, avoidance efforts are also increased, \( \tilde{z} > z^* \). This also implies that the expected sanction imposed on offenders is lower, i.e., 
\[
p(\tilde{e}, \tilde{z})\alpha s_{\text{max}} < p(e^*, z^*)\alpha s_{\text{max}}
\]
This, coupled with the fact that deterrence is unaffected, means that the public costs of punishment are lower as well, i.e., 
\[
[1 - F(\tilde{b})]p(\tilde{e}, \tilde{z})\gamma s_{\text{max}} < [1 - F(\tilde{b})]p(e^*, z^*)\gamma s_{\text{max}}
\]
Thus, not only are enforcement costs reduced, but the public costs of imprisonment are saved. Social welfare therefore rises, and \( s^* < s_{\text{max}} \), \( e^* > 0 \) could not be optimal.

The optimality of the maximum imprisonment term should be contrasted with Malik's (1990) result that fines need not be maximal if avoidance is considered. With imprisonment, increasing punishment to the maximum level is beneficial because it increases deterrence and induces avoidance efforts, thereby allowing the same level of deterrence to be achieved not only with less enforcement efforts, but also with less socially costly sanctions. Avoidance efforts effectively substitute for imprisonment.

With fines, avoidance efforts substitute for a socially costless sanctions. Therefore, instead of savings in the public costs of punishment, there is a revenue loss since
expected fines are lower. If this revenue loss is greater than the savings in the enforcement costs, maximal fines need not be optimal.\(^\text{17}\)

Given that \( s^* = s_{\text{max}} \), the optimal enforcement expenditures are determined by maximizing social welfare (6) with respect to \( e \). For an interior solution they should satisfy the first-order condition given by (9):

\[
(h - z - p(e, z)\alpha s_{\text{max}} + z + (\alpha + \gamma) p(e, z)s_{\text{max}}) \frac{dF(b)}{de} = 1 + [1 - F(\tilde{b})](\alpha + \gamma) p_*(e, z)s_{\text{max}}.
\]

The RHS and LHS of (9) reflect marginal benefits and costs, respectively, and have essentially the same form as without avoidance since the \( z' \)s on the LHS cancel out. The interpretation is also similar, except that the marginal benefits of increasing enforcement expenditures due to the reduction in the number of offenders saves avoidance efforts as well as the private and public costs of punishment, and the benefit to the marginal offender is defined at the optimum with regard to avoidance efforts, i.e., \( z^* + p(e, z^*)\alpha s_{\text{max}} \).

Condition (9), like condition (8), reveals that some degree of under or over-deterrence is socially desirable. The explanation is the same, so it is omitted.

3.3. Avoidance versus Non-Avoidance

As shown above, with or without avoidance, optimal enforcement is characterized by utilizing imprisonment to the maximum level and employing enforcement efforts

\(^{17}\) To be precise, interpret \( s \) as fines so that \( \gamma = -\alpha \). Suppose that \( s^* = s_{\text{max}} \) and \( e^* > 0 \). Now lower \( s \) to \( \tilde{s} \) and increase \( e \) to \( \tilde{e} \) such that \( p(e^*, z^*)\alpha s^* + z^* = p(\tilde{e}, \tilde{z})\alpha \tilde{s} + \tilde{z} \), where \( z^* \) and \( \tilde{z} \) are the optimal avoidance efforts before and after the enforcement scheme change. By construction, deterrence and offenders' total utility are unaffected. In addition, enforcement costs are increased. However, since \( \tilde{z} < z^* \), then \( p(e^*, z^*)\alpha s^* < p(\tilde{e}, \tilde{z})\alpha \tilde{s} \) so that government revenues are increased as well. \( [1 - F(\tilde{b})] p(\tilde{e}, \tilde{z})\alpha \tilde{s} > [1 - F(\tilde{b})] p(e^*, z^*)\alpha s^* \). If this increase is greater than the increase in enforcement costs, maximal fines are non-optimal.
so that under or over-deterrence results. Let us now show that a world in which avoidance is not controlled at all can be socially preferable to one without avoidance.

For that purpose, assume that the enforcement scheme without avoidance is set optimally, i.e., \( s_N^* = s_{\text{max}} \) and \( e_N^* \) satisfying condition (8), where the subscript \( N \) denotes a world without avoidance. Deterrence is determined by the expected sanction \( p(e_N^*, 0)\alpha s_{\text{max}} \). Now construct an enforcement scheme in the avoidance world giving rise to the same level of deterrence and compare the resulting social welfare.\(^ {18} \)

If the imprisonment term is set at \( s_{\text{max}} \), the enforcement expenditures, denoted \( e_A \) (the subscript \( A \) stands for a world with avoidance), should be set such that

\[
p(e_A, z^*) \alpha s_{\text{max}} + z^* = p(e_N^*, 0)\alpha s_{\text{max}},
\]

where \( z^* \) satisfies \( -p_A(e_A, z)\alpha s_{\text{max}} = 1 \). In other words, the expected sanction offenders face in the world with avoidance, \( p(e_A, z^*)\alpha s_{\text{max}} \), plus their optimal avoidance efforts, \( z^* \), should equal the expected sanction they face in the world without avoidance, \( p(e_N^*, 0)\alpha s_{\text{max}} \). Note that \( e_A \) is non-optimal (i.e., it does not satisfy condition (9)).

By construction, deterrence and the offenders' total utility are the same under both worlds. Since avoidance efforts reduce deterrence, enforcement expenditures must be higher in the world with avoidance than without it for deterrence to be maintained, i.e., \( e_A > e_N^* \). However, since avoidance efforts are positive, \( z^* > 0 \), the expected sanction that offenders face with avoidance is lower than without avoidance, i.e.,

\[
p(e_A, z^*)\alpha s_{\text{max}} < p(e_N^*, 0)\alpha s_{\text{max}}.
\]

This coupled with the fact that deterrence remains the

\(^{18} \text{Alternatively, one can construct an enforcement scheme giving rise to the same level of enforcement expenditures or to the same level of public costs of punishment. Such constructions, however, are generally more difficult to analyze, because the optimal enforcement scheme without avoidance may be characterized by under or over-deterrence. Therefore, if the level of deterrence is reduced, it will be considered a social cost (if there is under-deterrence) or a social benefit (if there is over-deterrence). Moreover, less deterrence may increase the public costs of punishment, because there are additional offenders. In any event, this paper completely avoids these difficulties by constructing throughout the proofs in which the level of deterrence is unaltered.}
same implies that the public costs of punishment are lower with avoidance than without it, i.e., $[1 - F(\tilde{b})]p(e_A, z^*)γs_{\max} < [1 - F(\tilde{b})]p(e_N^*, 0)γs_{\max}$; this is because avoidance efforts substitute for imprisonment. Thus, a world with avoidance that achieves the same optimal level of deterrence as a world without avoidance is associated with higher enforcement costs but lower public costs of punishment. Now, if savings in the public costs of punishment outweigh the increased enforcement costs, that is, if\(^{19}\)

$$[1 - F(\tilde{b})][p(e_N^*, 0)γs_{\max} - p(e_A, z^*)γs_{\max}] > [e_A - e_N^*],$$

avoidance will be socially desirable.

* A priori, it is difficult to tell if and under what circumstances condition (10) will hold, since it contains endogenous variables. Nevertheless, certain *necessary and sufficient* conditions regarding the values of the underlying primitive parameters can be demonstrated under which condition (10) holds. For example, if the public costs of imprisonment $γ$ were non-positive, avoidance would not confer any social benefits (the LHS in (10) would be non-positive). Therefore, condition (10) would not hold and avoidance would be socially undesirable. In other words, for avoidance efforts to be socially desirable, the public costs of punishment $γ$ must be positive.\(^{20}\) On the other hand, if the enforcement expenditures were very cheap and effective so that any level of detection (that is, any $p$) could be achieved at negligible costs, then avoidance would entail no social costs at all (the RHS of (10) would drop to zero); condition (10) would always hold, and, consequently, avoidance would always be

---

\(^{19}\) The avoidance world is associated with lower private punishment costs than the world without avoidance, but these costs are exactly offset by the avoidance costs offenders incur, since, by construction, the level of deterrence is maintained.

\(^{20}\) Generally speaking, the higher the public costs of punishment $γ$, the more likely it is that avoidance will be socially desirable, because the social benefits of avoidance will be higher.
socially desirable.\textsuperscript{21} In addition, the more costly the avoidance efforts, \( z^* \), and the less their effect on deterrence, \( p(e_N^*, 0)z_{\text{max}} - z^* - p(e_N^*, z^*)z_{\text{max}} \), the more likely it is that condition (10) will hold and avoidance will be socially desirable. This is because the more costly the avoidance efforts, the greater the savings in public punishment costs (the LHS in (10) will be higher), and the less the effect of avoidance on deterrence, the less the enforcement expenditures required for restoring deterrence (the RHS in (10) will be lower). Indeed, if avoidance has only a negligible effect on deterrence, i.e., \( p(e_N^*, z^*)z_{\text{max}} + z^* \approx p(e_N^*, 0)z_{\text{max}} \), the costs of maintaining deterrence will be very small (the RHS of (10) will tend to zero), and avoidance will definitely be socially desirable. This possibility is illustrated in Example 1 in the Introduction, in which the private costs and benefits of avoidance are almost equal, so deterrence is negligibly affected, but since avoidance efforts are positive, the public costs of punishment are reduced.\textsuperscript{22}

The analysis above underestimates the social desirability of avoidance in two important respects. First, since it compares an optimal enforcement scheme without avoidance to a non-optimal enforcement scheme with avoidance, condition (10) is a sufficient but not necessary condition for avoidance to be socially desirable. In other words, the analysis proves that avoidance \textit{can be} socially desirable, but it does not

\textsuperscript{21} Generally speaking, the less costly and more effective the enforcement expenditures, the more likely it is that avoidance will be socially desirable, because the social costs of avoidance in terms of increasing the costs of achieving any level of deterrence will be lower.

\textsuperscript{22} This last observation has an interesting implication regarding the social desirability of small changes in avoidance. Since the (private) optimal level of avoidance is determined such that the (private) marginal benefits and costs of avoidance are equal, i.e., \( -p_e(e, z)z = 1 \), it follows that at the optimum, small changes in avoidance have a negligible effect on deterrence. Therefore, regardless of whether (overall) avoidance is socially desirable, a small increase (decrease) in the private optimal level of avoidance is socially desirable (undesirable). A similar logic implies that starting from no avoidance, a small increase in avoidance is likely to be socially undesirable because the effect of avoidance on deterrence at that point is the greatest, i.e., \( -p_e(e, z)z_{\text{max}} - 1 \) is greatest at \( z = 0 \).
consider all the circumstances under which it will be.\textsuperscript{23} Second, and more importantly,
the analysis shows that a world in which avoidance is prohibitively costly to control
can be socially preferable to one in which avoidance is completely deterred without
costs. This does not imply that taking measures to control avoidance, if such measures
are not prohibitively costly, is socially undesirable. On the contrary, the analysis
suggests that taking steps that reduce the social costs of avoidance (the RHS of 10)
while, at the same time, maintaining or even enhancing its social benefits (the LHS of
10) will be socially desirable. Indeed, when such measures are considered, two
important results emerge: (1) avoidance becomes always socially desirable, and (2)
the optimal level of avoidance may actually be higher than if avoidance were not
controlled at all. The following section explores these possibilities.

4. THE SOCIAL DESIRABILITY OF PUNISHMENT AVOIDNACE

4.1. The Different Methods to Control Avoidance

Avoidance can be controlled in various ways. One simple way is to increase its
private costs by imposing, for example, a unit tax on the avoidance measure. This is
sometimes, but not always, feasible. Another way, suggested by Sanchirico (2006), is
to decrease the productivity of avoidance by employing different interrogation or
evidentiary rules that would make avoidance efforts less effective. For example,
forcing offenders to testify and respond to questions from memory makes lying less
effective (or more costly) in reducing the probability of punishment. A third way to
control avoidance is to increase the enforcement efforts and thereby the probability of
punishment for the underlying offense if the offenders engage in avoidance. This is, in
fact, a typical practice for prosecutors who direct investigative resources towards

\textsuperscript{23} For this reason, doing comparative analysis of (10) regarding different parameters is less instructive.
underlying offenses showing evidence of avoidance efforts. The fourth way to control avoidance is to punish it, either directly or indirectly, through an increase in the sanction for the underlying offense. This is apparently a method employed by lawmakers and judges, as evidenced, for example, by the recent legislation noted in the Introduction. In the present framework, this method presents a problem, because as is shown in Section 3.2, the optimal sanction for the underlying offense is maximal. Thus, the sanction cannot be further increased. To overcome this difficulty, $s_{\text{max}}$ can be interpreted not as a physical constraint reflecting, for example, the limited lifetime of individuals, but rather as an exogenous constraint on the maximum sanction for the underlying offense if there is no avoidance, demonstrating, for example, consideration for fairness. This interpretation is also consistent with the legal reality.

The various ways in which avoidance can be controlled differ in several important respects, including their feasibility, social costs, and possible effects on avoidance. However, under certain simplifying assumptions, all methods can discourage avoidance completely. To simplify matters, let us focus on the social desirability of punishing avoidance, observing, however, that our qualitative results generally apply to other methods of controlling avoidance.\textsuperscript{24}

4.2. Punishing Avoidance

Assume that offenders who engage in any level of avoidance face an additional fixed sanction $k$, which can be large enough to eliminate avoidance (the analysis if $k$ is not sufficiently large is straightforward). For simplicity, assume also that avoidance is always detected and punished if the underlying offense is detected and punished, and that detecting avoidance does not increase enforcement costs. Together, these

\textsuperscript{24} A full, complete comparison among the different methods is a task for future research.
assumptions imply that avoidance can be completely discouraged at no cost, thereby reflecting the best case against avoidance.

Given the enforcement scheme $e$, $s$ and $k$ offenders will engage in avoidance if and only if the expected sanction they face if they engage in avoidance plus their optimal avoidance efforts is less than or equal to the expected sanction they face if they do not engage in avoidance, that is, iff:

\[(11) \quad p(e, z^*)\alpha(s + k) + z^* \leq p(e,0)\alpha s ,\]

where $z^*$ satisfies $-p_z(e, z)\alpha(s + k) = 1.25$

Individuals will engage in the proscribed activity if and only if the benefits they derive are greater than their minimum expected losses, i.e., iff $b > \min[p(e, z^*)\alpha(s + k) + z^*, p(e,0)\alpha s]$. Thus, deterrence is determined by:

\[(12) \quad \tilde{b} = \min[p(e, z^*)\alpha(s + k) + z^*, p(e,0)\alpha s] .\]

The social problem lies in choosing the enforcement expenditures $e$, the punishment for the underlying offense $s (\leq s_{\text{max}})$, and the punishment for avoidance $k$, to maximize social welfare:

\[
W = \begin{cases} 
\int_{-\infty}^{\infty} [b - h - p(e, 0)(a + \gamma)s]f(b)db - e \\
\text{if } p(e,0)\alpha s < \min_{z}[z + p(e, z)\alpha(s + k)], \text{ or } \\
\int_{-\infty}^{\infty} [b - h - z - p(e, z)(\alpha + \gamma)(s + k)]f(b)db - e \\
\text{if } \min_{z}[z + p(e, z)\alpha(s + k)] \leq p(e,0)\alpha s .
\end{cases}
\]

\[25\] For convenience, it is assumed that offenders engage in avoidance even if $p(e, z^*)\alpha(s + k) + z^* = p(e,0)\alpha s$. 

http://law.bepress.com/taulwps/art101
4.3. Punishing Avoidance is Socially Desirable

Let us first show the social desirability of punishing avoidance. Suppose that the optimal enforcement scheme involves no punishment for avoidance at all, i.e., \( k^* = 0 \).

The social problem regarding \( s \) and \( e \) reduces to the problem analyzed in Section 3.2. The optimal solution is characterized by \( s_{\text{max}} \) and \( e_A^* \), satisfying condition (9).

This enforcement scheme, however, is non-optimal because punishment for avoidance could be slightly increased and enforcement expenditures appropriately reduced so as to achieve the same level of deterrence with lower enforcement and public punishment costs and without inducing zero avoidance. The reasoning is similar to the proof that the punishment for the underlying offense is maximal (Section 3.2). To illustrate, suppose that \( k^* = 0 \), \( e_A^* > 0 \), and \( s_{\text{max}} \). Increase \( k \) to \( \tilde{k} \) and lower \( e \) to \( \tilde{e} \) such that the expected sanction plus avoidance efforts are unaffected, i.e., \( p(\tilde{e}, \tilde{z}) \alpha (s_{\text{max}} + \tilde{k}) + \tilde{z} = p(e_A^*, z^*) \alpha s_{\text{max}} + z^* \), where \( z^* \) and \( \tilde{z} \) are the optimal avoidance efforts before and after the change in the enforcement scheme, respectively. If \( \tilde{k} \) is sufficiently small, the necessary decrease in \( e \) will also be small enough to guarantee that offenders will engage in avoidance, i.e., \( p(\tilde{e}, \tilde{z}) \alpha (s_{\text{max}} + \tilde{k}) + \tilde{z} \leq p(\tilde{e}, 0) \alpha s_{\text{max}} \). In fact, since punishment is increased, avoidance efforts will actually be higher, i.e., \( \tilde{z} > z^* \). By construction, deterrence and offenders' total utility are unaffected. Since punishment is increased but deterrence remains the same, enforcement expenditures are necessarily lower, i.e., \( \tilde{e} < e_A^* \). As noted, since punishment is increased, avoidance efforts are higher, \( \tilde{z} > z^* \). This means that the expected sanction imposed on offenders is lower,

\[
p(\tilde{e}, \tilde{z}) \alpha (s_{\text{max}} + \tilde{k}) < p(e_A^*, z^*) \alpha s_{\text{max}}.
\]

This, together with the fact that deterrence is

\[
\tilde{z} + p(\tilde{e}, \tilde{z}) \alpha (s_{\text{max}} + \tilde{k}) = z^* + p(e_A^*, z^*) \alpha s_{\text{max}} < p(e_A^*, 0) \alpha s_{\text{max}}.
\]

\[26\] This follows because: \( \tilde{z} + p(\tilde{e}, \tilde{z}) \alpha (s_{\text{max}} + \tilde{k}) = z^* + p(e_A^*, z^*) \alpha s_{\text{max}} < p(e_A^*, 0) \alpha s_{\text{max}} \).
unaffected, implies that the public costs of punishment are lower as well, i.e.,

\[1 - F(\tilde{b})]p(\tilde{c}, \tilde{z})\gamma(s_{\text{max}} + \tilde{k}) < [1 - F(\tilde{b})]p(e^*_A, z^*)s_{\text{max}}.\]

Thus, the enforcement and public costs of punishment are reduced, social welfare rises, and \( k^* = 0, \ e_A^* > 0, \) and \( s_{\text{max}} \) could not be optimal.

### 4.4 Punishing Avoidance Effectively (Discouraging Avoidance) is Socially Undesirable

The previous subsection shows that avoidance should be punished. This subsection shows that it should not be completely discouraged, however. In fact, the optimal enforcement scheme actually *encourages* avoidance.

Suppose that the sanction for avoidance could be set sufficiently high so that offenders prefer not to engage in avoidance at all. The social problem reduces to the standard problem of choosing the optimal enforcement scheme in a world without avoidance analyzed in Section 3.1. The optimal solution is characterized by \( s_{\text{max}} \) and \( e^*_N \), satisfying condition (8). Given this, the optimal \( k \) that induces no avoidance should satisfy

\[ p(e^*_N, 0)\alpha s_{\text{max}} < p(e^*_N, z^*)\alpha(s_{\text{max}} + k) + z^*, \]

where \( z^* \) satisfies

\[-p_z(e^*_N, z)\alpha(s_{\text{max}} + k) = 1.\]

In other words, the expected sanction that offenders face if they do not engage in avoidance, \( p(e^*_N, 0)\alpha s_{\text{max}} \), should be less than the expected sanction they face if they engage in avoidance, \( p(e^*_N, z^*)\alpha(s_{\text{max}} + k) \), plus their optimal avoidance efforts, \( z^* \). Note that \( k \) is not uniquely defined; any \( k \) which is greater than

\[
\frac{p(e^*_N, 0)\alpha s_{\text{max}} - z^* - p(e^*_N, z^*)\alpha s_{\text{max}}}{p(e^*_N, z^*)\alpha}
\]

will induce no avoidance. Deterrence is determined by \( p(e^*_N, 0)\alpha s_{\text{max}} \).

http://law.bepress.com/taulwps/art101
However, this enforcement scheme is non-optimal, because punishment for avoidance could be decreased so that avoidance is just induced. This would save the public costs of punishment without compromising deterrence. To illustrate this, suppose that \(s_{\text{max}}\), \(e_N^*\), and \(k^*\), inducing no avoidance. Now reduce \(k\) to \(\tilde{k}\) such that
\[
p(e_N^*, \tilde{z})\alpha(s_{\text{max}} + \tilde{k}) + \tilde{z} = p(e_N^*, 0)\alpha s_{\text{max}},
\]
where \(\tilde{z}\) satisfies \(-p_\gamma(e_N^*, z)\alpha(s_{\text{max}} + \tilde{k}) = 1\).

In other words, the expected sanction offenders face if they engage in avoidance, \(p(e_N^*, \tilde{z})\alpha(s_{\text{max}} + \tilde{k})\), plus their optimal avoidance efforts, \(\tilde{z}\), is equal to the expected sanction they face if they refrain from avoidance, \(p(e_N^*, 0)\alpha s_{\text{max}}\). Under these circumstances, offenders would engage in avoidance. By construction, offenders' total utility, deterrence, and enforcement expenditures are unaffected. However, since avoidance efforts are positive, \(\tilde{z} > 0\), the expected sanction offenders face is necessarily lower, i.e., \(p(e_N^*, \tilde{z})\alpha(s_{\text{max}} + \tilde{k}) < p(e_N^*, 0)\alpha s_{\text{max}}\). This, coupled with the fact that deterrence remains the same, implies that the public costs of punishment are lower, i.e., \([1 - F(\tilde{b})]p(e_N^*, \tilde{z})\gamma(s_{\text{max}} + \tilde{k}) < [1 - F(\tilde{b})]p(e_N^*, 0)\gamma s_{\text{max}}\). Thus, social welfare rises and \(s_{\text{max}}\), \(e_N^*\), and \(k^*\) that induce no avoidance could not be optimal.

Thus, the optimal enforcement scheme is not to leave avoidance alone, but rather to punish it. However, the optimal punishment for avoidance should not discourage it, but rather should induce more avoidance. The explanation is as follows. Punishing avoidance ineffectively induces more avoidance, but it also increases the costs of engaging in the underlying offense. Therefore, it reduces or even eliminates the social costs of avoidance in terms of increasing the enforcement costs of achieving

---

27 The existence of \(\tilde{k}\) is guaranteed since the minimum expected losses function
\[
L = \min\{z + p(e, z)\alpha(s_{\text{max}} + k)\}
\]
is continuous in \(k\), and since for \(k = 0\),
\[
p(e_N^*, \tilde{z})\alpha s_{\text{max}} + \tilde{z}^* < p(e_N^*, 0)\alpha s_{\text{max}},
\]
where \(\tilde{z}^*\) here satisfies \(-p_\gamma(e_N^*, \tilde{z})\alpha s_{\text{max}} = 1\).
deterrence, while at the same time increasing its social benefits by allowing a greater fraction of the expected sanction to be substituted with avoidance.  

Indeed, the optimal enforcement scheme is characterized by setting $s^* = s_{\text{max}}$, while $k^*$ and $e^*$ should satisfy (the formal proof is relegated to the appendix):

\[
(14) \quad (h + p(e^*, z^*)\gamma(s_{\text{max}} + k^*)) \frac{dF(\tilde{b})}{de} = 1 + [1 - F(\tilde{b})]p_e(e^*, z^*)(\alpha + \gamma)(s_{\text{max}} + k^*)
\]

\[
(15) \quad k^* = \frac{p(e^*, 0)\alpha s_{\text{max}} - z^* - p(e^*, z^*)\alpha s_{\text{max}}}{p(e^*, z^*)\alpha},
\]

where $\tilde{b} = z^* + p(e^*, z^*)\alpha(s_{\text{max}} + k^*)$ and $z^*$ satisfies $-p_e(e^*, z)\alpha(s_{\text{max}} + k^*) = 1$.

5. DISCUSSION AND EXTENSIONS

This paper argues that avoidance efforts to escape punishment may be socially desirable if the social cost of punishment is above and beyond the costs imposed on the offenders. Avoidance efforts are socially costly because they waste real resources.

28 The idea that avoidance should be controlled but should not be completely discouraged applies to other methods of controlling avoidance. To illustrate the latter claim, suppose that the private costs of avoidance $z$ could be increased by imposing a unit tax $\delta$ on the avoidance measure. Then, even though avoidance can be completely discouraged by setting $\delta$ sufficiently high such that $p(e^*, 0)\alpha s_{\text{max}} - z^* - p(e^*, z^*)\alpha s_{\text{max}} + z^* + \tilde{b}$, this will not be optimal. To see this, suppose that $\delta$ satisfies $p(e^*, 0)\alpha s_{\text{max}} = p(e^*, z^*)\alpha s_{\text{max}} + z^* + \tilde{b}$. Then avoidance will not be discouraged, but offenders' total utility and deterrence will not change. Social welfare will increase because the public costs of punishment will be lower. Similarly, suppose that the productivity of avoidance could be reduced at no cost, which can be represented by assuming that the minimum expected losses takes the form: $L = \min[p(e, \beta z)\alpha + z]$, where $\beta$ is a productivity control parameter taking values between 0 and 1. If $\beta = 1$, avoidance productivity is not affected at all, while if $\beta = 0$, avoidance productivity is eliminated. Observe that if $\beta = 0$ avoidance will be completely discouraged. However, this is not optimal. To see this, set $\beta$ such that $p(e^*, 0)\alpha s_{\text{max}} = p(e^*, \beta z)\alpha s_{\text{max}} + \tilde{z}$. where $\tilde{z}$ satisfies $-p_e(e, \beta \tilde{z})\alpha s_{\text{max}} = 1$. Then, avoidance will not be completely discouraged, but offenders' total utility and deterrence will not change. Social welfare, however, will rise, because the public costs of punishment will be saved.

29 The analysis assumes that $k$ is fixed. If $k$ could vary with the level of avoidance, it would be socially desirable to do so. However, contrary to common practice and intuition, it would be optimal to design $k(z)$ such that $k'(z) < 0$. That is, the marginal punishment for avoidance should decrease with avoidance. This will allow for elimination of the deterrence dilution effect of avoidance, but it will have the advantage of inducing even more avoidance.
and reduce the expected punishment, thereby increasing the enforcement costs of achieving deterrence. However, since avoidance efforts are also costly to offenders, they may substitute for socially costlier sanctions, such as imprisonment, thereby saving punishment costs. This benefit of avoidance may outweigh its social costs, suggesting that avoidance should not necessarily be discouraged. This, however, does not imply that punishing avoidance or otherwise increasing its private costs or decreasing its productivity is socially undesirable. Rather, it suggests that such measures, when employed, should discourage avoidance as little as possible or even encourage it. This way, the social costs of avoidance can be reduced, while its social benefits are maintained or even enhanced. Indeed, the possibility to effectively punish avoidance makes avoidance and its inducement always socially desirable.

Several issues that may have implications for the conclusions of this paper are worth discussing.


It was assumed that the illicit benefits offenders derive from the proscribed activity as well as their avoidance efforts and private suffering from punishment were taken into account in the social welfare calculus. While this is probably valid for minor or regulatory offenses, some may claim that for malicious crimes such as theft, burglary or murder, these benefits and perhaps costs should have little or no weight in calculating social welfare. Nevertheless, the social desirability of avoidance is not diminished and may even be enhanced by taking into account such views. If offenders' illicit benefits from the proscribed activity were not given any weight in the social welfare function, complete deterrence would generally not be optimal because enforcement and punishment are socially costly. Avoidance could still improve social welfare since it can substitute for imprisonment. Indeed, what is really important is
that avoidance should be very costly to offenders on the one hand, but not very effective in reducing deterrence on the other hand. Moreover, if avoidance efforts were not given any weight in the social welfare function, but the private suffering of imprisonment was, as many argue is appropriate, the social desirability of avoidance would merely strengthen because the social costs of avoidance would actually decrease. Thus, the conclusions of this paper are relevant for a large set of criminal activities, including malicious crimes.

5.2. *Types of Avoidance - Other Negative/Positive Externalities of Avoidance*

It was assumed that avoidance is socially costly in terms of utilizing real resources and increasing the enforcement costs of obtaining deterrence but may generate social benefits by substituting for imprisonment. It was also assumed that avoidance creates no other social costs or benefits (i.e., it generates no additional externalities). This captures a wide variety of avoidance measures offenders utilize such as concealment or destruction of incriminating evidence (disposing of a gun or other weapon, destruction of a car, paper shredding, evidence non-creation), lying, hiding, litigating, and so on. To clarify, offenders who avoid apprehension and punishment by hiding from the police incur costs upon themselves and decrease the probability of punishment, but presumably do not impose additional costs or benefits. Similarly, offenders who lie or engage in zealous litigation incur costs while increasing the costs of their convictions, but presumably do not generate additional social costs and benefits. Likewise, generally speaking, offenders who conceal or destroy incriminating evidence do not cause other costs and benefits than those formalized in the model.
However, certain avoidance measures may cause additional negative or positive externalities of various types and magnitudes than those envisioned in the model that should be taken into account and may dramatically affect the results. For example, if avoidance takes the form of *killing or threatening a witness*, the social costs should definitely include such grave harm and would clearly render it socially undesirable. Similarly, if avoidance (e.g., lying) involves efforts to *incriminate innocent individuals*, the social costs associated with this, including the private and public costs of punishing the innocent, should be calculated too and would render avoidance socially undesirable. In addition, if fleeing the scene of a crime involves dangerous driving (during a car chase) which puts other individuals at risk, the social costs of avoidance should incorporate these social costs as well. Cooking the books in financial transactions may significantly distort the ability of financial markets to allocate risk; these social costs should be also calculated.\(^{30}\) On the other hand, avoidance may create additional positive externalities that should also be taken into consideration. For example, offenders who attempt to avoid punishment for one offense may choose to refrain from committing other inefficient offenses. In addition, if offenders pay witnesses (in order to silence them or to influence their testimony) instead of killing or threatening them, then presumably this reduces the social costs of avoidance and should be counted as a social benefit. If avoidance takes the form of using seal-proof containers, which reduce the probability of detecting illegal moving of toxic waste, it also reduces the possibility of leakage and pollution, thereby generating a social benefit that should be taken into account.\(^{31}\) Thus, while the

\(^{30}\) I thank one of the referees for the last two examples. Observe, however, that the social undesirability of avoidance in all of the above examples stems from the additional social costs of avoidance and not because it is an avoidance measure per se.

\(^ {31}\) This example is taken from Stanley (1995).
analysis and conclusions of this paper apply to a wide range of avoidance measures, it is sensitive to certain avoidance measures that cause additional negative externalities.

5.3 Bribery

A special form of avoidance which deserves a separate analysis is bribery. Bribery is generally defined as a payment made from the offenders to law enforcement agents (police officers, prosecutors, or judges) in order to get better than fair treatment, i.e., to reduce or eliminate the probability or magnitude of punishment (Ayers, 1997). As with other avoidance measures, the common wisdom in the literature is that bribery is generally socially undesirable because it dilutes deterrence (Polinsky and Shavell, 2001, 2007).\(^{32}\) Apparently, if punishment for the underlying offense took the form of imprisonment, bribery would confer a social benefit because it could substitute for a socially costly punishment. This social benefit may outweigh the social costs, thereby increasing social welfare. However, there are reasons to believe that bribery is a socially undesirable form of avoidance. First, since bribes take a monetary form which is paid to a government agent, there is no reason to prefer them over fines, particularly if the government can be compensated for the lower expected revenues (fines) by reducing the wages of law enforcement agents. In other words, the justifications for not using fines as a punishment instrument (e.g., offenders are wealth-constrained or the government should not benefit from punishment) also justify that government agents will not take bribes. Second, bribery is probably an avoidance measure that causes other social costs (i.e., negative externalities). In

\(^{32}\) Polinsky and Shavell (2001) devote an entire section to discussing why corruption, including bribery, is socially undesirable and why it may be optimal to control it. Polinsky and Shavell (2001, p.3) restrict their analysis to fines, but point out that the general nature of their conclusion would not change if punishment took the form of imprisonment. However, as this paper explains, since the optimal enforcement scheme with imprisonment may be characterized by over-deterrence, the fact that bribes reduce deterrence should sometimes be counted as a social benefit. See also Polinsky and Shavell (2007) and the related literature.
particular, law enforcement agents who take bribes may also tend to engage in extortion or framing of innocents individuals (Ayers, 1997), which are clearly socially undesirable. Therefore, these additional, significant social costs of extortion should be taken into account in evaluating bribery and would render it socially undesirable. Furthermore, even if bribery is assumed to be a socially desirable form of avoidance, it does not follow that it should be legalized. On the contrary, as is explained in Section 4, avoidance efforts should be punished (even if they should not be completely discouraged). In any event, the special characteristics of bribery and its interactions with extortion call for a more elaborate analysis.

5.4. Screening Offenders

The benefits offenders derive from engaging in the proscribed activity were implicitly assumed to be private information not known to the social planner. This implied that the enforcement scheme and, in particular, sanctions, could not be tailored to the specific offender. If offenders could be perfectly screened at no cost, then the result would be markedly different. As Malik (1990) shows, with perfect screening, there should be no fines for offenders engaging in socially efficient crimes, in the sense that the benefits exceed the harm. This would save wasteful avoidance efforts altogether. This argument is only reinforced if the sanctions are themselves socially costly, since with costless and perfect screening, the costs of imposing sanctions are also saved.33 All in all, screening offenders, even if costly or imperfect, may be socially beneficial under the present model.

33 See Shavell (1987). Indeed, the argument for screening offenders also holds for undesirable but undeterrable activities (in the sense that the benefits to offenders are greater than any conceivable expected punishment). Screening offenders would save punishment costs (as in Shavell (1987)) as well as avoidance efforts. Consider, for example, the act of killing in the heat of passion. If screening took place and no punishment was imposed, the offender who by assumption could not have been deterred would not engage in costly and wasteful avoidance efforts.
5.5. Technological Relationship of Enforcement and Avoidance:

Enforcement expenditures and avoidance efforts were assumed to be independent in the sense that the level of one factor had no effect on the marginal effectiveness of the other (formally, $p_e(e,z) = 0$). This implied that punishment, but not enforcement expenditures, affected avoidance. More realistically, however, avoidance and enforcement efforts may be complements or substitutes in the production of detection. This may affect the results regarding the optimality of maximal sanctions.

For example, as Langlais (2006) shows, if the elasticity of avoidance efforts with respect to fines is equal to or less than that with respect to enforcement efforts, then, in contrast to Malik’s (1990) argument, the maximum fines are optimal in the presence of avoidance. The explanation is that although more severe punishment increases wasteful avoidance, a proportionate decrease in enforcement efforts reduces avoidance even more. Analogous analysis suggests that imprisonment need not be maximal.

However, the fundamental results of this paper: (1) that avoidance may be socially desirable, and (2) that avoidance is socially desirable if it can be effectively punished, are unaltered, because as can be verified, the arguments and proofs of these results

---

34 See, for example, Sanchirico (2006), Langlais (2006), and Nussim and Tabbach (2006).

35 The proof can be sketched as follows. Suppose that $s_{\text{max}}$ and $e^*$ are optimal such that deterrence is determined by $z^* + p(e^*, z^*)\alpha s_{\text{max}}$. Suppose now that imprisonment is slightly reduced to $\tilde{s}$ and enforcement efforts are slightly increased to $\tilde{e}$ so that deterrence is maintained; that is, $z^* + p(e^*, z^*)\alpha s_{\text{max}} = \tilde{s} + p(\tilde{e}, \tilde{z})\alpha \tilde{s}$, where $\tilde{z}$ is the optimal avoidance effort given the change in enforcement scheme. If enforcement and avoidance efforts are complementary, avoidance may be increased or decreased depending on its elasticity with respect to enforcement efforts relative to imprisonment, that is, $\frac{\partial z}{\partial e} > \frac{\partial z}{\partial f}$. If avoidance efforts are increased, that is, $\tilde{z} > z^*$, then the public costs of imprisonment per offender are reduced, $p(\tilde{e}, \tilde{z})\tilde{s} < p(e^*, z^*)s_{\text{max}}$, possibly by more than the increase in enforcement efforts. Consequently, social welfare would be increased, and $s_{\text{max}}$ and $e^*$ would not be optimal.
were not conditional in any way on the independence of avoidance and enforcement efforts.

5.6. Imprisonment and Fines:

Throughout this analysis, it was assumed that imprisonment is the sole form of punishment. More realistically, offenses are punished by both imprisonment and fines. This has an interesting implication regarding Malik's (1990) argument against maximal fines. As is well known, if optimal enforcement requires using both fines and imprisonment, then fines should be set to the maximum possible level, while imprisonment should generally not be maximal (Polinsky and Shavell, 1984). This is simply because fines, as a socially costless means of punishment, should be used to the limit before employing socially costly sanctions. This logic is valid regardless of avoidance, so Malik's argument is inapplicable. Since avoidance efforts depend on the total rather than the mix of punishment, changing the mix while keeping the total punishment constant does not alter avoidance at all. Thus, substituting costly with costless sanctions is socially desirable.\(^{36}\)

In any event, the use of both fines and imprisonment may affect the social desirability of punishment avoidance. This is so because, as shown, avoidance may be socially desirable if and only if the social costs of punishment are greater than the costs incurred by offenders, or equivalently, if and only if the public costs of

\(^{36}\)To illustrate, suppose that \(f^\ast < f_{\text{max}}, s^\ast > 0\) and \(e^\ast > 0\) so deterrence is determined by \(p(e^\ast, z^\ast)(f^\ast + \alpha s^\ast) + z^\ast\). Raise \(f\) to \(f_{\text{max}}\) and lower \(s\) to \(\hat{s}\) such that total punishment remains the same, i.e., \(f^\ast + \alpha s^\ast = f_{\text{max}} + \alpha \hat{s}\). Since avoidance depends on total punishment, there is no change in avoidance efforts. Consequently, deterrence and the offenders' total utility are unaffected. The only effect on social welfare is the reduced public costs of punishment since \(\hat{s} < s^\ast\) and the increased governmental revenues from fines since \(f_{\text{max}} > f^\ast\), which effectively represents the decrease in the private costs of punishment \(\alpha s < \alpha s^\ast\). Since social welfare is increased, \(f^\ast < f_{\text{max}}, s^\ast > 0\), and \(e^\ast\) cannot be optimal.
punishment are positive (see Section 3.3). However, if both fines and imprisonment are used, the satisfaction of these conditions is not guaranteed. Rather, it depends on the magnitude of the optimal fine $f_{\text{max}}$ relative to the optimal imprisonment term $s^*$ in a world without avoidance. Avoidance efforts may be socially desirable if and only if the total social costs of punishment are greater than the costs incurred by offenders, i.e., iff $(\alpha + \gamma)s^* > (\alpha s^* + f_{\text{max}})$, or equivalently, if and only if the total public costs of punishment are positive, i.e., iff $\gamma s^* - f_{\text{max}} > 0$. Otherwise, avoidance efforts are definitely socially undesirable.

**APPENDIX:**

The social problem as stated in (13) can be written as follows: choose $e$, $s$ and $k$ to maximise $[\max W_N, \max W_A]$ where:

(1A)  
\[
W_N = \int_{b=p(e,0)\alpha s}^{\infty} [b - h - p(e,0)(\alpha + \gamma)s] f(b)db - e;
\]

S.t.  
$s - s_{\text{max}} \leq 0$

\[
p(e,0)s - \tilde{z} - p(e,\tilde{z})\alpha(s + k) < 0
\]

$\tilde{z}$ satisfies $-p_\tilde{z}(e, z)\alpha(s + k) = 1$

(1B)  
\[
W_A = \int_{b=p(e,z^*)\alpha(s+k)}^{\infty} [b - h - z - p(e,z)(\alpha + \gamma)(s + k)] f(b)db - e;
\]

S.t.  
$s - s_{\text{max}} \leq 0$

\[
z^* + p(e, z^*)\alpha(s + k) - p(e,0)s \leq 0
\]

$z^*$ satisfies $-p_{z^*}(e, z)\alpha(s + k) = 1$
The maximization of $W_N$: Since $k$ is not bounded, it could be set sufficiently large to assure that the constraint is satisfied for any positive $e$ and $s$. Therefore, this constraint can be ignored at first. Define the Lagrangian function as:

$$(1A') \quad LG_N = W_N + \eta (s_{\text{max}} - s),$$

The optimal $e$ and $s$ should satisfy the Kuhn-Tucker first-order conditions (throughout second-order conditions are assumed to be satisfied):

$$(2A') \quad \frac{\partial LG_N}{\partial s} = (h + p(e,0)\gamma s)p(e,0)\alpha f(\tilde{b}) - [1 - F(\tilde{b})]p(e,0)(\alpha + \gamma) - \eta = 0$$

$$(3A') \quad \frac{\partial LG_N}{\partial e} = (h + p(e,0)\gamma s)p_e(e,0)\alpha s f(\tilde{b}) - [1 - F(\tilde{b})]p_e(e,0)(\alpha + \gamma)s - 1 = 0$$

$$(4A') \quad \frac{\partial LG_N}{\partial \eta} = s_{\text{max}} - s \geq 0 \quad \eta \geq 0 \quad \text{and} \quad \eta (s_{\text{max}} - s) = 0$$

Suppose now that $\eta = 0$. This implies (from (2A')) that:

$$(h + p(e,0)\gamma s)\alpha f(\tilde{b}) = [1 - F(\tilde{b})](\alpha + \gamma),$$

but then:

$$\frac{\partial LG_N}{\partial e} = (h + p(e,0)\gamma s)p_e(e,0)\alpha s f(\tilde{b}) - [1 - F(\tilde{b})]p_e(e,0)(\alpha + \gamma)s - 1 = -1 < 0.$$

This contradicts (3A'). Therefore, $\eta > 0$ and $s^*_N = s_{\text{max}}$.

Now $e^*_N$ should satisfy (3A'), and $k^*_N$, which is not uniquely determined, should satisfy the inequality constraint: $k > \frac{p(e^*_N,0)\alpha s_{\text{max}} - \bar{z} - p(e^*_N,\bar{z})\alpha s_{\text{max}}}{p(e^*_N,\bar{z})\alpha}$.

Denote the maximum social welfare $W_N^*$.

The maximization of $W_A$:

Define the Lagrangian function as:

$$(1B') \quad LG_A = W_A + \lambda (s_{\text{max}} - s) + \mu(p(e,0)\alpha s - z^* - p(e, z^*)\alpha(s + k))$$

The optimal $e$, $s$ and $k$ should satisfy the Kuhn-Tucker first-order conditions:

$$(2B') \quad \frac{\partial LG_A}{\partial s} = (h + p(e, z^*)\gamma(s + k)) p(e, z^*)\alpha f(\tilde{b}) +$$
\[-[1-F(\hat{b})][p_z(e,z\*)\gamma(s+k)\frac{dz\*}{ds} + p(e,z\*)(\alpha + \gamma) +
- \lambda + \mu(p(e,0)\alpha - p(e,z\*)\alpha) = 0\]

(3B') \[\frac{\partial LG_A}{\partial k} = (h + p(e,z\*)\gamma(s+k))p(e,z\*)af(\hat{b}) + \]
\[-[1-F(\hat{b})][p_z(e,z\*)\gamma(s+k)\frac{dz\*}{dk} + p(e,z\*)(\alpha + \gamma) - \mu p(e,z\*)\alpha = 0\]

(4B') \[\frac{\partial LG_A}{\partial e} = (h + p(e,z\*)\gamma(s+k))p_z(e,z\*)\alpha(s+k)f(\hat{b}) + \]
\[-[1-F(\hat{b})]p_z(e,z\*)(\alpha + \gamma)(s+k) - 1 = 0\]

(5B') \[\frac{\partial LG_A}{\partial \lambda} = s_{\text{max}} - s \geq 0 \quad \lambda \geq 0 \quad \text{and} \quad \lambda(s_{\text{max}} - s) = 0\]

(6B') \[\frac{\partial LG_A}{\partial \mu} = p(e,0)\alpha s - z\* - p(e,z\*)\alpha(s+k) \geq 0 \quad \mu \geq 0 \quad \text{and} \quad \frac{\partial LG_A}{\partial \mu} \mu = 0 .\]

Suppose that \(\lambda = 0\) and \(\mu = 0\).

This implies (either from (2B') or (3B'), since \(\frac{dz\*}{ds} = \frac{dz\*}{dk} = -\frac{p_z(e,z)(s+k)}{p_z(e,z)} > 0\))

that:

\[(h + p(e,z\*)\gamma(s+k))af(\hat{b}) = [1-F(\hat{b})][\frac{p_z(e,z\*)}{p(e,z\* )}\frac{dz\*}{ds}\gamma(s+k) + (\alpha + \gamma)\]

However,

\[\frac{\partial LG_A}{\partial e} = [1-F(\hat{b})]\frac{p_z(e,z\*)p_z(e,z\*)\gamma(s+k)^2\frac{dz\*}{ds}}{p(e,z\*)} - 1 < 0\], since \(p_z(e,z\*) < 0\),

which contradicts condition (4B'). Therefore \(\lambda = \mu = 0\) is not possible.

Suppose that \(\lambda > 0\) and \(\mu = 0\).

This implies that conditions (2B') and (3B') are inconsistent since \(\frac{dz\*}{ds} = \frac{dz\*}{dk}\).

Therefore, \(\lambda > 0\) and \(\mu = 0\) is also not possible.

Suppose that \(\lambda = 0\) and \(\mu > 0\).
This implies that conditions (2B’) and (3B’) are again inconsistent, since \( \frac{dz^*}{ds} = \frac{dz^*}{dk} \)
and \( p(e,0)\alpha > 0 \). Therefore, \( \lambda = 0 \) and \( \mu > 0 \) is also not possible.

Thus, \( \lambda > 0 \) and \( \mu > 0 \), which means that \( s^*_A = s_{\text{max}} \), and that \( e^*_A \) and \( k^*_N \) should satisfy:

\[
\text{(7B')} \quad (h + p(e^*_A, z^*)\gamma(s_{\text{max}} + k^*_N))p_e(e^*_A, z^*)\alpha(s_{\text{max}} + k^*_N)f(\hat{b}) = \\
= [1 - F(\hat{b})]p_e(e^*_A, z^*)(\alpha + \gamma)(s_{\text{max}} + k^*_N) + 1 \quad \text{and} \\
\text{(8B')} \quad k^*_N = \frac{p(e^*_A, 0)\alpha s_{\text{max}} - z^* - p(e^*_A, z^*)\alpha s_{\text{max}}}{p(e^*_A, z^*)\alpha},
\]

where \( z^* \) satisfy \( -p_z(e^*_A, z^*)\alpha(s_{\text{max}} + k^*_N) = 1 \).

Denote the maximum social welfare \( W^*_A \).

Now let us show that \( W^*_A > W^*_N \).

Choose the following enforcement scheme: \( s_{\text{max}} \), \( e^*_N \) and \( \hat{k} \) such that

\[
\hat{z} + p(e^*_N, \hat{z})\alpha(s_{\text{max}} + \hat{k}) = p(e^*_N, 0)\alpha s_{\text{max}},
\]

where \( \hat{z} \) satisfy \( -p_z(e^*_N, z)\alpha(s_{\text{max}} + \hat{k}) \), which gives rise to \( \hat{W}_A(< W^*_A) \). Subtract \( W^*_N \) from \( \hat{W}_A \), and obtain:

\[
[1 - F(\hat{b})][p(e^*_N, 0)\alpha s_{\text{max}} - p(e^*_N, \hat{z})\gamma(s_{\text{max}} + \hat{k})],
\]

which is positive since

\[
p(e^*_N, \hat{z})(s_{\text{max}} + \hat{k}) < p(e^*_N, 0)s_{\text{max}}.
\]

Therefore, \( W^*_A > \hat{W}_A > W^*_N \), and the optimal solution is characterized by \( s^*_A = s_{\text{max}} \), and \( e^*_A \) and \( k^*_N \) satisfying (7B’) and (8B’).

REFERENCES


