Avoidance, Errors and the Optimal Standard of Proof

Murat Mungan∗
Antonin Scalia Law School at George Mason University

Andrew Samuel†
Loyola University Maryland

April 19, 2018

Abstract

Criminals often engage in costly avoidance to lower their probability of being detected and sanctioned. Such avoidance, in turn, affects the optimal enforcement policy. This paper studies the impact of avoidance on a specific type of enforcement policy - the standard of proof. We show that when avoidance is possible the optimal standard of proof is weaker than preponderance of the evidence. This stands in contrast to much of the literature, which shows that non-deterrence costs usually cause the standard of proof to be stronger than preponderance. Our results suggest that it is important not to ignore criminals’ secondary behavior when determining the optimal standard of proof.

1 Introduction

Legal systems are imperfect. They are incapable of imposing liability for every wrong-doing (type-2 errors), and similarly, they may impose liability on individuals who have not violated the law (type-1 errors). Many legal systems endeavor to reduce the frequency of both types of errors, presumably because judicial errors are detrimental to both fairness and deterrence. However, it is either costly to reduce the frequency of a specific type of error, or doing so leads to an increase in the frequency of the other type of error. As noted by courts,

∗Associate Professor, Antonin Scalia Law School, George Mason University, Fairfax, VA.
†Associate Professor of Economics, Sellinger School of Business, Loyola University Maryland, Baltimore, MD

1See, e.g., Png (1986), which makes the observation that both types of errors are detrimental to deterrence because they reduce the expected opportunity cost of committing a crime. There is a sizable literature that builds on this observation, which is reviewed in Lando and Mungan (2018).

standards of proof (SoP) used in trials can be regarded as tools that address the latter types of trade-offs that courts inevitably face.

In the American legal system, stronger standards are used in criminal trials than in civil trials. This situation has attracted the attention of many law and economics scholars, who have shown a particular interest in exploring the economic rationale behind this distinction. The result was a focus on factors that could cause the optimal standard to be different, and in many cases, stronger, than the standard frequently used in civil cases, namely preponderance of the evidence. Much of the work in this literature has focused on the asymmetric effects of wrongful convictions and wrongful acquittals. For instance, Kaplow (2011) and Mungan (2011) both note that increases in the probability of wrongful conviction can cause chilling of desirable behavior, whereas Rizzolli and Saraceno (2013) focuses on increases in imprisonment costs caused by wrongful convictions.

Interestingly, the (non-deterrence) costs considered in these articles share an important property; they can be reduced by strengthening the SoP. In Kaplow (2011) and Mungan (2011), this is because stronger standards lead to fewer wrongful convictions, which, in turn, causes a reduction in chilling effects. In Rizzolli and Saraceno (2013), stronger standards lead to fewer convictions, and, this causes a reduction in imprisonment costs. In this article, we identify a cost that is unlikely to have this property, namely investments that criminals make to mimic the behavior of innocent individuals to reduce their likelihood of being convicted. The expected value a criminal reaps from mimicking an innocent individual is naturally decreasing in the probability of wrongful convictions. There is not much reason to attempt to appear innocent, if innocents are convicted almost as frequently as guilty individuals. We refer to these costs as avoidance costs (see, e.g. Malik (1990)), because they allow the criminal to avoid punishment more frequently.

To isolate and illustrate the effects of avoidance costs on the optimal SoP, we focus on a case where SoP’s only affect deterrence and avoidance activity. We show that the presence of avoidance costs push the optimal SoP to be weaker than preponderance of the evidence. This observation implies that abstractions made in prior analyses of optimal the SoP may have serious effects on normative conclusions, and that potential impacts on criminals’ secondary behavior ought to be taken into account when discussing the optimality of various standards of proof.

The remainder of the article is organized as follows. In the next section, we provide a brief review of the related literature. In section 3 we present a simple model of crime and deterrence, study the equilibrium level of avoidance and its relationship to the standard of proof, and present our main welfare result.

---


4 For example, a firm may choose to hide traces of pollution or waste that it produced. See Sanchirico (2010) for a discussion of the types of avoidance measures that criminals may undertake.
2 Literature Review

This paper is related to two strands within the literature on enforcement; namely one concerning “the standard of proof” and another concerning “avoidance and deterrence.” In the first strand of the literature, there is a considerable focus on identifying rationales for the optimality of strong standards of proofs in criminal trials. Specifically, Rizzolli and Saraceno (2013) shows that when the costs of punishment are positive, then wrongful convictions are more costly to society than wrongful acquittals. Similarly, Mungan (2011) shows that if the innocent undertake precautionary activities to avoid being wrongfully convicted, then a utilitarian planner will choose a standard of proof that is higher than preponderance. Kaplow (2011) identifies a similar dynamic, and notes that lowering the standard of proof can cause a chilling of socially desirable behavior. Another function of the standard of proof is to alter the stigma that attaches to a defendant upon conviction: because a lower standard of proof provides a weaker signal about the actual guilt of a person, it can dilute the stigma imposed upon conviction (Fluet and Mungan (2017) and Mungan (2018)). Finally, most recently Garoupa (2018) has demonstrated that if law enforcers have punitive preferences, then high standards of proof may be necessary to correct for their bias towards punishment. In contrast to the common theme running through this literature, we focus on how the standard of proof may affect the non-criminal behavior of offenders, and show that such the incorporation of these costs has the opposite of the effect that has been identified in the literature.

A survey of the second strand of the literature that relates to our article, is provided by Sanchirico (2010). Most of this literature focuses on studying the optimal fine under avoidance, and finds that when offenders can invest in costly avoidance activities then the optimal fine is not maximal (Malik 1990). Intuitively, because avoidance is costly it serves as a substitute for the fine. Hence, if increasing the fine also raises the equilibrium level of avoidance, then the optimal fine may not be maximal. Langlais (2008) and Sanchirico (2010), however, show that when avoidance and monitoring efforts are complements in the sense that an increase in monitoring increases the marginal impact of avoidance, then the optimal fine is maximal. Dechenaux and Samuel (2014), build on this work to study the effect of announced versus surprise inspections within the context of avoidance. Surprise inspections are especially useful when criminals can engage in avoidance or hide evidence because it enables the auditor to catch them off guard. Dechenaux and Samuel, therefore, find that the optimal fine is maximal when there are surprise inspections, but not when they are announced.6

---

6Dechenaux and Samuel (2014) also show that compliance is not always increasing in the fine because the increase in avoidance as a result of the higher fine, can offset the deterrence
In contrast to this literature, our paper focuses on the interaction between avoidance and the standard of proof, which to our knowledge has yet to be studied. Here we find that the effect on avoidance of lowering the standard of proof is different from the effect of increasing the level of monitoring. This is because lowering the standard of proof has two effects on avoidance. First, weakening the standard of proof makes correct convictions more likely which increases the marginal productivity of avoidance. This encourages more avoidance and is identical to the affect of monitoring on avoidance that has previously been identified in the literature (e.g. Langlais 2008). However, there is a second effect, because avoidance also distorts the evidence so that a guilty individual’s evidence more closely mimics the innocent. When the standard of proof is weakened it makes wrongful convictions more likely, which discourages avoidance because avoidance makes it more likely that a criminal is perceived as an innocent (who is now more likely to be punished). Consequently, the equilibrium level of avoidance may rise or fall with the standard of evidence depending on which effect dominates. As we show in the next section, the interactions between these two effects cause the optimal standard of proof to be weaker than preponderance of the evidence.

3 Model

A continuum of potential offenders decide whether or not to commit an act which increases (expected) social harm by \( h \). The act confers a benefit, which differs across individuals, such that the density of individuals who derive benefit \( b \) is given by \( f(b) \) with support \( b \in [0, \infty) \). To deter the commission of the act, individuals are audited: a person who commits the act is audited with probability \( p \), and a person who has not committed the act is audited with probability \( q \leq p \). An audit triggers an investigation about whether the person committed the act. The investigation is imperfect, in that it can fail to detect wrongdoing (i.e. generate type-2 error), with probability \( 1 - \beta \), and it can detect wrongdoing where none exists (i.e. generate type-1 error), with probability \( \alpha \). Thus, a person who has committed the act is found liable with probability \( p\beta \) and a person who has not committed the act is found liable with probability \( q\alpha \). Liability triggers a sanction whose size is normalized to 1, and we assume that \( h > 1 \) to formalize the idea that sanctions under-deter.

People who commit the act can invest in avoidance activity, which reduces the probability of conviction. Therefore, we capture the effect of avoidance activity as a distortion in the amount of exculpatory evidence. Avoidance decisions must be made subsequent to committing the act, but, prior to being audited.\(^7\) Moreover, avoidance investments cannot lower an offender’s probability of being found guilty in an investigation to one that is lower than the probability of an innocent individual being found guilty during an investigation. The rationale

---

\(^7\)This corresponds to the case of “surprise inspections” identified in Dechenaux and Samuel (2014).
here is that the offender invests in mimicking an innocent individual, and even an innocent individual is found guilty with a probability of \( \alpha \), conditional on investigation. To formalize avoidance activities, we denote by \( a \) the investment in avoidance, such that \( \beta \) becomes a function of \( a \). We specify the precise relationship between \( \beta \) and \( a \) after we explain how the government can affect \( \beta \) as well as \( \alpha \).

### 3.1 Investigation Standards, \( \alpha \) and \( \beta \)

The government chooses an investigation standard that determines the requirements for inferring that the investigate has committed the act. This standard influences both \( \alpha \) and \( \beta \). To formalize this idea, suppose that each individual emits a noisy signal \( x \in [x, x] \) regarding his guilt where smaller values of \( x \) are more consistent with guilt. Suppose further that the probability with which a person emits various signals depends on whether he is guilty, and if he is, on how much he spends on avoidance. To formalize this, let \( g_0(x) \) and \( n(x) \) respectively denote the probability density functions over \( x \) for guilty individuals who have spent \( a \) on avoidance activities, and for innocent individuals. Let \( G_a \) and \( N \) denote the cumulative distribution functions associated with \( g_0 \) and \( n \). We assume that \( g_0(x)/n(x) \) satisfies the monotone likelihood ratio property (MLRP), i.e. that \( \frac{\partial (g_0(x)/n(x))}{\partial x} < 0 \) for all \( x \). Moreover, we assume that avoidance activities allow guilty individuals to mimic innocent individuals, i.e., we assume that:

\[
g_a(x) = \gamma(a)n(x) + (1 - \gamma(a))g_0(x), \tag{1}
\]

where \( \gamma(0) = 0, \lim_{a \to \infty} \gamma = 1, \gamma' > 0, \gamma'' < 0 \). Here \( g_0 \) can be thought of as describing the natural, or uninterfered, process through which evidence is ordinarily generated. By spending \( a > 0 \), an offender mimics innocent individuals.

The government chooses an investigation standard in the form of a threshold signal, \( x^* \), such that a person is found liable upon investigation only if he emits a signal \( x < x^* \).\(^8\) This implies that an innocent individual’s probability of being found liable upon investigation is

\[
\alpha(x^*) = \int_{x^*}^{x^*} n(x)dx = N(x^*), \tag{2}
\]

and a guilty person’s probability of being found liable upon investigation is:

\[
\beta(x^*, a) = \int_{x^*}^{x^*} g_a(x)dx \tag{3}
\]

which, by using (1) and (2) can be re-written as:

\[
\beta(x^*, a) = \gamma(a)\alpha(x^*) + (1 - \gamma(a))\beta_0(x^*) \tag{4}
\]

\(^8\)This threshold rule is efficient in the sense that it achieves any targeted level of type-2 error through the smallest type-1 error possible.
where
\[ \beta_0(x^*) \equiv \beta(x^*, 0). \] (5)

Since \( \alpha \) is an increasing function of \( x^* \), to simplify notation, we can assume that the government chooses \( \alpha \) which implies an \( x^* \) given by \( N^{-1}(\alpha) = x^* \). This choice of \( \alpha \), in turn, generates a unique \( \beta_0 \), because \( \beta_0 \) is also an increasing function of \( x^* \). In symbols, we may assume that the government chooses an \( \alpha \in [0, 1] \) which, in turn generates a \( \beta_0 \) such that:
\[ \beta_0(\alpha) = G_0(x^*(\alpha)) = G_0(N^{-1}(\alpha)). \] (6)

This convention allows the elimination of \( x^* \) as a variable that needs tracking, and allows \( \beta \) to be expressed as a function of \( \alpha \) and \( a \) as follows:
\[ \beta(\alpha, a) = \gamma(a)\alpha + (1 - \gamma(a))\beta_0(\alpha). \] (7)

Moreover, as noted in prior work,\(^9\) MLRP implies that \( \beta_0 \) has the following properties:
\[ \beta_0(0) = 0, \beta_0(1) = 1, \frac{d\beta_0}{d\alpha} > 0, \text{ and } \frac{d^2\beta_0}{d\alpha^2} < 0. \] (8)

Finally, it is worth noting that these properties imply the existence of a unique type-1 error that satisfies the following:
\[ \alpha^m \equiv \arg \max_{\alpha} \beta_0(\alpha) - \alpha. \] (9)

A property of \( \beta_0 \), which relates to \( \alpha^m \) and which we will use repeatedly later, is that
\[ \frac{d\beta_0(\alpha)}{d\alpha} \geq 1 \text{ iff } \alpha \leq \alpha^m. \] (10)

### 3.2 Preponderance of the Evidence Standard

#### 3.2.1 Without Avoidance Activity

In our description of the process through which individuals are found liable, we referred to an investigation process, without specifying how this process takes place. In general, this process can be in the form of a trial, arbitration, or any other form of dispute resolution. Prior articles which have conceived of this investigation process as a trial have articulated a preponderance of the evidence standard through which a person’s liability is determined. This standard corresponds to one where a person is found guilty, if he produces a signal, \( x \), which is more likely to be produced by a guilty individual than an innocent individual. In the prior literature guilty individuals produce signals through the density function which we have denoted \( g_0 \), because avoidance activity has not previously been incorporated in this framework. Thus, in the prior literature, because signals are produced only upon audit, a signal is more likely to

\(^9\)See, e.g. Fluet and Mungan (2017) and the references cited therein.
be produced by a guilty individual if \( \frac{p_{0}(x)}{n(x)} > 1 \).\(^{10}\) Due to MLRP, this standard would correspond to a threshold \( x^{P} \) such that \( \frac{p_{0}(x^{P})}{n(x^{P})} = \frac{q}{p} \). Noting that \( \frac{d\beta_{0}(\alpha)}{d\alpha} = \frac{g_{0}(x^{*}(\alpha))}{n(x^{*}(\alpha))} \), in the prior literature the preponderance of the evidence can be formulated as the government choosing an \( \alpha^{P} \) such that\(^{11}\)

\[
\frac{d\beta_{0}(\alpha^{P})}{d\alpha} = \frac{q}{p} \tag{11}
\]

It is worth noting that such a standard need not exist, because although MLRP guarantees that \( \frac{d\beta_{0}(\alpha)}{d\alpha} < 1 \) for all \( \alpha > \alpha^{m} \), the evidence generating process can be such that \( \frac{d\beta_{0}(\alpha)}{d\alpha} > \frac{q}{p} \) for all \( \alpha \). In these cases, preponderance of the evidence would not be a feasible standard, and \( \alpha = 1 \) would be as close as one can get to this standard. As will become apparent from the proceeding analysis, in such cases the analysis is relatively uninteresting, because \( \alpha = 1 \) becomes the deterrence maximizing as well as optimal standard. Thus, in the next subsection (section 3.3), we introduce an assumption that guarantees that preponderance of the evidence is a feasible standard.

### 3.2.2 With Avoidance Activity

The above discussion pertains to cases where avoidance activity is not possible. When individuals can invest in avoidance, preponderance of the evidence has to be re-interpreted. This is because, the equilibrium \( \beta \) through which signals are generated depends on offenders’ avoidance activity. Thus, given some avoidance activity, \( a \), the probability of type-1 error that corresponds to preponderance of the evidence can be defined as:

\[
\partial\beta(\alpha^{P}, a) = \frac{q}{p} \tag{12}
\]

In the proceeding analysis we will make frequent references to these standards, and therefore it is important to note the properties of this standard. Specifically, by differentiating \( \beta \) with respect to \( \alpha \), we can re-write (12) as follows:

\[
\frac{d\beta_{0}(\alpha^{P})}{d\alpha} = \frac{q - p\gamma(a)}{p - p\gamma(a)} \tag{13}
\]

Thus,

\[
\frac{d\alpha^{P}}{da} = -\frac{\gamma(a)(p - q)}{p^{2}\beta_{0}(1 - \gamma)^{2}} > 0 \tag{14}
\]

\(^{10}\)This approach (followed in Fluet and Demougin (2006)) formulates the standard of proof with reference only to priors, and does not incorporate the odds with which a randomly chosen person may be guilty, i.e. the crime rate. Kaplow (2011), on the other hand, argues that preponderance of the evidence ought to be formulated in terms of posterior probabilities. Since the objective of this article is to demonstrate that avoidance activity pushes the optimal SoP to be weaker than is suggested in the literature, we use the former type of formulation, which corresponds to that studied in most of the articles on the topic.

\(^{11}\)See, e.g. Demougin and Fluet (2006) and Fluet and Mungan (2017).
where the inequality follows from the facts that $\frac{d^2\beta}{d\alpha^2} < 0$ and $p > q$. Moreover, (13) implies that $\alpha^P(0) > \alpha^m$, since $\frac{q-p\gamma(0)}{p-P\gamma(0)} = \frac{q}{p} < 1$. The fact that $\frac{da^P}{da} > 0$ implies that $\alpha^P(a) > \alpha^m$ for all $a$. Thus, (14) and (13) together reveal the following.

**Observation 1:** The $\alpha$ that corresponds to the preponderance of the evidence standard (when feasible) is increasing in the amount of avoidance and $\alpha^P(a) > \alpha^m$ for all $a$.

The intuition behind observation 1 is that when offenders invest more in avoidance, they produce signals that better mimic innocent individuals, i.e. $g_a$ gets closer to $n$. Countering this effect requires a standard that better discriminates between offenders and non-offenders on the margin. This requires lowering the threshold evidence required for a finding of liability, because this widens the gap between the likelihood with which the threshold evidence is produced by innocent versus guilty individuals.

### 3.3 Criminal’s Avoidance Choice

The previous discussion explains the machinery, in the form of type-1 and type-2 errors faced by potential offenders, and also explains how the preponderance of the evidence standard depends on avoidance. However, an offenders’ choice of avoidance is naturally also affected by the standard chosen by the government. This can be formalized by analyzing individuals’ decisions, starting with the last decision they make, which relates to avoidance. An offender chooses his avoidance investment, $a$, after he chooses to commit the harmful act, to minimize the expected costs associated with his action. In particular, for any given $\alpha$, he minimizes:

$$p\beta(\alpha, a) + a$$

(15)

The first order condition for this minimization problem characterizes offenders’ best responses, $a^*(\alpha)$ as follows:

$$p\frac{\partial \beta(\alpha, a^*)}{\partial a} + 1 = 0$$

(16)

Using the implicit function theorem, we have that

$$\frac{da^*}{d\alpha} = \frac{-\frac{\partial^2 \beta}{\partial a^2}}{\frac{\partial^2 \beta}{\partial a^2} - \frac{\partial^2 \beta}{\partial a^2}} = \frac{\gamma_\alpha(1 - \frac{d\beta}{d\alpha})}{\gamma_\alpha(\beta_0 - \alpha)}$$

(17)

Since the denominator is negative, it follows that

$$\frac{da^*(\alpha)}{d\alpha} \geq 0 \text{ iff } \alpha \leq \alpha^m$$

(18)

since $\frac{d\beta}{d\alpha} \geq 1$ iff $\alpha \leq \alpha^m$ (see (9) and (10)). From this it follows that $a^*$ is quasi-concave in $\alpha$ and is maximized at $\alpha^m$. Differentiating $a^*$ once more with
This observation allows us to formalize the assumption that guarantees the feasibility of the preponderance of the evidence standard that corresponds to the equilibrium avoidance investment.

**Assumption 1**: \( \frac{\partial \beta(1, a^m)}{\partial a} < \frac{q}{p} \).

This assumption implies that \( \frac{\partial \beta(1, a^*(\alpha))}{\partial a} < \frac{q}{p} \) for all \( \alpha \in [0, 1] \), because \( \frac{\partial \beta(1, a)}{\partial a} = \gamma_a(1 - \frac{\partial \beta(1)}{\partial a}) > 0 \) for all \( a \). This, in turn, implies that for all possible \( a \in [0, a^*(\alpha)] \), there exists \( a^P < 1 \), such that \( \frac{\partial \beta(a^P, a)}{\partial a} = \frac{q}{p} \), which corresponds to the preponderance of the evidence standard.

### 3.4 Equilibrium Preponderance Standard v. Chosen Standard

As sections 3.2 and 3.3 reveal, the preponderance of the evidence standard depends on the investment in avoidance, and the equilibrium investment of avoidance, in turn, depends on the standard chosen by the government. Thus, a question that needs to be answered is whether there exists any standard which corresponds to preponderance of the evidence in equilibrium. The next proposition answers this question affirmatively, and makes important comparisons between the equilibrium preponderance standard and other standards that can potentially be chosen. The relationship between \( a^P(a) \) and \( a^*(\alpha) \) is depicted in figure 1, below, which can be used to follow the steps taken in the proof of proposition 1.

**Proposition 1** (i) There exists a unique \( \hat{\alpha} \in (0, 1) \) such that \( a^P(a^*(\hat{\alpha})) = \hat{\alpha} \).

(ii) All \( \alpha < \hat{\alpha} \) correspond to standards that are stricter than preponderance of the evidence in equilibrium (i.e. \( \alpha < a^P(a^*(\alpha)) \) for all \( \alpha < \hat{\alpha} \)). (iii) All \( \alpha > \hat{\alpha} \) correspond to standards that are weaker than preponderance of the evidence (i.e. \( \alpha > a^P(a^*(\alpha)) \) for all \( \alpha > \hat{\alpha} \)).

**Proof.** As noted in observation 1, \( a^m < a^P(a) \) for all \( a \in [0, a^*(\alpha)] \), thus \( a^P(a^*(\alpha^m)) - a^m > 0 \). Moreover, \( a^P(a^*(1)) - 1 = a^P(0) - 1 < 0 \). Thus, there exists \( \hat{\alpha} \) such that \( a^P(a^*(\hat{\alpha})) = \hat{\alpha} \). Moreover, \( \frac{\partial a^P(a^*(\alpha))}{\partial a} < 0 \) for all \( \alpha > a^m \) as implied by (14) and (18). Thus, \( \hat{\alpha} \) is the only \( \alpha \in (a^m, 1) \) for which this equality holds. Moreover, for all \( \alpha \leq a^m \), it follows that \( \alpha < a^P(a^*(\alpha)) \), since \( a^m < a^P(a) \) for all \( a \). Thus, \( \hat{\alpha} \) is the unique \( \alpha \) such that \( a^P(a^*(\hat{\alpha})) = \hat{\alpha} \).

(ii) and (iii) follow from the previously established facts that (a) \( \frac{\partial a^P(a^*(\alpha))}{\partial a} < 0 \) for all \( \alpha > a^m \), (b) \( \alpha > a^m \), and (c) \( a^m < a^P(a) \) for all \( a \). ■

---

\(^{12}\) The second derivative is given by: \( \frac{d^2a^*}{dx^2} = \frac{\gamma_a}{\gamma_a} \frac{d^2\beta(1)}{dx^2} (\beta_0 - \alpha) + (1 - \frac{d\beta_1}{dx}) \cdot \frac{d^2\beta_1}{dx^2} < 0 \).
Figure 1: Equilibrium preponderance. 

Proposition 1 reveals that there is a unique $\hat{\alpha} \in [0, 1]$ that corresponds to the preponderance of the evidence standard in equilibrium. $\hat{\alpha}$ plays an important role in the proceeding analysis, because it maximizes deterrence, and serves as
a benchmark to compare the welfare maximizing standard of proof.

3.5 Choice to Commit Crime

A potential offender commits crime if \( b - p\beta(\alpha, a^*(\alpha)) - a^* > -\alpha q \), since the left hand side captures the net expected benefit from crime, and the right hand side captures the net-expected cost from compliance. Thus, a person commits crime if his benefit is such that:

\[
b^* \equiv (p\beta(\alpha, a^*(\alpha)) - q\alpha) + a^*(\alpha) < b \tag{20}
\]

Using the envelop theorem, it follows that,

\[
\frac{db^*}{d\alpha} = p\frac{\partial \beta(\alpha, a^*(\alpha))}{\partial \alpha} - q \tag{21}
\]

Note that

\[
\frac{\partial \beta(\alpha, a^*(\alpha))}{\partial \alpha} \geq \frac{q}{p} \iff \alpha \leq \alpha' \iff \alpha \leq \hat{\alpha} \tag{22}
\]

were the last implication is due to proposition 1. A higher \( b^* \) corresponds to greater deterrence, and thus, we have the following result.

**Proposition 2**  
Deterrence is maximized by the preponderance of the evidence standard.

Although \( \hat{\alpha} \) maximizes deterrence, it need not maximize welfare. This is because, in addition to affecting deterrence, \( \alpha \) has an impact on the equilibrium level of avoidance, \( a^* \). The effects of \( \alpha \) on both \( a^* \) and \( b^* \) are illustrated in figure 2, below, and the effects of \( \alpha \) on welfare is analyzed, next.

3.6 Welfare

If welfare were to exclude the benefits and costs to offenders, an immediate implication of proposition 2 would be that \( \hat{\alpha} \) also maximizes welfare, because it minimizes the harm from crime. On the other hand, when offenders’ benefits and costs are included in the welfare function, it can be expressed as:

\[
W = \int_{b^*}^{\infty} (b - h - a^*)f(b)d(b) \tag{23}
\]

An important preliminary observation is that, welfare is increasing in \( b^* \), because \( h > 1 \), and it is decreasing in \( a^* \). Thus, if a policy leads to greater deterrence than another, second policy, without increasing avoidance activity compared to that policy, it unambiguously dominates that policy. This observation immediately implies that no standard of proof in the range \( [\alpha^m, \hat{\alpha}] \) can be optimal, because, as depicted in figure 2, an increase in this range leads to a reduction in avoidance activity and increases deterrence, too. Moreover, perhaps more importantly, because \( \hat{\alpha} > \alpha^m \), for each \( \alpha' \in [0, \alpha^m) \), one can find a
policy $\alpha'' \in (\alpha^m, 1]$ which leads to the same amount of avoidance investment while leading to greater deterrence. This idea is depicted in figure 2, and is formalized in the proof of the next proposition.
Proposition 3 When offenders can invest in avoidance, it is optimal to use a standard that is weaker than preponderance of the evidence.

Proof. Suppose the maximizer is not interior. In this case, it trivially follows that $W(1) = \int_{f(b)}^{\infty} (b-h)f(b)db > \int_{0}^{\infty} (b-h)f(b)db = W(0)$, and, thus, $\alpha^* = 1$, and, therefore, the optimal standard of proof is weaker than preponderance of the evidence.

When the maximizer is interior, note that welfare is increasing in deterrence (i.e. $b^*$) and decreasing in $\alpha^*$. Thus, if one could show that for any $\alpha^* < \alpha_m$, there exists $\alpha'' > \alpha_m$ such that $a^*(\alpha') = a^*(\alpha'')$, and $b^*(\alpha') < b^*(\alpha'')$, then it would follow that $\alpha^* \in [\alpha_m 1)$. Next, we prove that for all $\alpha^* < \alpha_m$ there in fact exists $\alpha'' > \alpha_m$ which satisfy these conditions.

First note, per (18), that $\frac{da^*(\alpha')}{\alpha}$ $> 0$ iff $\alpha \leq \alpha_m$, and that $a^*(0) = a^*(1) = 0$. Thus, for all $\alpha^* < \alpha_m$, there exists $\alpha'' > \alpha_m$ such that $a^*(\alpha') = a^*(\alpha'') = \pi^*$. Next, note that per (16) $\frac{\partial \beta(\alpha', \pi)}{\partial a} = \frac{\partial \beta(\alpha'') \pi}{\partial a}$. This implies that

$$\beta_0(\alpha') - \alpha' = \beta(\alpha'') - \alpha''$$

Multiplying both sides by $(1 - \gamma(\pi))$ and using the definition of $\beta$ we have:

$$\beta(\alpha', \pi) - \alpha' = \beta(\alpha'', \pi) - \alpha''$$

Multiplying both sides by $p$, and adding $\pi - q\alpha' - qa''$ we have:

$$p\beta(\alpha', \pi) + \pi - q\alpha' - qa'' - pa' = p\beta(\alpha'', \pi) + \pi - q\alpha' - qa'' - pa''$$

Using the definition of $b^*$ in (20) and re-arranging terms yields:

$$b^*(\alpha'') - b^*(\alpha') = (p - q)(\alpha'' - \alpha') > 0$$

where the inequality follows from the facts that $p > q$ and $\alpha'' > \alpha'$. Thus, $\alpha^* \in [\alpha_m, 1]$. This, combined with the observations that $\frac{da^*(\alpha')}{\partial a} > 0$ iff $\alpha \leq \tilde{\alpha}$; and $\frac{da^*(\alpha')}{\partial a} \leq 0$ iff $\alpha \geq \alpha_m$, imply that $\alpha^* \in (\alpha_m, 1)$.

It is important to note that proposition 3 does not rule out the possibility of strict liability being optimal, i.e. $\alpha^* = 1$. This becomes possible when the objective of minimizing avoidance activity, which is achieved by strict liability, is relatively important compared to the objective of maximizing deterrence. On the other hand, preponderance of the evidence, reflected through $\tilde{\alpha}$, maximizes deterrence. Thus, the optimal standard is interior and closer to $\tilde{\alpha}$ when deterrence is the more important objective (e.g. when $h$ is large).

4 Conclusion

The goal of this paper is to show how avoidance affects the optimal standard of proof in a setting where the standard of proof can be used to trade-off type-1 and type-2 errors. We show that when criminals can engage in costly avoidance to reduce their likelihood of being punished, then the optimal standard of
proof is weaker than preponderance of the evidence. This result stands in sharp contrast to much of the literature which has shown that (non-deterrence) costs usually cause the optimal standard of proof to be stronger than preponderance. While the primary objective of this paper is to study the relationship between avoidance and the standard of evidence, it also contributes to the literature on avoidance and deterrence more generally: deterrence and avoidance can move in the same or opposite directions as one alters the standard of proof, depending on how strong the standard of proof is to begin with. This dynamic is consistent with the overall implications of the avoidance literature, which shows that deterrence and avoidance may be positively or negatively related to each other depending on the enforcement and avoidance technology in place.

We conclude by noting that in other contexts avoidance has been shown to reverse well-established results (Sanchirico, 2010). Most notable among these is that Becker’s (1968) result that the optimal fine is maximal is contradicted by Malik (1990) and Langlais (2008) who show that the optimal fine is not maximal when criminals engage in avoidance. Thus, our paper can be thought of as complementing the existing literature by showing that the incorporation of avoidance activity reverses another well established result, namely that the presence of non-deterrence-related objectives cause a strengthening of the optimal standard of proof.

References


