CENTRE DE RECHERCHES EN ÉCONOMIE ET DROIT
Paris Center for Law and Economics

CRED WORKING PAPER N^o 2020-3

Public Law Enforcement under Ambiguity

February 2020

Bertrand Chopard*, Marie Obidzinski†

*Department of Health Policy and LIRAES, University Paris Descartes, Sorbonne Paris Cité, email: bertrand.chopard@parisdescartes.fr

†Department of Economics and CRED (TEPP), University Paris II Panthéon Assas, email: marie.obidzinski@paris2.fr

Public Law Enforcement under Ambiguity[1]

Bertrand Chopard
Department of Health Policy and LIRAES
University Paris Descartes, Sorbonne Paris Cité
bertrand.chopard@parisdescartes.fr

Marie Obidzinski
Department of Economics and CRED (TEPP)
University Paris II Panthéon Assas
marie.obidzinski@paris2.fr

Abstract: In real life situations, potential offenders may only have a vague idea of their own probability of getting caught and possibly, convicted. As they have beliefs regarding this probability, they may exhibit optimism or pessimism. Thus there exists a discrepancy between the objective expected fine and the subjective expected fine. In this context, we investigate how the fact that the choice whether or not to commit an offense is framed as a decision under ambiguity can modify the standard Beckerian results regarding the optimal fine and the optimal resources we should invest in detection and conviction. We discuss the validity of our results under three different social welfare functions.

Keywords: deterrence, pessimism, optimism, ambiguity.
JEL Codes: D81, K42.

[1] We are very grateful for their comments and suggestions to the participants at the "Behavioral Economics" seminar at Economix CNRS and Paris Nanterre University, and the 3rd French Law and Economics Annual Conference, and the American Law and Economics Association's annual conference (2019), and the Montpellier Research in Economics seminar, and the CRED seminar in Paris II. We are especially indebted to Alvaro Bustos, Johana Etner, Luigi Franzoni, Meglena Jeleva, Nuno Garoupa, Etienne Lehmann, Barbara Luppi, Lisa Morhaim, Mitchell Polinsky, Jennifer Reinganum for their fruitful suggestions. Any remaining errors or omissions are, of course, our own.

1 Introduction

In the public law enforcement framework *à la* Becker (1968), individuals are assumed to perfectly assess the probability of being fined if they commit an offense. However, in real life situations, potential offenders generally only have a vague idea of their own probability of getting caught and possibly, convicted. They have beliefs regarding the probability of detection and may exhibit optimism or pessimism. For instance, in the context of tax compliance, taxpayers tend to overestimate the probability of facing an investigation by the tax authority (Alm et al. 1992; Andreoni et al. 1998). In other situations, people are somewhat optimistic about their chances of not meeting with misfortune such as a car accident or illness (Jolls, 1998).

In our context, the way individuals estimate the probability of being detected will affect their decision whether or not to obey the law. This probability of detection and conviction is ambiguous while potential offenders know full well the amount of the fine. The main justification for this hypothesis is that sanctions are often detailed in sentencing guidelines or penal codes, while information about the probability of detection cannot be given. Furthermore, uncertainty over the size of fines raises issues regarding the principle of equality before the law (Universal Declaration of Human Rights, 1948, article 7). Imagine Mr. A and Mr. B commit the same crime under the same circumstances. If Mr. A is sentenced to 5 years while Mr. B is sentenced to 3 years, the difference seems quite unfair[2]. Next, we refer to Chateauneuf *et al.*'s (2007) framework to specify how each potential offender estimates the probability of detection and conviction.

The aim of our paper is to investigate how the fact that the choice whether or not to commit an offense is now framed as a decision under ambiguity[3] can modify the results regarding the optimal fine and the resources a benevolent public law enforcer should invest in detection and conviction. We consider successively three different objective functions for the authorities. In most of the paper, the aim of the law enforcer is to maximize *subjective* social welfare which is computed with potential offenders's beliefs about the probability of being detected and convicted (sections 3 to 5). Like most papers in the traditional public law enforcement literature (e.g., Garoupa 1997; Polinsky and Shavell 2007), by social welfare we refer to the benefits that individuals obtain from their behavior, the harm they cause, the cost of catching offenders, and also the cost of imposing sanctions on them — including here any psychological costs associated with ambiguity. In such a case, the law enforcer may be denoted *populist*. Salanié and Treich (2009) distinguish the *populist* regulator who "maximizes social welfare computed with citizens' beliefs" from the *paternalistic* regulator

[2] In practice, circumstances may vary a lot, and plea bargaining or prosecutorial discretion for instance might alter this principle.

[3] Ambiguity, as defined by Snow (2010), is uncertainty about probability, created by missing information that is relevant and could be known. See Etner et al. (2012) for a survey on decision theory under ambiguity.

who "maximizes social welfare computed with the regulator's own belief".[4] The key difference is whether the social planner takes into account the difference between how much offenders expect to pay and how much they might indeed pay, in other words, the "mental suffering associated with overestimation of the risk" as pointed out by Johansson-Stenman (2008). The discrepancy might generate a perception bias cost/gain (as it does with the subjective social welfare function) or not (as with the objective one). This question is closely related to a debate in public economics about whether such "mental suffering" (or the opposite) should be considered (Pollak 1998; Viscusi 2000; Salanié and Treich 2009; Johansson-Stenman 2008).[5]

Section 6 is devoted to the case of a paternalistic public law enforcer. A paternalistic public law enforcer does not care about the discrepancy between the expected and the actual expected fine. One reason may be that the law enforcer cannot observe these costs or gains, or that what matters for society is what citizens are actually paying and not what they subjectively expect to pay.

In section 7 we consider the case where the public law enforcer's objective is to minimize the social cost of crime, including both the detection cost and the external cost of crime. In such a case, the law enforcer does not take into account the gains from crime.

Our results call for the degree of pessimism of potential offenders in determining deterrence policy to be taken into account. Indeed, recommendations on deterrence policy can be widely affected by beliefs. Assume that individuals are pessimistic: they overestimate the probability of detection and conviction. We find that optimal fines are lower for two reasons. First, fines may be considered as costly transfers if society takes mental suffering into account. Second, the subjective probability of detection is higher than the objective probability. Regarding the optimal means to invest in detection, our results go in different directions depending on the objective function of the law enforcer. When the law enforcer is populist (he/she takes into account the perception cost), we show that it may be socially desirable to raise the probability and lower the magnitude of fine accordingly (in order to keep the deterrence level constant) if the marginal cost of detection is sufficiently small. In such a case, the fine is not necessarily maximal. When the law enforcer is paternalistic (he/she ignores the perception cost), the optimal fine is always maximal (as fines are costless transfers). And it is possible that the means invested in detection are lower than those in the absence of ambiguity only under certain conditions. In such a case, the objective probability of detection appears to be a less efficient deterrence tool due to the weight of the beliefs. If the law enforcer aims to minimize the social cost of crime, the results are similar: the optimal fine is maximal, and the optimal probability of detection may anew be lower than in the absence of ambiguity.

[4] The term "paternalism" refers to the protective attitude of an authority reminiscent of a father (*pater* in Latin).

[5] Note that the case where regulators themselves misperceive the risk is developed in Viscusi and Hamilton (1999). Arguably, it may not be the case in our framework since the law enforcer decides about the resources spent on detection.

The remainder of the paper is organized as follows. Section 2 presents related papers. In section 3, we present the model of public enforcement of law under ambiguity. In section 4, we derive a set of results when the probability of detection and conviction is exogenous. In section 5, both the monetary sanction and the probability of detection and conviction are endogenous. The paternalistic case is developed in section 6. In section 7, the public authority aims to minimize the cost of crime. Section 8 concludes.

2 Related literature

To our knowledge, a limited number of papers have addressed the link between crime deterrence and ambiguity. On the contrary, there is a larger number of articles introducing ambiguity in a tort law framework, such as Teitelbaum (2007) and Franzoni (2017). The closest contribution is certainly the one by Harel and Segal (1999). Their contribution focuses on describing *how* the legal system actually favors certainty relative to the sanction (for instance, through specifying the penalties in the criminal code or sentencing guidelines) and uncertainty towards the probability of detection and conviction and *why*. Regarding this second step, they invoke behavioral and psychological insights, namely ambiguity aversion (Ellsberg, 1961). In the analysis developed by Harel and Segal (1999), the authorities' objective is to choose the criminals' most disfavored law enforcement scheme in order to induce more deterrence at the lowest cost. They show that this scheme should consist in a certain sanction and an uncertain probability of getting caught. They base their analysis on contributions and results from the behavioral economic literature and Prospect theory (Kahneman and Tversky, 1979).

Our analysis differs in many aspects. We do not aim to determine whether uncertainty should be favored, but instead to examine the consequences of ambiguity for the optimal probability of detection and the amount of the fine. In addition, we consider both the cases where individuals might exhibit optimism and pessimism in a Chateauneuf et al. (2007) framework, while Harel and Segal (1999) consider only the case of ambiguity aversion. Furthermore, in our contribution the benevolent law enforcer aims to maximize social welfare. Conversely, Harel and Segal (1999) aim at determining the law enforcement scheme which maximizes deterrence at the cheapest cost.

A body of literature draws attention to the behavioral analysis of crime control (McAdams and Ulen 2009, Harel 2014, van Winden and Ash 2012). Garoupa (2003) discusses the relevance of a behavioral approach to the theory of public law enforcement. Many biases are presented and discussed in the context of public law enforcement. Jolls (2005) provides a detailed analysis on the negative impact of optimism (of agents with bounded rationality) on deterrence. Teichman (2011) argues that the behavioral analysis of crime control is quite limited by the indeterminacy bias effects. For instance, the evaluation of the probabilities might depend on whether they concern losses or gains, whether the probability is close to zero or to one, whether or not people are risk-seeking regarding punishment. Moreover, several opposite or "counter" biases might co-exist. For instance, the "availability bias" might recommend making enforcement highly visible, while ambiguity aversion may promotes concealing enforcement. Horovitz and Segal (2007) also note that the effect of ambiguity in crime deterrence should be carefully assessed according to whether the ambiguity refers to likely or unlikely events and whether ambiguity concerns gains or losses.

Our contribution is also related to imperfect information regarding the probability of arrest. This issue has been investigated notably by Bebchuk and Kaplow (1992), and more recently by Buechel et al. (2018). Bebchuk and Kaplow (1992) analyze the case where potential offenders get a noisy signal regarding the probability of getting caught. They show that the optimal sanction may be less than the maximal feasible sanction.[6] Buechel et al. (2018) consider two types of potential offenders; the naive ones, who are informed about the resources invested in law enforcement only if the authority decides to reveal that information, and the sophisticated ones, who are perfectly informed. They investigate when it is optimal to hide or reveal the enforcement effort.

3 Model and assumptions

3.1 Assumptions and notations

Our framework elaborates on the conventional model of public law enforcement (Polinsky and Shavell, 2007). Risk-neutral individuals choose whether or not to commit an act that yields a private benefit b and generates an external per act harm D. The public law enforcer does not observe any type b but knows their distribution described by a general density function $f(b)$ with support $[0, B]$ and a cumulative distribution function $F(b)$, with $D < B$. The proportion of offenders is equal to $1 - F(\widetilde{b})$, with $\widetilde{b}$ the deterrence threshold endogenously determined later.

The decision whether or not to commit an offense is represented in a Choquet expected utility framework, where individuals have difficulty assessing the probability of detection and conviction, while the magnitude of fines is perfectly known. Each potential offender estimates the probability of detection and conviction, and we denote this estimation as α which may be either greater than, less than, or equal to the objective probability of detection and conviction denoted p, with $0 < \alpha < 1$. Following Chateauneuf et al. (2007), pessimistic individuals will overestimate the probability of detection and conviction ($\alpha > p$), while optimistic individuals will underestimate their probability of being fined ($\alpha < p$). Furthermore, with the probability δ, an individual will wrongly estimate that the probability of detection and conviction equals α instead of p. With a probability $1-\delta$, the individual will correctly estimate the probability of detection and conviction, p. The *subjective* probability of being fined is written:

[6] In their setting, the relative size of the error decreases as the probability of detection rises.

$$\tilde{p} = \delta\alpha + (1-\delta)p = p + \delta(\alpha - p)$$

where δ measures the degree of ambiguity, that is the weight of the belief α in the subjective probability. If $\delta = 0$, there is no ambiguity as in the standard Beckerian framework. If $\delta = 1$, potential offenders consider themselves to be confronted with complete uncertainty. In the remainder of the paper, we will assume that $0 < \delta < 1$.

We denote the fine s, t is the tax imposed on each individual to finance detection and conviction, and w the individual level of (legal) wealth w, with $s \leq w$. The expected utility level of an individual who decides to commit an offense is written:

$$u_c = w + b - \tilde{p}s - t - (1 - F(\tilde{b}))D$$

Conversely, the utility of an individual who decides to abide by the law is:

$$u_a = w - t - (1 - F(\tilde{b}))D$$

Therefore, the individual commits an offense if and only if:

$$b \geq [\delta\alpha + (1-\delta)]ps = ps + \delta(\alpha - p)s = \widetilde{b}(p, s)$$

where $\widetilde{b}$ denotes the deterrence threshold that verifies $u_c = u_a$.

Potential offenders compare the benefit of committing the offense b with two terms: the objective expected fine ps, and a positive or negative additional term, $\delta(\alpha - p)s$. This new term results from the difference between the degree of pessimism and the objective probability of detection and conviction, $\alpha - p$, weighted by the degree of ambiguity δ. For pessimistic individuals, the deterrence threshold $\widetilde{b}(p, s)$ is now higher than the objective expected fine ps, while for optimistic individuals, the deterrence threshold is lower than the objective expected fine.

With no ambiguity ($\delta = 0$), the deterrence threshold is written $\widetilde{b}(p, s)|_{\delta=0} = ps$. In such a case, individuals are equally affected by a percentage increase in the fine or in the probability of detection and conviction, that is $e_s^{\widetilde{b}|_{\delta=0}} = e_p^{\widetilde{b}|_{\delta=0}} = 1$ where $e_s^{\widetilde{b}|_{\delta=0}}$ and $e_p^{\widetilde{b}|_{\delta=0}}$ are the elasticities of the threshold benefit value $\widetilde{b}$ with respect to the two deterrence tools in the absence of ambiguity. On the contrary, if there is ambiguity ($\delta > 0$), we have $e_s^{\widetilde{b}} = 1$ and $e_p^{\widetilde{b}} = \frac{(1-\delta)p}{\delta\alpha + (1-\delta)p} < e_s^{\widetilde{b}}$. A potential offender's decision to commit a crime is more sensitive to a percentage increase in the fine than to an equal percentage increase in the probability of detection and conviction. And it is all the more true when the degree of ambiguity or the degree of pessimism is high.[7]

[7] Because $\frac{\partial e_p^{\widetilde{b}}}{\partial \alpha} < 0$ and $\frac{\partial e_p^{\widetilde{b}}}{\partial \delta} < 0$.

3.2 The social welfare function

The social welfare function W is defined as the sum of the utilities of all individuals, given their decisions whether or not to commit an offense:

$$W = \int_{0}^{\tilde{b}(p,s)} u_a f(b)db + \int_{\tilde{b}(p,s)}^{B} u_c f(b)db \tag{1}$$

where potential offenders make their decision on the basis of the *subjective* probability of being fined.

The per capita cost to achieve the probability of detection and conviction $p \in [0, 1)$ is given by $m(p)$ where $m'(p) > 0$ and $m''(p) \geq 0$ (Polinsky and Shavell, 1979). Imposing and collecting fines is costless. Enforcement expenditures are financed through a lump sum tax t plus the fine s imposed on the offenders detected. The per capita public budget constraint is written:

$$m(p) = t + \left(1 - F(\widetilde{b})\right) ps \tag{2}$$

Only balanced-budget policies are considered.

Substituting the budget constraint (2) in (1), the social welfare function simplifies as:

$$W = w + \int_{\tilde{b}(p,s)}^{B} (b - D)f(b)db - \left(1 - F(\widetilde{b})\right)(\tilde{p} - p)s - m(p) \tag{3}$$

This expression exhibits an additional term by comparison with the standard case with no ambiguity: $(1 - F(\widetilde{b}))(\tilde{p} - p)s$. This term is referred to as the expected *perception bias*. If individuals are pessimistic, they overestimate the probability of getting fined ($\tilde{p} > p$), and the perception bias is a cost. As $\tilde{p} - p = (\alpha - p)\delta$, we have $\tilde{p} > p$ as long as $\alpha > p$, the reverse being true. If individuals are optimistic, then they underestimate the probability of getting caught ($\tilde{p} < p$), and the perception bias is a gain. As a consequence, the fine is no longer a pure transfer.

Remark 1: *Under ambiguity, monetary sanctions are not a pure transfer, as they induce either a cost or a gain.*

4 The optimal fine under ambiguity

We assume that enforcement expenditures are fixed, resulting in a given probability of detection and conviction. We will study the case with both an endogenous fine and an endogenous probability of being fined in the next section. We also divide everyone into two groups, optimistics on one side and pessimistics on the other. In order to simplify the exposition, only the case of pessimistics is presented below. We study optimistic individuals in appendix 7.

4.1 The no ambiguity case

In the absence of ambiguity surrounding the probability of detection and conviction, the public law enforcer chooses s_n^* which solves:

$$\max_s \left\{ W_n = w + \int_{ps}^{B} (b - D) f(b) db - m(p) \right\} \text{ u.c. } s \leq w.$$

The first-order condition is:

$$f(\tilde{b}_n) p (D - p s_n^*) = 0.$$

Thus we have $s_n^* = \frac{D}{p}$. The first-best outcome can be achieved as long as $\frac{D}{p} \leq w$.

4.2 The pessimistic case

Individuals are pessimistic ($\tilde{p} > p$). Under ambiguity, the law enforcer chooses s^* which solves:

$$\max_s \left\{ W = w + \int_{\tilde{b}(p,s)}^{B} (b - D) f(b) db - \left(1 - F(\widetilde{b})\right) (\tilde{p} - p) s - m(p) \right\} \text{ u.c. } s \leq w$$

The first-order condition is:[8]

$$f(\tilde{b}) \tilde{p} [(D - \tilde{b}) + s^* (\tilde{p} - p)] = \left(1 - F(\widetilde{b})\right) (\tilde{p} - p) \tag{4}$$

The left-hand side in (4) is the social marginal benefit of deterrence. An increase of the fine by one unit reduces the probability of offending by $-\frac{\partial(1-F(\tilde{b}))}{\partial s} = f(\tilde{b})\tilde{p}$ units. Thus, it diminishes the occurrence of both the net harm $D - \tilde{b}$ and the perception bias cost $(\tilde{p} - p)s$. The right-hand side equals the marginal cost of fines. Fines are costly because individuals overestimate the probability of getting fined.

We can rewrite (4) as:

$$\underbrace{h(\widetilde{b})\tilde{p}}_{(+)} (D - p s^*) = \underbrace{(\tilde{p} - p)}_{(+)} \tag{5}$$

where $h(\tilde{b}) = \frac{f(\tilde{b})}{1 - F(\tilde{b})}$ is the hazard rate function, that is the relative likelihood that $b = \widetilde{b}$ conditional on $b \geq \widetilde{b}$. According to (5), we find that $D - ps^* > 0$ or $s^* < s_n^* = \frac{D}{p}$.

Remark 2: *When individuals are pessimistic, the optimal fine under ambiguity is lower than the optimal fine in the standard Beckerian framework. Thus, it is socially desirable to set a lower fine under ambiguity as pessimism deters offenses and sanctions are no longer a costless transfer.*

[8] We report the analysis of second-order condition in appendix 1.

4.2.1 Under or overdeterrence

Underdeterrence is defined as a situation where $\tilde{b} < D$ at equilibrium: the subjective threshold benefit is lower than the harm. In such a case, some undeterred offenders obtain a benefit b lower than the harm D. Overdeterrence is defined as a situation where $\tilde{b} > D$: the subjective threshold benefit is higher than the harm. Some efficient offenses (such that $b > D$) are deterred.

We can rewrite (4) as follows:

$$\underbrace{f(\widetilde{b})\tilde{p}}_{(+)}(D - \widetilde{b}) = \underbrace{(\tilde{p} - p)}_{(+)}[\Big(1 - F(\widetilde{b})\Big) - f(\widetilde{b})\widetilde{b}] \tag{6}$$

where

$$[\Big(1 - F(\widetilde{b})\Big) - f(\widetilde{b})\widetilde{b}] < 0 \Longleftrightarrow 1 + e_{\widetilde{b}}^{1-F(\tilde{b})} < 0 \tag{7}$$

with $e_{\widetilde{b}}^{1-F(\tilde{b})} = \frac{-f(\tilde{b})\tilde{b}}{(1-F(\tilde{b}))} < 0$ standing for the elasticity of the proportion of offenders. This elasticity can also be interpreted as the sensitivity of the supply of offenses to the deterrence threshold[9]

From (6) and (7), we show that overdeterrence is socially desirable if and only if the proportion of offenders is elastic to the deterrence threshold ($1 < -e_{\widetilde{b}}^{1-F(\tilde{b})}$). It is socially optimal to increase the magnitude of fines when potential offenders are sensitive to fines and enforcement expenditures, up to a point making overdeterrence socially optimal. Conversely, underdeterrence is socially desirable if and only if the proportion of offenders is inelastic to the deterrence threshold ($1 > -e_{\widetilde{b}}^{1-F(\tilde{b})}$).

Remark 3: *When individuals are pessimistic, there is overdeterrence if and only if the proportion of offenders is elastic to the deterrence threshold. There is underdeterrence if and only of the proportion of offenders is inelastic to the deterrence threshold.*

4.2.2 Comparative statics

Let us first consider the impact of pessimism and ambiguity on the optimal magnitude of fine. Proposition 1 summarizes our results.

Proposition 1 *Assume that individuals are pessimistic. Then the optimal magnitude of fines increases with the degree of pessimism α and the degree of ambiguity δ if and only if:*

$$e_{\widetilde{b}}^{h(\widetilde{b})} > \frac{p}{\underbrace{(\tilde{p} - p)}_{(+)}}$$

where $e_{\widetilde{b}}^{h(\widetilde{b})} = \frac{h'(\widetilde{b})\widetilde{b}}{h(\widetilde{b})}$ stands for the elasticity of the hazard rate function with respect to the deterrence threshold.

[9] The term "supply" is used for instance by Becker (1968) and Ehrlich (1996).

Proof. See Appendix 2. ■

When individuals are pessimistic, the more pessimistic they are (or the more ambiguous the probability of detection and conviction is), the more it is socially desirable to increase the magnitude of fines if and only if the elasticity of the hazard rate function with respect to the deterrence threshold takes a sufficiently high positive value. As $\frac{p}{(\tilde{p}-p)} > 0$, this condition may hold if the hazard rate function $h(b)$ is increasing.[10] Note that $h'(b) > 0$ for most usual functions, such as the uniform one (Bagnoli and Bergstrom, 2005).

The link between the monotonicity of the hazard rate function and the marginal effect of pessimism and ambiguity on the optimal magnitude of fine appears more clearly by analyzing the first-order condition under this form:

$$\underbrace{h(\widetilde{b})}_{(+)}(D - ps^*) = \underbrace{\frac{(\tilde{p}-p)}{\tilde{p}}}_{(+)} \Longleftrightarrow \underbrace{h(\widetilde{b})[\delta\alpha + (1-\delta)p]}_{(+)}(D - ps^*) = \underbrace{\delta(\alpha - p)}_{(+)}.$$

Let us consider the marginal perception cost of deterrence on the right-hand side of second equation. If the degree of pessimism or the degree of ambiguity increases, the marginal cost also increases, thereby making it more socially desirable to reduce the magnitude of fines. Next, consider the marginal benefit of deterrence on the left-hand side. We have shown before that a one-unit increase in the fine reduces the probability of offending, thereby reducing the social harm at the margin by D per capita, net of the objective expected fine, ps, paid by offenders. Here, two effects play a role. First, an increase of the degree of the pessimism or the degree of ambiguity increases the subjective probability of detection and conviction, $\delta\alpha + (1-\delta)p$, which weights the marginal benefit of deterrence (defined by the difference between the social harm per capita and the objective expected fine, $D - ps$). Intuitively, an increase in pessimism or ambiguity increases the deterrence value of any given fine. Second, this marginal benefit of deterrence is weighted by the value of the hazard rate function at the threshold benefit of offense, $h(\widetilde{b})$. As a consequence, if the hazard rate function is increasing with the benefit of offense (and is high enough) then it is more plausible that the degree of pessimism or the degree of ambiguity will positively affect the optimal magnitude of fines. Deterrence becomes more desirable up to a point that when either pessimism or ambiguity increases, the optimal magnitude of fines also increases.

Finally, notice that there may also be a standard Beckerian trade-off in the sense that the optimal magnitude of fines can be reduced when confronted with an increase in pessimism or ambiguity, achieving the same deterrence level. This situation emerges if and only if $e_{\widetilde{b}}^{h(\widetilde{b})}$ takes a low positive value (or a negative

[10] By definition, it means that the distribution significantly puts less probability on extreme values of the benefit of crime. Put differently, the distribution is said to be a light- (resp. heavy-) tailed distribution when it has an increasing (resp. decreasing) hazard rate function.

value). Assume that the hazard rate function is decreasing, that is, the distribution significantly puts more probability on extreme values of the benefit of offenses. If larger values of benefit from offending are more plausible, then deterrence is less socially desirable. And it becomes more desirable to reduce the magnitude of fines as long as pessimism and ambiguity have a deterrence effect.

In proposition 2 below, we study the effect of the probability of detection and conviction on the optimal magnitude of fines.

Proposition 2 *Assume that individuals are pessimistic. The effect of the probability of detection and conviction on the optimal magnitude of fines is depicted in the table below.*

Pessimistic individuals	**Inelastic probability of offending**	**Elastic probability of offending**
$e_{\tilde{b}}^{h(\tilde{b})} > \underbrace{\frac{p}{\tilde{p}-p}}_{(+)}$	$\frac{ds^*}{dp} > 0$	unclear
$e_{\tilde{b}}^{h(\tilde{b})} < \underbrace{\frac{p}{\tilde{p}-p}}_{(+)}$	unclear	$\frac{ds^*}{dp} < 0$

Comment: The probability of offense is inelastic if $-e_{\tilde{b}}^{1-F(\tilde{b})} < 1$ and elastic if $-e_{\tilde{b}}^{1-F(\tilde{b})} > 1$.

Proof. See Appendix 3. ■

The marginal effect of the (exogenous) investment in detection and conviction on the optimal magnitude of fines depends on the elasticity of the hazard rate function with respect to the deterrence threshold. As for the degree of pessimism and the degree of ambiguity, the interaction between the monotonicity of the hazard rate function and the marginal effect of investment in detection and conviction may be understood *via* the first-order condition. We do not report all the reasoning as it is rather similar to the previous comparative static results.

We show in Proposition 2 that there is an additional condition to explain the way that the magnitude of investment in detection and conviction will affect the optimal magnitude of fines. The probability of offense must be elastic in order to have a substitution effect between the two deterrence tools. It is socially desirable to reduce the magnitude of fines when investment in detection and conviction (exogenously) rises on condition that the probability of offense is elastic. Under this condition, we have shown before that overdeterrence is socially optimal thus leading to a high level of deterrence at equilibrium. It is then plausible to replace one instrument of deterrence with the other as long as deterrence is high enough. The reverse is true when the probability of offense is inelastic. Deterrence is much lower at equilibrium, thus leading to underdeterrence being socially desirable. In such a case, the optimal magnitude of fines tends to increase with the investment in detection and conviction to keep deterrence at equilibrium.

5 Fines and detection under ambiguity

In this part, we assume that the public law enforcer determines both the size of the fine and the probability of detection. We determine first the optimal magnitude of fine and the optimal probability of detection, and second we analyze the substitution between the two deterrence tools.

5.1 Optimal fine and detection

5.1.1 The no ambiguity case

In the absence of ambiguity, the public law enforcer solves:

$$\max_{s,p}\left\{W_n = w + \int_{ps}^{B}(b-D)f(b)db - m(p)\right\} \text{ u.c. } s \leq w.$$

The first-order conditions relative to s and p respectively are:

$$f(p_n^* s_n^*)p(D - p_n^* s_n^*) = 0 \tag{8}$$

$$f(p_n^* s_n^*)s_n^*(D - p_n^* s_n^*) = m'(p_n^*) \tag{9}$$

where p_n^* is defined by (9) with $s_n^* = w$ as in the standard Beckerian framework.

5.1.2 The pessimistic case

Under ambiguity, the law enforcer solves:

$$\max_{s,p}\left\{W = w + \int_{\tilde{b}(p,s)}^{B}(b-D)f(b)db - \left(1 - F(\widetilde{b})\right)(\tilde{p}-p)s - m(p)\right\} \text{ u.c. } s \leq w.$$

The first-order conditions relative to s and p respectively are:

$$f(\tilde{b})\tilde{p}^*(D - \tilde{p}^* s^*) + f(\tilde{b})\tilde{p}^* s^*(\tilde{p}^* - p^*) = \left(1 - F(\widetilde{b})\right)(\tilde{p}^* - p^*)$$

$$f(\widetilde{b})(1-\delta)s(D - \widetilde{b}) + (1 - F(\widetilde{b}))\delta s + f(\widetilde{b})(1-\delta)s(\tilde{p}^* - p^*)s = m'(p^*) \tag{10}$$

The interpretation of the derivative of W relative to the fine is the same as in section 3. The interpretation of the derivative of W relative to the probability of detection and conviction deserves more attention. The right-hand side of equation (10) stands for the marginal cost of detection and conviction. The left-hand side of equation (10) is the social marginal benefit of deterrence, which sums three terms.

First, an increase in the probability of detection and conviction by one unit reduces the proportion of offenders by $-\frac{\partial(1-F(\widetilde{b}))}{\partial p} = f(\widetilde{b})(1-\delta)s$ units, thereby diminishing the occurrence of net harm by $(D - \widetilde{b})$.

Next, there is a perception bias cost of deterrence as each offender does not expect to pay ps but $\tilde{p}s$. If offenders are pessimistic ($\tilde{p} > p$), the discrepancy between how much they expect to pay and the objectively expected fine $(\tilde{p}-p)s$ reflects the perception bias cost. A one-unit increase of the probability of detection and conviction diminishes this discrepancy, thus reducing this perception cost by δs units, which occurs with probability $(1 - F(\widetilde{b}))$.

Finally, an increase in p raises the expected gains provided by fines. The subjective probability $\tilde{p}$ increases as well, but by less than one unit ($\frac{\partial \tilde{p}}{\partial p} = 1-\delta$). At the same time, the proportion of offenders decreases by $-\frac{\partial(1-F(\widetilde{b}))}{\partial s} = f(\widetilde{b})(1-\delta)s$ units, thereby saving the perception bias cost $(\tilde{p} - p)s$.

5.2 Substitution between deterrence tools

In this subsection, we determine the conditions under which it is socially desirable to replace one deterrence tool by the other. We report our results in Proposition 3 below.

Proposition 3 *Assume that individuals are pessimistic. It is socially desirable to increase the probability of detection and conviction while decreasing the magnitude of fines if and only if*

$$e_{\alpha}^{\widetilde{b}} s > \frac{m'(p)}{(1 - F(\widetilde{b}))}$$

where $e_{\alpha}^{\widetilde{b}}$ *is the elasticity of the deterrence threshold relative to the degree of pessimism.*

Proof. See Appendix 4. ■

Assume that there is no ambiguity ($\delta = 0$). Then, we have $e_{\alpha}^{\widetilde{b}|_{\delta=0}} = 0$. According to Proposition 3, it is then socially desirable to decrease the probability of detection and conviction while increasing the magnitude of fines. Intuitively, there is neither any perception cost, nor any cost to collect fines, while detection and conviction is costly. We thus find the standard Beckerian result: $s_n^* = w$.

Next, assume that $\delta > 0$: the probability of detection and conviction becomes ambiguous. The condition under which it is optimal to replace investment in ambiguous detection and conviction with the well known fine at the margin is written $e_{\alpha}^{\widetilde{b}} s < \frac{m'(p)}{1-F(\widetilde{b})}$. The intuition behind this result runs as follows. First, detection and conviction must be costly enough (see $\frac{m'(p)}{1-F(\widetilde{b})} > 0$). Second, potential offenders must not be too sensitive to the degree of pessimism (see the term $e_{\alpha}^{\widetilde{b}}$). Recall that individuals are more affected by a percentage increase in the magnitude of fines than by an equal percentage increase in the probability of detection and conviction ($e_p^{\widetilde{b}} < e_s^{\widetilde{b}}$). Third, it is all the more socially desirable to increase the optimal magnitude of fines while reducing the investment in detection and conviction when fines are little used to deter offenses (see the variable s to the left-hand side of the inequality above).

6 The paternalistic social welfare function

Up to now, the law enforcer was seen as populist because social welfare was based on individuals' beliefs. At the opposite, a paternalistic benevolent law enforcer takes into account the expected cost ps actually experienced by individuals who decide to commit crimes. Thus, the paternalistic social welfare function is written as (1) with $u_c = w + b - ps - t - (1 - F(\widetilde{b}))D$ instead of $u_c = w + b - \tilde{p}s - t - \left(1 - F(\widetilde{b})\right) D$. Thus, the social welfare function is written:

$$W_a = w + \int_{\tilde{b}(p,s)}^{B} (b - D) f(b) db - m(p)$$

The social welfare equals the individual wealth w, plus the gains from crime net of the harm caused $\int_{\tilde{b}(p,s)}^{B} (b - D) f(b) db$, and less the cost of detection $m(p)$.

Remark 4: *If the public law enforcer is paternalistic, monetary sanctions under ambiguity are seen as a costless transfer.*

6.1 The optimal fine

Under ambiguity, the paternalistic law enforcer solves:

$$\max_s \left\{ W_a = w + \int_{\tilde{b}(p,s)}^{B} (b - D) f(b) db - m(p) \right\} \text{ u.c. } s \leq w$$

where $\widetilde{b}(p, s) = ps + \delta(\alpha - p)s$. The first-order condition is:

$$f(\tilde{b})\tilde{p}(D - \tilde{b}) = 0 \tag{11}$$

Thus, $s^* = \frac{D}{\tilde{p}}$. As $\tilde{p} = \delta(\alpha - p)$, the optimal fine is decreasing with the degree of pessimism and the degree of ambiguity because pessimistic individuals will overestimate the expected fine. The first-best outcome can be achieved as long as $\frac{D}{\tilde{p}} \leq w$.

Remark 5: *With a paternalistic law enforcer, the first-best outcome $\widetilde{b} = D$ can be achieved as long as $\frac{D}{\tilde{p}} \leq w$.*

If individuals are pessimistic ($\tilde{p} > p$), the optimal sanction under ambiguity with a paternalistic public law enforcer is still lower than the optimal fine in the absence of ambiguity: $s^* < s_n^* = \frac{D}{p}$.

Remark 6: *With a paternalistic law enforcer, the optimal fine under ambiguity is lower than the optimal fine without ambiguity.*

6.2 Optimal fine and detection

Both deterrence tools are now endogenous. Under ambiguity, the paternalistic law enforcer solves:

$$\max_{s,p} \left\{ W_a = w + \int_{\tilde{b}(p,s)}^{B} (b - D) f(b) db - m(p) \right\} \text{ u.c. } s \leq w.$$

The derivatives of W_a relative to s and p respectively are given by:

$$f(\tilde{b})\tilde{p}^*(D - \tilde{p}^* s^*) = 0 \tag{12}$$

$$f(\tilde{b})(1 - \delta)s^*(D - \tilde{b}) - m'(p^*) = 0 \tag{13}$$

As the fine is cost-free in this setting (as in the standard Beckerian model), the fine is set at its maximum: $s^* = w$. Note also that $D > \tilde{b} = \tilde{p}w$ at equilibrium as the first-order condition on p indicates that:

$$\underbrace{f(\tilde{b})(1 - \delta)w}_{(+)}(D - \tilde{b}) = \underbrace{m'(p^*)}_{(+)}$$

Proposition 4 *Assume that individuals are pessimistic. With a paternalistic public law enforcer choosing both the fine and the probability of detection, the optimal fine is maximal. The optimal probability of detection and conviction under ambiguity is higher (resp. lower) than the optimal probability of detection in the absence of ambiguity if:*

$$e_{\delta}^{\tilde{b}} \frac{1 - \delta}{\delta} \left(e_{\tilde{b}}^{f(\tilde{b})} - \frac{\tilde{b}}{D - \tilde{b}} \right) > 1 \text{ (resp. } < 1). \tag{14}$$

Proof. See Appendix 5. ■

The comparison between the optimal expenditures in detection and conviction in the no ambiguity case versus the pessimistic case is not straightforward. Let us compare the first-order condition with p in cases (9) and (13). Notice that the marginal cost of detection and conviction (to the right-hand side above) does not depend on the degree of ambiguity, while the marginal benefit of detection and conviction does. It depends on a large set of factors, such as the elasticity of the threshold benefit value with respect to the degree of ambiguity ($e_{\delta}^{\tilde{b}}$), a relative measure of the degree of ambiguity ($\frac{1-\delta}{\delta}$), the elasticity of the density function with respect to the threshold value of benefit ($e_{\tilde{b}}^{f(\tilde{b})}$) minus a relative measure of the underdeterrence at equilibrium ($\frac{\tilde{b}}{D-\tilde{b}} > 0$). For sketch of simplicity, assume that the benefit of crime follows a uniform distribution over the interval $(0, B)$. Under this condition, the optimal probability of detection and conviction under ambiguity is lower than the optimal probability of detection in the no ambiguity case.

7 Minimizing the social cost of crime

One may consider, as Dau-Schmidt (1990), that it is morally shocking to include criminal benefits in the social welfare function. More generally, criminals' preferences should not be considered as belonging to social preferences, since criminals are people society wants to exclude. In this view, the public law enforcer aims to minimize the social cost of crime defined as the sum of the expected external harm induced by offenses, and the cost of detection and conviction. Thus, the public law enforcer solves:

$$\min_{s,p} \left\{ SC = \left(1 - F(\widetilde{b})\right) D + m(p) \right\} \text{ u.c. } s \leq w.$$

The derivatives of SC relative to s and p respectively are given by:

$$-f(\tilde{b})\tilde{p}^* D = 0 \tag{15}$$

$$f(\tilde{b})(1-\delta)s^* D = m'(p^*) \tag{16}$$

In this framework, increasing the fine reduces the social cost. Therefore, the fine should be set at its maximum $s^* = w$. The optimal probability of detection is defined by equation (16). The left-hand side of (16) represents the marginal benefit. Increasing the probability of detection decreases the proportion of offenders, thereby reducing the occurrence of social harm. The right-hand side of (16) represents the marginal cost of increasing detection.

In the absence of ambiguity, the program is:

$$\min_{s,p} \left\{ SC = \left(1 - F(\widetilde{b})\right) D + m(p) \right\} \text{ u.c. } s \leq w.$$

The optimal fine s_n^* equals w because the fine is cost-free. The probability of being fined is defined by:

$$f(ps)wD = m'(p^*) \tag{17}$$

To determine how the probability of detection and conviction is affected by ambiguity at equilibrium, we compare left-hand sides of equations (16) and (17). It is sufficient to show that the marginal benefit of detection in (16) increases with the degree of ambiguity to conclude that the optimal probability of detection and conviction under ambiguity is higher than the optimal probability of detection and conviction in the no ambiguity case, the reverse being true.

Proposition 5 *Assume that individuals are pessimistic. If the public law enforcer aims to minimize the social cost of crime, then the optimal probability of detection and conviction under ambiguity is higher (resp. lower) than the optimal probability of detection in the absence of ambiguity if:*

$$e_{\tilde{b}}^{f(\tilde{b})} > \underbrace{\frac{\tilde{p}}{(\alpha - \tilde{p})}}_{(+)} \text{ (resp. } e_{\tilde{b}}^{f(\tilde{b})} < \frac{\tilde{p}}{(\alpha - \tilde{p})} \text{).}$$

Proof. See Appendix 6. ■

The comparison between the optimal expenditures in detection and conviction in the no ambiguity case versus the ambiguity case anew depends on the distribution of benefit of crime. For instance, if crime benefit follows a uniform distribution, then the optimal probability of detection and conviction under ambiguity is lower than the optimal probability of detection in the no ambiguity case.

Remark 8: *If the public law enforcer aims to minimize the social cost of crime, then the optimal magnitude of fine should be set at its maximum. The optimal probability of detection and conviction under ambiguity is lower than the optimal probability of detection in the no ambiguity case if the benefit of crime is uniformly distributed.*

8 Conclusion

The aim of this paper is to provide some insights into how the degree of pessimism influences the socially optimal amount of fines and enforcement expenditures when potential offenders are unable to perfectly estimate the probability of detection and conviction.

By comparison with the standard Beckerian framework (Garoupa, 1997; Polinsky and Shavell, 2000), fines are no longer a costless transfer if indivi duals overestimate their probability of being caught with a populist law enforcer. The reason is that populist law enforcers take into account the discrepancy between the actual expected fine and the subjectively expected fine. This discrepancy, denoted as perception bias, reflects disutility when potential offenders are pessimistic. Consequently, the optimal fine is lower than the maximal one, contrary to the standard Beckerian framework. Further, it may be socially optimal to raise the probability of detection while decreasing the magnitude of fine under a certain condition. This result, as well as the comparative statics analysis, considerably depends on the distribution of crime's benefit.

With a paternalistic law enforcer, the perception cost is no longer taken into account. Fines are a pure transfer between the detected offenders and society, as fines are used to fund detection. In such a case, the first-best outcome can be achieved as long as the optimal fine is lower or equal than the maximal one, as in the standard Beckerian framework. We show that the optimal fine adjusts to the impact of pessimism on the deterrence threshold. The more pessimistic the offenders, the lower the optimal fine. When both the fine and the probability of detection are endogenous, the optimal fine is maximal as in the standard Beckerian framework. We also show that the optimal probability of detection under ambiguity may be lower than the optimal probability in the absence of ambiguity under certain conditions. For instance, this result holds if the benefit of crime is uniformly distributed. The means invested in detection should be reduced due to the weight of beliefs, making potential offenders less sensitive to a rise in the probability of detection.

In order to complete the picture, we consider the case where the law enforcer's aim is to minimize the social cost function, defined as the external harm induced by offenses, plus the cost of detection and conviction. When the public law enforcer seeks to minimize the social cost of crime, ambiguity plays no role regarding the fine: it should be set at its maximum. However, the means invested in detection may be anew reduced due to the weight of beliefs, making potential offenders less sensitive to a rise in the probability of detection under a certain condition.

Let us say a word about the optimistic case (appendix 7). If potential offenders are optimistic, it is still true that fines are no longer a pure transfer with a populist law enforcer. But the similarities with the pessimistic case end there or almost. When potential offenders are optimistic, they underestimate the probability of detection. Thus the deterrence threshold is lower than the objective one. And the perception bias now reflects a gain, as individual expect to pay less that what will actually go into the coffers of the State (to fund deterrence policy). Consequently, the results go in the opposite direction than in the pessimistic case. The optimal fine under ambiguity should be higher than the standard Beckerian fine, and the comparative statics results are reversed. Further, with a paternalistic law enforcer, the optimal fine perfectly adjusts to the belief to reach the optimal objective deterrence threshold. With optimistic potential offenders, it means that the optimal fine under ambiguity is higher than the optimal fine in absence of ambiguity. Finally, with a public law enforcer aiming at minimizing the social cost of crime, the optimal probability of detection and conviction under ambiguity is lower than the optimal probability of detection in the no ambiguity case if the benefit of crime is uniformly distributed. The explanation is the same as with pessimistic law offenders: due to the weight of beliefs, potential offenders less sensitive to a rise in the probability of detection.

Our contribution has several limits. We do not consider the possibility of "debiasing" potential offenders (Jolls and Sunstein, 2006). We could imagine that the law enforcer attempts to educate potential offenders about the risk of being optimist, or to reduce individuals' cost of information through public warnings. However, the effect of such attempts on optimistic individuals' beliefs might be limited. Indeed, the optimistic bias is often associated with a "blind spot bias" (Pronin et al., 2002). The "blind spot bias" is the illusion that one is less prone to bias, and notably optimism. Therefore, it is quite difficult if not impossible to modifiy optimistic individuals' beliefs (Luppi and Parisi, 2016).

Another limitation of our contribution is that we consider a homogeneous population in terms of beliefs. Individuals share either the same degree of ambiguity, or the same degree of pessimism. Furthermore, we do not consider risk-aversion or risk-seeking. In a sense, our contribution may be seen as a first step towards producing a formal representation of ambiguity in a public law enforcement model.

References

Alm, J., McClelland, G. H. and W. D. Schulze. 1992. "Why do People Pay Taxes?" *Journal of Public Economics* 48(1): 21-38.

Andreoni, J., Erard, B. and J. Feinstein. 1998. "Tax Compliance," *Journal of Economic Literature* 36(2): 818-860.

Bagnoli, M. and Bergstrom, T. 2005. "Log-concave probability and its applications," *Economic Theory* 26: 445-469.

Becker, G. 1968. "Crime and Punishment: an Economic Approach," *Journal of Political Economy* 76: 169-217.

Bebchuk, L. A., and L. Kaplow. 1992. "Optimal Sanctions when individuals are imperfectly informed about the probability of apprehension," *The Journal of Legal Studies* 21.2 (1992): 365-370.

Buechel, B., E. Feess, and G. Muehlheusser. 2018 "Optimal Law Enforcement with Sophisticated and Naive Offenders," Working Paper No. 7106. CESifo Group Munich.

Chateauneuf, A., Eichberger, J., and Grant, S. 2007. "Choice under Uncertainty with the Best and Worst in Mind: Neo-Additive Capacities," *Journal of Economic Theory*, 137(1), 538-567.

Ellsberg, D. 1961. "Risk, ambiguity, and the Savage axioms," *The Quarterly Journal of Economics* 643-669.

Ehrlich, I. 1996. "Crime, Punishment, and the Market for Offenses." *Journal of Economic Perspectives* 10.1: 43-67.

Etner, J., M. Jeleva and J.-M. Tallon. 2012. "Decision Theory Under Ambiguity," *Journal of Economic Surveys*, 26(2) : 234-270.

Franzoni, L.A. 2017. "Liability Law under Scientific Uncertainty." *American Law and Economics Review* 19(2): 327-360.

Garoupa, N. 1997. "The Theory of Optimal Law Enforcement." *Journal of Economic Surveys*, 11(3), 267-295.

Garoupa, N. 2003. "Behavioral Economic Analysis of Crime: A Critical Review." *European Journal of Law and Economics*, 15(1), 5-15.

Harel, A. 2014. "Behavioural Analysis of Criminal Law: a Survey." *Bergen Journal of Criminal Law & Criminal Justice*, 2(1), 32-55.

Harel, A., and U. Segal. 1999 "Criminal Law and Behavioral Law and Economics: Observations on the Neglected Role of Uncertainty in Deterring Crime." *American Law and Economics Review* 1.1: 276-312.

Horovitz, A., and U. Segal. 2006. "The Ambiguous Nature of Ambiguity and Crime Control," *New York University Journal of Law and Liberty* 2: 541.

Johansson-Stenman, O. 2008. "Mad Cows, Terrorism and Junk Food: Should Public Policy Reflect Perceived or Objective Risks?." *Journal of Health Economics* 27(2) : 234-248.

Jolls, C. 1998. "Behavioral Economics Analysis of Redistributive Legal Rules," *Vanderbilt Law Review* 51: 1653.

Jolls, C. 2005. "On Law Enforcement with Boundedly Rational Actors," In: in F. Parisi and V.L. Smith, eds. *The Law and Economics of Irrational Behavior*. Stanford University Press, 268-286.

Jolls, C., and Sunstein, C. R. 2006. "Debiasing Through Law." *The Journal of Legal Studies* 35.1 (2006): 199-242.

Kahneman, D. and Tversky, A. 1979. "Prospect Theory: An Analysis of Decision Theory Under Risk," *Econometrica* 47: 263-91.

Luppi, B. and F. Parisi. 2016. "Optimal Liability for Optimistic Tortfeasors." *European Journal of Law and Economics* 41.3 (2016): 559-574.

McAdams, R. H. and Ulen, T.S. 2009. "Behavioral Criminal Law and Economics". In *Criminal Law and Economics*, Nuno Garoupa eds. (Edward Elgar, 2009).

Polinsky, A. M., and Shavell, S. 1979. "The Optimal Tradeoff Between the Probability and Magnitude of Fines". *The American Economic Review*, 69(5):880-891.

Polinsky, A. M., and Shavell, S. 2007. "The Theory of Public Enforcement of Law," *Handbook of Law and Economics*, 1:403-454.

Pollak, R. A. 1998. "Imagined Risks and Cost-Benefit Analysis," *The American Economic Review*, 88(2), 376-380.

Pronin, E., Lin, D., and Ross, L., 2002, "The Bias Blind Spot: Perceptions of Bias in Self Versus Others," *Personality and Social Psychology Bulletin*, 28: 369-381.

Salanié, F., and Treich, N. 2009. "Regulation in Happyville." The Economic Journal 119.537: 665-679.

Snow, A. 2010. "Ambiguity and the Value of Information." *Journal of Risk and Uncertainty* 40: 133-145.

Teichman, Doron. 2011. "The Optimism Bias of the Behavioral Analysis of Crime Control." *University of Illinois Law Review* 1697.

Teitelbaum, J. C. 2007. "A Unilateral Accident Model Under Ambiguity." *The Journal of Legal Studies* 36.2: 431-477.

van Winden, F A., and E. Ash. 2012. "On the Behavioral Economics of Crime." *Review of Law and Economics* 8.1 (2012): 181-213.

Viscusi, W. K. 2000. "Risk equity." *The Journal of Legal Studies*, 29(S2), 843-871.

Viscusi, W. K., and Hamilton, J. T. 1999. "Are Risk Regulators Rational? Evidence from Hazardous Waste Cleanup Decisions." *American Economic Review* 89.4: 1010-1027.

9 Appendix

9.1 A1: Second-order condition

The second-order condition for s is written:

$$h'(\widetilde{b})(\delta\alpha + (1-\delta)p)(D - ps) < ph(\widetilde{b})$$

where $h(\widetilde{b}) = \frac{f(\widetilde{b})}{1-F(\widetilde{b})}$ is the hazard rate function estimated at the threshold value $\widetilde{b}$. The sign of the right-hand side depends on the individuals being either optimistic or pessimistic, while the sign of the left-hand side depends on the monotonicity of the hazard rate function.

Assume that individuals are pessimistic: $\alpha - p > 0$. Using (5), we have shown that $D - ps^* > 0$ at equilibrium. Thus $h'(\widetilde{b}) < 0$ is a sufficient condition for the second-order condition to hold. If the hazard rate function is monotonically decreasing, it is an indication that the distribution significantly puts more probability on larger values of the benefit of illegal activity. Put differently, the distribution of the benefit of committing the harmful act has a heavy tail, meaning a tail that is heavier than an exponential distribution. And the heavier the tail, the higher the probability that illegal benefits will take one or more disproportionate values in the population.

Using first-order condition, we can replace $ph(\widetilde{b}) = \frac{\delta(\alpha-p)p}{(\delta\alpha+(1-\delta)p)(D-ps)}$ in second order condition above. If the hazard rate function is increasing ($h'(\widetilde{b}) > 0$) as for frequently used elementary distributions such as the uniform, normal and exponential ones, then the second-order condition condition can be rewritten:

$$h'(\widetilde{b}) < \underbrace{\frac{\delta(\alpha - p)p}{(\delta\alpha + (1-\delta)p)^2(D - ps)^2}}_{(+)}.$$

If the hazard rate function is increasing then the distribution puts less probability on larger values of the illegal benefit. The probability distribution has a thinner tail than an exponential distribution. This means that it goes to zero much faster than the exponential, and so has less mass in the tail.

Assume that individuals are optimistic. According to (5), we now have $D - ps^* < 0$ at equilibrium because $\widetilde{p} < p$. Thus, $h'(\widetilde{b}) > 0$ is a sufficient condition for the second-order condition to hold. If $h'(\widetilde{b}) < 0$, then the second-order condition writes:

$$h'(\widetilde{b}) > \underbrace{\frac{\delta(\alpha - p)p}{(\delta\alpha + (1-\delta)p)^2(D - ps)^2}}_{(-)}.$$

9.2 A2: Proof of Proposition 1

We start by computing the sign of the derivative $\frac{ds^*}{d\alpha}$. The first-order condition can be written:

$$h(\widetilde{b})\,(\delta\alpha + (1-\delta)p)\,(D - ps) - \delta(\alpha - p) = 0.$$

Using the implicit function theorem and the second-order condition, we obtain:

$$\frac{ds^*}{d\alpha} > 0 \Leftrightarrow h'(\widetilde{b})\widetilde{b}(D - ps)\delta + h(\widetilde{b})\delta(D - ps) - \delta > 0.$$

Dividing by δ:

$$\frac{ds^*}{d\alpha} > 0 \Leftrightarrow h'(\widetilde{b})\widetilde{b}(D - ps) + h(\widetilde{b})(D - ps) > 1.$$

Using the first-order condition, we replace $h(\widetilde{b})(D - ps) = \frac{\delta(\alpha-p)}{\delta\alpha+(1-\delta)p}$ and put it on the right-hand side. So

$$\frac{ds^*}{d\alpha} > 0 \Leftrightarrow h'(\widetilde{b})\widetilde{b}(D - ps) > 1 - \frac{\delta(\alpha - p)}{\delta\alpha + (1-\delta)p} = \frac{p}{\delta\alpha + (1-\delta)p}.$$

Still using the first-order condition, we now replace $D - ps = \frac{\delta(\alpha-p)}{(\delta\alpha+(1-\delta)p)h(\widetilde{b})}$ and simplify by $(\delta\alpha + (1-\delta)p)$ to have:

$$\frac{ds^*}{d\alpha} > 0 \Leftrightarrow \frac{h'(\widetilde{b})\widetilde{b}}{h(\widetilde{b})}\delta(\alpha - p) > p.$$

If individuals are pessimistic $(\alpha - p > 0)$ then $\frac{ds^*}{d\alpha} > 0 \Leftrightarrow \frac{h'(\widetilde{b})\widetilde{b}}{h(\widetilde{b})} > \frac{p}{\delta(\alpha-p)}$ where $\frac{h'(\widetilde{b})\widetilde{b}}{h(\widetilde{b})}$ stands for the elasticity of the hazard rate function estimated at the threshold value. If individuals are optimistic $(\alpha - p < 0)$ then $\frac{ds^*}{d\alpha} > 0 \Leftrightarrow \frac{h'(\widetilde{b})\widetilde{b}}{h(\widetilde{b})} < \frac{p}{\delta(\alpha-p)}$.

Now, we study the sign of the derivative $\frac{ds^*}{d\delta}$. Using the implicit function theorem, the first and second-order conditions, we get:

$$\frac{ds^*}{d\delta} > 0 \Leftrightarrow (\alpha - p)((h'(\widetilde{b})\widetilde{b}(D - ps) + h(\widetilde{b})(h - ps) - 1)$$

We have previously shown the conditions under which the expression $(h'(\widetilde{b})\widetilde{b}(D-ps)+h(\widetilde{b})(D-ps)-1)$ is either positive or negative. Thus, if $\alpha-p>0$ (pessimistic case) then $\frac{ds^*}{d\delta} > 0 \Leftrightarrow \frac{h'(\widetilde{b})\widetilde{b}}{h(\widetilde{b})} > \frac{p}{\delta(\alpha-p)}$. If $\alpha - p < 0$ (optimistic case) then $\frac{ds^*}{d\delta} > 0 \Leftrightarrow \frac{h'(\widetilde{b})\widetilde{b}}{h(\widetilde{b})} < \frac{p}{\delta(\alpha-p)}$ if $(\alpha - p) < 0$.

9.3 A3: Proof of Proposition 2

The first-order condition is:

$$h(\widetilde{b})\left(\delta\alpha + (1-\delta)p\right)(D - ps) - \delta(\alpha - p) = 0.$$

Using the implicit function theorem and the second-order condition, we have:

$$\frac{ds^*}{dp} > 0 \Leftrightarrow h'(\widetilde{b})\widetilde{b}(D-ps)(1-\delta) + h(\widetilde{b})(1-\delta)(D-ps) - h(\widetilde{b})\widetilde{b} + \delta > 0.$$

Adding 1 while subtracting 1 on the left-hand side of the inequality above gives:

$$\frac{ds^*}{dp} > 0 \Leftrightarrow h'(\widetilde{b})\widetilde{b}(D-ps)(1-\delta) + h(\widetilde{b})(1-\delta)(D-ps) - h(\widetilde{b})\widetilde{b} + \delta \underbrace{+1-1} > 0.$$

After rearranging, we have:

$$\frac{ds^*}{dp} > 0 \Leftrightarrow (1-\delta)[h'(\widetilde{b})\widetilde{b}(D-ps) + h(\widetilde{b})(D-ps) - 1] - h(\widetilde{b})\widetilde{b} + 1 > 0.$$

We have shown in proof of Proposition 1 the conditions under which the expression $(h'(\widetilde{b})\widetilde{b}(D-ps) + h(\widetilde{b})(D-ps) - 1)$ is either positive or negative. Further, we have: $-h(\widetilde{b})\widetilde{b} + 1 > 0 \Leftrightarrow -e_{\widetilde{b}}^{1-F(\widetilde{b})} > 1$ where $-e_{\widetilde{b}}^{1-F(\widetilde{b})} > 0$ is the elasticity of the probability of offense with respect to the threshold value. All the cases are depicted in table 1.

9.4 A4: Proof of Proposition 3

By definition, $\widetilde{b} = (\delta\alpha + (1-\delta)p)\, s$. So,

$$\frac{ds}{dp}_{\widetilde{b}=cst} = \frac{-(1-\delta)s}{\delta\alpha + (1-\delta)p} < 0.$$

Fully differentiating the social welfare function, we have:

$$dW = -\left(1 - F(\widetilde{b})\right)(\delta(\alpha - p)ds - \delta s dp) - m'(p)dp.$$

Factorizing by dp , we have:

$$dW = dp\left((1-F(\widetilde{b}))\left(-\delta(\alpha-p)\frac{ds}{dp}+\delta s\right)-m'(p)\right).$$

Thus

$$\begin{aligned} dW_{\widetilde{b}=cst} &= dp\left((1-F(\widetilde{b}))\left(\frac{\delta(\alpha-p)(1-\delta)s}{\delta\alpha+(1-\delta)p}+\delta s\right)-m'(p)\right) \\ &= dp\left(\underbrace{(1-F(\widetilde{b}))e_{\alpha}^{\widetilde{b}}s}_{(+)}-m'(p)\right) \end{aligned}$$

where $e_{\alpha}^{\widetilde{b}}$ is the elasticity of the deterrence threshold with respect to the degree of pessimism.

As a consequence, if $e_{\alpha}^{\widetilde{b}}s > \frac{m'(p)}{(1-F(\widetilde{b}))}$ then it is socially desirable to increase the probability of detection and conviction while decreasing the magnitude of fines. If $e_{\alpha}^{\widetilde{b}}s < \frac{m'(p)}{(1-F(\widetilde{b}))}$ then it is socially desirable to decrease the probability of detection and conviction while increasing the magnitude of fines, and to choose $s^* = w$ at equilibrium.

9.5 A5: Proof of Proposition 4

In the no ambiguity case, the optimal probability of detection and conviction is given by:

$$f(p_n^* w)w(D-p_n^* w) = m'(p_n^*)$$

while in the paternalistic case (under ambiguity), it is given by:

$$f(\tilde{b})(1-\delta)w(D-\tilde{b}) = m'(p^*)$$

with $\tilde{b} = [\delta\alpha + (1-\delta)p]w$ and $D > \tilde{b}$ at equilibrium.

The marginal cost of detection and conviction (at the right-hand side above) does not depend on the degree of ambiguity while the marginal benefit of detection and conviction does in the paternalistic case. Remark that the two values of marginal benefit of detection and conviction depicted above are equal if $\delta = 0$. Thus it is sufficient to show that the marginal benefit of detection in the paternalistic case is increasing with the degree of ambiguity (δ) to conclude that the optimal probability of detection and conviction under ambiguity is higher than the optimal probability of detection and conviction in the no ambiguity case. Conversely, if the marginal benefit of detection and conviction in the paternalistic case decreases with the degree of ambiguity, then the optimal probability of detection and conviction under ambiguity is now lower than the optimal probability of detection and conviction in the no ambiguity case.

Using the implicit function theorem and the first-order condition in the paternalistic case, we show that the marginal benefit $f(\tilde{b})(1-\delta)w(D-\tilde{b})$ is increasing with δ if and only if:

$$f'(\tilde{b})(\alpha-p)w(D-\tilde{b})(1-\delta) - f(\tilde{b})(1-\delta)(\alpha-p)w - f(\tilde{b})(D-\tilde{b})w > 0$$

or

$$f(\tilde{b})\underbrace{(D-\tilde{b})}_{(+)}\left(\frac{f'(\tilde{b})}{f(\tilde{b})}(\alpha-p)(1-\delta)-\frac{(1-\delta)(\alpha-p)}{D-\tilde{b}}-1\right)>0$$

or

$$e_{\tilde{b}}^{f(\tilde{b})}\left(\frac{(\alpha-p)(1-\delta)\delta}{\delta\tilde{b}}\right)-\frac{\delta(1-\delta)(\alpha-p)\tilde{b}}{\delta(D-\tilde{b})\tilde{b}}-1>0$$

where $e_{\tilde{b}}^{f(\tilde{b})}=\frac{f'(\tilde{b})}{f(\tilde{b})}\tilde{b}$ is the elasticity of the density function with respect to the threshold value of benefit. Next, replacing $e_{\delta}^{\tilde{b}}=\frac{\delta(\alpha-p)w}{\tilde{b}}$ the elasticity of the threshold value of benefit with respect to the degree of ambiguity in the inequality above, we have:

$$e_{\tilde{b}}^{f(\tilde{b})}e_{\delta}^{\tilde{b}}\frac{1-\delta}{\delta}-e_{\delta}^{\tilde{b}}\frac{1-\delta}{\delta}\frac{\tilde{b}}{D-\tilde{b}}-1>0$$

or

$$e_{\delta}^{\tilde{b}}\frac{1-\delta}{\delta}\left(e_{\tilde{b}}^{f(\tilde{b})}-\frac{\tilde{b}}{D-\tilde{b}}\right)>1.$$

9.6 A6: Proof of Proposition 5

If the law enforces seeks to minimize the social cost of crime under ambiguity, then $s^*=w$ and p^* is defined by $f(\tilde{b})(1-\delta)wD=m'(p^*)$. The marginal benefit of detection, $f(\tilde{b})(1-\delta)wD$, is increasing with δ if and only if $f'(\tilde{b})(\alpha-p)w(1-\delta)-f(\tilde{b})>0$. Factorizing by $\frac{-f'(\tilde{b})}{f(\tilde{b})}\tilde{b}=e_{\tilde{b}}^{f(\tilde{b})}$ gives $e_{\tilde{b}}^{f(\tilde{b})}(1-\delta)(\alpha-p)w>\tilde{b}$. Next replace $(1-\delta)(\alpha-p)$ by $\alpha-\widetilde{p}$ at left-hand side, and replace $\widetilde{b}$ by $\widetilde{p}w$ at right-hand side. We obtain that the marginal benefit of detection under ambiguity increases with the degree of ambiguity if and only if $e_{\tilde{b}}^{f(\tilde{b})}(\alpha-\widetilde{p})>\widetilde{p}$. This is sufficient to conclude that the optimal probability of detection and conviction under ambiguity is higher than the optimal probability of detection and conviction in the no ambiguity case. In the pessimistic case $(\alpha-p>0)$, this condition rewrites $e_{\tilde{b}}^{f(\tilde{b})}>\frac{\widetilde{p}}{\alpha-\widetilde{p}}>0$ while in the optimistic case $(\alpha-p<0)$, it rewrites $e_{\tilde{b}}^{f(\tilde{b})}<\frac{\widetilde{p}}{\alpha-\widetilde{p}}<0$.

9.7 A7: Optimistic potential law offenders

There exists some crucial differences between the optimistic case and the pessimistic case. First, when individuals are optimistic, they underestimate the probability of detection and conviction: $\tilde{p}<p$ or $\alpha<p$. As a consequence, the deterrence threshold ($\tilde{p}s$) is inferior to the objective expected fine (ps). Second, the elasticity of the deterrence threshold relative to the degree of ambiguity is negative: $e_{\delta}^{\tilde{b}}<0$. It means that an increase of the degree of ambiguity has a deterrence effect while the reverse is true if invididuals are pessimistic. Third, the perception bias defined as $(p - \widetilde{p})s > 0$ is now a gain while it is a cost in the pessimistic case.

9.7.1 Populistic law enforcer

The populist public law enforcer chooses s^* which solves:

$$\max_s \left\{ W = w + \int_{\tilde{b}(p,s)}^{B} (b - D)f(b)db - \left(1 - F(\widetilde{b})\right)(\tilde{p} - p)s - m(p) \right\} \text{ u.c. } s \leq w$$

The first-order condition (4) is:

$$f(\tilde{b})\tilde{p}(D - \tilde{b}) + \left(1 - F(\widetilde{b})\right)(p - \tilde{p}) = f(\tilde{b})\tilde{p}s^*(p - \tilde{p}).$$

The left-hand side in equation above is the marginal benefit of deterrence. It combines two terms. First, an increase of the fine by one unit reduces the probability of offending by $-\frac{\partial(1-F(\tilde{b}))}{\partial s} = f(\tilde{b})\tilde{p}$ units thereby diminishing the occurrence of the net harm $D - \tilde{b} > 0$. Second, the positive term $\left(1 - F(\widetilde{b})\right)(p - \tilde{p})$ means that raising the fine increases the expected perception bias gain that is the difference between the expected fine and the subjective expected fine. The right-hand side equals the marginal cost of fines. An increase of the fine by one unit reduces the probability of offending by $\frac{-\partial\left(1-F(\widetilde{b})\right)}{\partial s}$ units, thereby preventing individuals from getting the perception bais (gain): $(p - \tilde{p})s > 0$.

According to (5), we have $h(\widetilde{b})\tilde{p}(D - ps^*) = (\tilde{p} - p) < 0$ where $h(.) > 0$ is the hazard rate function. Thus we find that $D - ps^* < 0$ or $s^* > s_n^* = \frac{D}{p}$. When individuals are optimistic, the optimal fine under ambiguity is higher than the optimal fine in the standard Beckerian framework while it is lower when individuals are pessimistic. It is socially desirable to increase the magnitude of fines beyond the value $\frac{D}{p}$ for two reasons. First, there is an additional marginal benefit associated with the use of fines due to the perception bias. Second, optimistic potential offenders are less sensitive to fines.

From equations (6) and (7), we explain how the psychological group (pessimists or optimists) has an influence on the social desirability of under or overdeterrence. If the proportion of offenders is elastic to the deterrence threshold, then overdeterrence is socially desirable if individuals are pessimistic while underdeterrence is socially desirable if individuals are optimistic. At the opposite, if the proportion of offenders is inelastic to the deterrence threshold, then overdeterrence is socially desirable if individuals are optimistic while underdeterrence is socially desirable if individuals are pessimistic.

We turn to the comparative statics analysis. We show in appendix 2 that the optimal magnitude of fine is decreasing (increasing) with the degree of pessimism α and the degree of ambiguity δ if and only if $e_{\widetilde{b}}^{h(\widetilde{b})} > \frac{p}{(\tilde{p}-p)}$ under the condition that individuals are optimistic (resp. pessimistic).

Next, we report in table below the effect of the probability of detection and conviction on the optimal magnitude of fines for each psychological group (see appendix 3 for the proof).

		Optimists	Pessimists
Inelastic probability	$e_{\widetilde{b}}^{h(\widetilde{b})} > \frac{p}{\delta(\alpha-p)}$	unclear	$\frac{ds^*}{dp} > 0$
of offense	$e_{\widetilde{b}}^{h(\widetilde{b})} < \frac{p}{\delta(\alpha-p)}$	$\frac{ds^*}{dp} > 0$	unclear
Elastic probability	$e_{\widetilde{b}}^{h(\widetilde{b})} > \frac{p}{\delta(\alpha-p)}$	$\frac{ds^*}{dp} < 0$	unclear
of offense	$e_{\widetilde{b}}^{h(\widetilde{b})} < \frac{p}{\delta(\alpha-p)}$	unclear	$\frac{ds^*}{dp} < 0$

Now, assume that the two deterrence tools are endogenous. If offenders are optimistic ($\tilde{p} < p$), the optimal probability of detection and conviction is implicitly defined by:

$$\left(1 - F(\widetilde{b})\right)\delta s + f(\widetilde{b})(1-\delta)s(D-\widetilde{b}) = -f(\widetilde{b})(1-\delta)s\delta(\alpha - p^*)s + m'(p^*)$$

The left-hand side equals the marginal benefit. We have shown before that the discrepancy between how much they expect to pay and the objective expected fine (that is $(\tilde{p}-p)s = \delta(\alpha-p)s$) is positive, and represents the perception bias gain. Thus, a one-unit increase of the probability of detection and conviction raises this gain by δs units, which occurs with probability $\left(1 - F(\widetilde{b})\right)$. In other words, an increase in detection raises the expected sanction more than the subjective expected sanction, and this net marginal difference equals δs. Next, the proportion of offenders decreases by $-\frac{\partial(1-F(\widetilde{b}))}{\partial s} = f(\widetilde{b})(1-\delta)s$ units, thereby diminishing the occurrence of the net harm $(D-\widetilde{b})$. The right-hand side stands for the marginal cost. There are two terms: $m'(p)$ and $-f(\widetilde{b})(1-\delta)s\delta(\alpha-p^*)s$. The second term means that a one-unit increase of the probability of detection and conviction reduces the probability of offense by $-\frac{\partial\left(1-F(\widetilde{b})\right)}{\partial s} = f(\widetilde{b})(1-\delta)s$ units, thereby reducing the occurrence of the perception bias gain $\delta(\alpha-p^*)s = (\tilde{p}-p)s$.

9.7.2 Paternalistic law enforcer

A paternalistic benevolent law enforcer solves:

$$\max_s \left\{ W_a = w + \int_{\widetilde{b}(p,s)}^{B} (b-D)f(b)db - m(p) \right\} \text{ u.c. } s \leq w.$$

The first-order condition is $f(\tilde{b})\tilde{p}(D-\tilde{b}) = 0$. So we have $s^* = \frac{D}{\tilde{p}}$ whatever the individual being optimistic or pessimistic. As $\tilde{p} = \delta\alpha + (1-\delta)p$, the optimal magnitude of fine is still decreasing with the degree of pessimism, but is increasing with the degree of ambiguity because optimistic individuals underestimate the probability of detection. Next, the first-best outcome $(\widetilde{b} = D)$ can be achieved as long as $\frac{D}{\tilde{p}} \leq w$. Finally, if individuals are optimistic ($\tilde{p} < p$), the optimal sanction under ambiguity (with a paternalistic public law enforcer) is higher than the optimal fine in absence of ambiguity $s^* > s_n^* = \frac{D}{p}$ while the reverse is true if individuals are pessimistic.

9.7.3 Minimizing social costs

The public law enforcer solves:

$$\min_{s,p} \left\{ SC = \left(1 - F(\widetilde{b})\right) D + m(p) \right\} \text{ u.c. } s \leq w.$$

As the derivatives of SC relative to s is $-f(\tilde{b})\tilde{p}^* D = 0$, the fine should be set at its maximum $s^* = w$. And the optimal probability of detection is still defined by the equation $f(\tilde{b})(1 - \delta)wD = m'(p^*)$. We show in Appendix 6 that the optimal probability of detection and conviction under ambiguity is higher than the optimal probability of detection and conviction in the no ambiguity case if and only if $e_{\tilde{b}}^{f(\tilde{b})} < \frac{\widetilde{p}}{\alpha - \widetilde{p}} < 0$. It follows that the optimal probability of detection and conviction under ambiguity is lower than the optimal probability of detection in the no ambiguity case if the benefit of crime is uniformly distributed (as $e_{\tilde{b}}^{f(\tilde{b})} = 0$).