1 Model

A society is made by a unit mass of individuals. Individuals differ in their propensity to commit a crime and this propensity is distributed in the population according to an absolutely continuous cdf G with support normalied in $[0,1]$. The individual benefit from taking action $a = 1$ is given by θ. At each point in time an individual is randomly drawn from the population and has to decide whether to engage in an illegal activity (action $a = 1$) or not (action $a = 0$). The illegal activity generates a harm ρ to the society. Neither the agent's action a, nor the harm ρ is observable.[1]

The action chosen by an individual stocastically generates an output $e \in [0,1]$ that is observable and verifiable in court. We refer to e as to the *incriminating evidence*. In particular, if the individual takes action a, incriminating evidence is produced according to an absolutely continuous cdf H_a; the pdf associated with H_a is denoted with h_a. Since the action chosen by an individual is not observable, while incriminating evidence e is, the society will try to infer a based on the observation of e. The next assumption, known as Monotone Likelihood Ratio Property (MLRP), implies that such inference will put higher and higher weight on $a = 1$ as e increases.

Assumption 1 $\frac{h_1(e)}{h_0(e)}$ *is increasing in* e.

Assumption 1 has two immediate implications: (i) H_1 first-order stocastically dominates H_0, and (ii) $\frac{1-H_0(e)}{h_0(e)} < \frac{1-H_1(e)}{h_1(e).}$ for every e.[2]

To deter crime, a social planner sets up a *judicial system* represented by a triple $(p, \bar{e}, F)$. p represents the probability with which the individual is subject to the scrutiny

[1] We can assume that the harm ρ is observable after the end of the trial.

[2] To see these two results, notice that assumption 1 implies that for every $e' > e$:

$$h_0(e)\, h_1(e') \geq h_0(e')\, h_1(e).$$

Thus, integrating e from 0 to e', we get that for every $e' \in [0,1]$:

$$\frac{H_0(e')}{H_1(e')} \geq \frac{h_0(e')}{h_1(e')}$$

Similarly, integrating with respect to e' from e to 1, we get that for every $e \in [0,1]$

$$\frac{h_0(e)}{h_1(e)} \geq \frac{(1 - H_0(e))}{(1 - H_1(e))}.$$

Implication (ii) follows immediately from this last inequality. Implication (i), instead, follows from combining the two inequalities and noticing that

$$\frac{H_0(e)}{H_1(e)} \geq \frac{(1 - H_0(e))}{(1 - H_1(e))} \; \forall e \in [0,1]$$

implies that H_1 first order stocastically dominates H_0.

of the judicial sysyem (*inspection probability*); $\bar{e} \in [0, 1]$ is the level of incriminating evidence necessary to convict an agent who is scrutinized (*standard of proof*); F is the fine levied on convicted individuals (*penalty*). The judicial system chosen by the planner is known by everybody.

By construction, a given judicial system $(p, \bar{e}, F)$ is prone to two different types of errors. On the one hand it could convict an innocent and this will happen with probability $1 - H_0(\bar{e})$; on the other hand it could acquit a guilty individual with probability $H_1(\bar{e})$. We refer to these two types of error as *to type I and type II errors*, respectively. We will define $1 - (1 - H_0(\bar{e})) - H_1(\bar{e}) = H_0(\bar{e}) - H_1(\bar{e})$ as the *accuracy of the judicial system*.

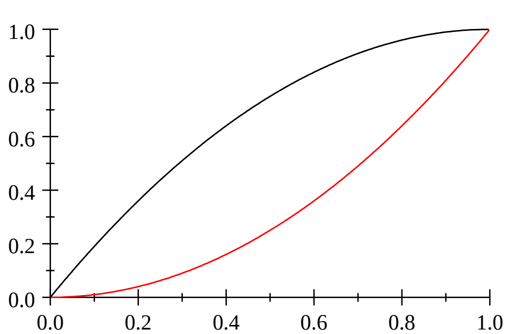
Finally, we will assume that $F < \bar{F}$ as individuals are subject to limited liabilities. We start assuming that increasing p is costly, whereas changing the level of incriminating evidence $(\bar{e})$ and of the inspection probability is costless [**Justify**]. In particular, let $C_p(p)$ be the cost associated with inspection probability p; where $C_p(p)$ is a strictly increasing and strictly convex functions with the property that $\lim_{x \to 0} C_p(x) = 0$.

The goal of the social planner aims to minimize the amount of harm taking into account the cost associated with crime detterence. Formally, the social planner chooses the judicial system $(p, \bar{e}, F)$ in order to maximize the following objective function:

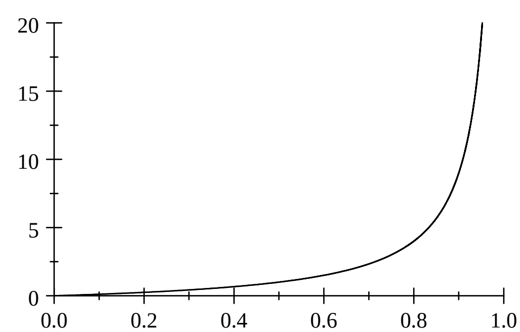
$$W(p, \bar{e}, F) = -\int_0^1 \rho \cdot a(\theta) \cdot dG(\theta) - C_p(p)$$

The social planner faces a trade off: he wants to minimize social harm, at the same time he wants to save on the costs associated with crime deterrence.

For analytical convenience, we will sometime assume the following functional forms: $G(\theta) = 1$ for every $\theta \in [0, 1]$, $C_p(x) = c_p \cdot \frac{x^2}{2}$, $H_0(e) = \int_0^e 2(1 - x)\, dx$ and $H_1(e) = \int_0^e 2x dx$. Below, we plot the cdf functions and the likelihood ratio $(\frac{h_1(e)}{h_0(e)})$of the two pdfs:



$H_0(e)$ in black and $H_1(e)$ in red.



Likelihood ratio of pdfs: $\frac{h_1(e)}{h_0(e)}$

2 Crime Deterrence with Standard Risk-Neutral Agent

Suppose that the agent is a standard expected utility maximizer with linear vNM utility index. Then if an agent with type θ faces a judicial system $(p, \bar{e}, F)$, he will engage in the illegal activity if and only if:[3]

$$\theta - p \cdot (1 - H_1(\bar{e})) \cdot F > -p \cdot (1 - H_0(\bar{e})) \cdot F$$

Thus the marginal individual who is indifferent between committing the crime or not is identified by:

$$\theta^*(p, e, F) = R(p, \bar{e}, F) \tag{1}$$

where $R(p, \bar{e}, F) \equiv p \cdot (1 - (1 - H_0(\bar{e})) - H_1(\bar{e})) \cdot F$.

(1) immediately implies that illegal activity can be discouraged using three different tools: (i) an increase in the inspection probability (p), (ii) an improvement in the accuracy of the judicial system $((1 - (1 - H_0(\bar{e})) - H_1(\bar{e})))$, and (iii) an increase in the amount of fines (i.e., an increase in F). Notice that, this behavioral rule implies that the deterrence power of a judicial system depends on the incriminating evidence only through total accuracy $((1 - H_0(\bar{e})) - H_1(\bar{e}))$ and not on the actual composition of the two errors.

Thus, the maximization problem of the social planner can be written as follows:

$$\max_{(p,\bar{e},F)} - \int_{\theta^*(p,e,F)}^{1} \rho \cdot dG(\theta) - C_p(p)$$

The first order necessary condition for an internal solution are given by:

$$\rho \cdot g\left(\theta^*\left(p^{RN}, \bar{e}^{RN}, F^{RN}\right)\right) \cdot \left(H_0\left(\bar{e}^{RN}\right) - H_1\left(\bar{e}^{RN}\right)\right) \cdot F^{RN} = \frac{\partial C_p\left(p^{RN}\right)}{\partial p} \tag{2}$$

$$\rho \cdot g\left(\theta^*\left(p^{RN}, \bar{e}^{RN}, F^{RN}\right)\right) \cdot F^{RN} \cdot \left(h_0\left(\bar{e}^{RN}\right) - h_1\left(\bar{e}^{RN}\right)\right) \cdot p^{RN} = 0 \tag{3}$$

$$\rho \cdot g\left(\theta^*\left(p^{RN}, \bar{e}^{RN}, F^{RN}\right)\right) \cdot p^{RN} \cdot \left(H_0\left(\bar{e}^{RN}\right) - H_1\left(\bar{e}^{RN}\right)\right) \geq 0 \tag{4}$$

Since the last condition is everywhere positive, F will be set equal to $\bar{F}$. Moreover, from the second first order necessary conditiom, we can conclude that $\bar{e}^*$ will be chosen so that:

$$h_0\left(\bar{e}^{RN}\right) = h_1\left(\bar{e}^{RN}\right)$$

Thus, the standard of proof will be chosen in order to maximize the accuracy of the

[3] For simplicity, we assume that, if indifferent, the agent breaks such an indifference by choosing $a = 0$.

judicial system. Finally, the optimal level of detection will be determined by:

$$\rho \cdot g\left(\theta^*\left(p^{RN}, \bar{e}^{RN}, \bar{F}\right)\right) \cdot \left(H_0\left(\bar{e}^{RN}\right) - H_1\left(\bar{e}^{RN}\right)\right) \cdot \bar{F} = \frac{\partial C_p\left(p^{RN}\right)}{\partial p}$$

Given the convexity of the cost function, it is immediate to check that the optimal inspection probability p^{RN} depends positively on: (i) the amount of social harm caused by $a = 1$ (ρ), (ii) the maximal fine which can be levied on agents ($\bar{F}$) if convicted, (iii) the accuracy of the judicial system ($H_0\left(\bar{e}^{RN}\right) - H_1\left(\bar{e}^{RN}\right)$).

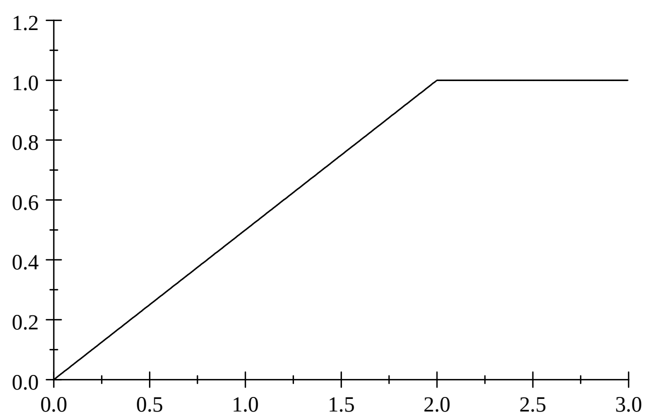
Furthermore, if we assume $G(\theta) = 1$ for every $\theta \in [0,1]$, $C_p(x) = c_p \cdot \frac{x^2}{2}$, $H_0(e) = \int_0^e 2(1-x)\,dx$, $H_1(e) = \int_0^e 2x dx$, the first order necessary conditions for an interior optimum becomes:

$$\rho \cdot \left(1 - \left(1 - \int_0^{\bar{e}^{RN}} 2(1-x)\,dx\right) - \int_0^{\bar{e}^{RN}} 2xdx\right) \cdot F^{RN} = p^{RN}$$

$$2\left(1 - \bar{e}^{RN}\right) = 2 \cdot \bar{e}^{RN}$$

$$F^{RN} = \bar{F}$$

and the optimal judicial system will be given by: $F^{RN} = \bar{F}$, $\bar{e}^{RN} = 0.5$, $p^{RN} = \frac{1}{2} \cdot \rho \cdot \bar{F}$. Below, we plot the optimal level of inspection probability as a function of ρ. Notice that for every $\rho > 2$, the optimal solution involves the highest possible inspection probability $p^{RN} = 1$.



p^{RN} as a function of ρ in the special case

3 Crime Deterrence under Reference-Dependence and Loss Aversion a là Kőszegi and Rabin

Now, we will consider the case in which individuals exhibit reference-dependent preferences a là Kőszegi and Rabin. These preferences are meant to capture two behavioral phenomena that have received strong support in experimental settings, reference dependence and loss aversion. *Reference dependence* refers to individuals' tendency to evaluate the events they experience not only in absolute terms, but also relatively to their expectations. In particular, we will assume individuals care both about the *consumption utility* induced by a certain outcome (represented by the utility index associated with a certain event), but also on how this utility compares with a reference outcome and determine the so-called *gain/loss utility*. *Loss aversion* further captures the tendency of individuals to dislike losses (i.e., negative deviations from the reference utility) more than they enjoy equal-size gains (i.e., positive deviations from the reference utility).

Formally, maintaining the assumption of linear vNM utility indexes, the *total utility* associated with outcome x, when the reference point is r would be given by:

$$v\left(x,r\right)=x+\mu\left(x-r\right),$$

where $\mu\left(\cdot\right):\mathbb{R}\rightarrow\mathbb{R}$ is a function defined as follows:

$$\mu\left(x-r\right)=\eta\cdot\max\left\{x-r,0\right\}+\eta\lambda\cdot\min\left\{x-r,0\right\}$$

with $\eta\in\mathbb{R}_{+}$ and $\lambda\geq1$. Thus, η captures the relative importance of reference dependence, whereas λ measures loss aversion.

Moreover, in line with Kőszegi and Rabin, we will assume that individuals' reference points are determined through a rational expectation approach, namely by the distribution over outcomes associated with the action they plan to take. Thus, if an individual facing judicial system $(p,\bar{e},F)$ takes action $a\in\{0,1\}$, his reference point will be given by a random measure that assigns probability $H_a\left(\bar{e}\right)$ to $a\cdot\theta$ and $p\cdot\left(1-H_a\left(\bar{e}\right)\right)$ to $a\cdot\theta-F$.

Thus, an individual who faces judicial system $(p,\bar{e},F)$ and engages in the illegal activity $(a=1)$ will have an expected total utility equal to:

$$\begin{aligned}u\left(1,p,\bar{e},F\mid\theta\right) &= \theta-p\cdot\left(1-H_1\left(\bar{e}\right)\right)\cdot F-\\ &\quad -\eta\cdot\left(\lambda-1\right)\cdot\left(1-p\cdot\left(1-H_1\left(\bar{e}\right)\right)\right)\cdot p\cdot\left(1-H_1\left(\bar{e}\right)\right)\cdot F\end{aligned}$$

Similarly, if an individual behaves legally $(a=0)$ his expected total utility will be given by:

$$u(0,p,\bar{e},F \mid \theta) = -(1-H_0(\bar{e}))\cdot F - \\ -\eta\cdot(\lambda-1)\cdot(1-p\cdot(1-H_0(\bar{e})))\cdot p\cdot(1-H_0(\bar{e}))\cdot F$$

Subtracting $u(0,p,\bar{e},F \mid \theta)$ from $u(1,p,\bar{e},F \mid \theta)$ and equating to 0, we can identify a threshold on the benefit from committing a crime that makes an individual indifferent between the two actions. Such a threshold is given by:

$$\theta^{**}(p,\bar{e},F,\eta,\lambda) = L(p,\bar{e},F,\eta,\lambda), \tag{5}$$

where:

$$L(p,\bar{e},F,\eta,\lambda) = p\cdot(1-(1-H_0(\bar{e}))-H_1(\bar{e}))\cdot F\cdot\big(1+ \\ +\eta\cdot(\lambda-1)\cdot(1-p\cdot((1-H_0(\bar{e}))+(1-H_1(\bar{e}))))\big)$$

Comparing 1 with 5, we can draw some interesting conclusions. First of all, for a given judicial system an increase in the parameters capturing reference dependence and loss aversion (respectively, η and λ) deters crime if and only if

$$p < \frac{1}{(1-H_0(\bar{e}))+(1-H_1(\bar{e}))}, \tag{6}$$

namely, if and only if convictions are relatively likely or, equivalently, the standard of proof is sufficiently high. Moreover, it is easy to check that the composition of errors plays a role in crime deterrance. Indeed, whereas a decrease in type I errors $(1-H_0(\bar{e}))$ is unambiguously associated with an increase in crime deterrance, a decrease in type II errors $(H_1(\bar{e}))$ leads to a reduction in the deterrance power of loss aversion. The intuition behind these results is as follows: an increase in the standard of proof $(\bar{e})$ has two effects: (i) it leads to a decrease in type I errors, and (ii) to an increase in type II erorrs. Under reference dependence and loss aversion, these changes affect not only the overall accuracy of the judicial system, but also the expected losses associated with different actions. As a result, a raise in $\bar{e}$ may lower the gain/loss utility component associated with $a=1$ vis-a-vis the one associated with $a=0$ diminishing the deterrance power of the judicial system. Thus, whereas a decrease in type I errors $(1-H_0(\bar{e}))$ deter crime both through its effect on the accuracy of the judicial system and through the one on gain/loss utility components, a decrease in type II $(H_1(\bar{e}))$ will result in the two effects partially offsetting each other.

In this setting, the social planner will choose the judicial system in order to solve the following maximization problem:

$$\max_{(p,\bar{e},F)} -\int_{\theta^{**}(p,\bar{e},F,\eta,\lambda)}^{\bar{\theta}} \rho \cdot dG(\theta) - C_p(p)$$

The first order necessary conditions for an interior solution associated to this problem are:

$$\begin{aligned}
\rho \cdot g\left(\theta^{**}\left(p^{RD}, \bar{e}^{RD}, F^{RD}, \eta, \lambda\right)\right) \cdot \frac{\partial L\left(p^{RD}, \bar{e}^{RD}, F^{RD}, \eta, \lambda\right)}{\partial p} &= \frac{\partial C_p\left(p^{RD}\right)}{\partial p} \\
\rho \cdot g\left(\theta^{**}\left(p^{RD}, \bar{e}^{RD}, F^{RD}, \eta, \lambda\right)\right) \cdot \frac{\partial L\left(p^{RD}, \bar{e}^{RD}, F^{RD}, \eta, \lambda\right)}{\partial \bar{e}} &= 0 \\
\rho \cdot g\left(\theta^{**}\left(p^{RD}, \bar{e}^{RD}, F^{RD}, \eta, \lambda\right)\right) \cdot \frac{\partial L\left(p^{RD}, \bar{e}^{RD}, F^{RD}, \eta, \lambda\right)}{\partial F} &\gtrless 0
\end{aligned}$$

Proposition 1 *Under reference dependence and loss aversio, the optimal penalty is equal to* $\bar{F}$, *while the optimal detection probability is greater than* 0.

Proof. Consider the first order necessary condition with respect to p. This can be rewritten as:

$$\begin{aligned}
&\left(H_0\left(\bar{e}^{RD}\right) - H_1\left(\bar{e}^{RD}\right)\right) \cdot F^{RD} \cdot \Bigg(1 + \eta \cdot (\lambda - 1) \cdot \\
&\quad \cdot \left(1 - 2p^{RD} \cdot \left(2 - H_0\left(\bar{e}^{RD}\right) - H_1\left(\bar{e}^{RD}\right)\right)\right)\Bigg) = \frac{1}{\rho} \cdot \frac{\frac{\partial C_p\left(p^{RD}\right)}{\partial p}}{g\left(\theta^{**}\left(p^{RD}, \bar{e}^{RD}, \bar{F}, \eta, \lambda\right)\right)}
\end{aligned}$$

Notice that when $p^{RD} = 0$ and $F^{RD} > 0$, the left hand side is positive and decreasing in p, whereas the right hand side is equal to 0 and increasing in p. As a result, $F^{RD} > 0$ implies $p^{RD} > 0$. Since F can be set equal to $\bar{F} > 0$ at no cost, we conclude that $p^{RD} > 0$. Moreover, since we are looking for an internal solution, the following condition must be true:

$$p^{RD} < \frac{1 + \eta(\lambda - 1)}{2 \cdot \eta(\lambda - 1) \cdot \left(2 - H_0\left(\bar{e}^{RD}\right) - H_1\left(\bar{e}^{RD}\right)\right)}. \tag{7}$$

Now, consider the first order conditon with respect to $\bar{F}$ and observe that its sign depends on:

$$\begin{aligned}
\frac{\partial L\left(p^{RD}, \bar{e}^{RD}, F^{RD}, \eta, \lambda\right)}{\partial F} &= p^{RD} \cdot \left(H_0(\bar{e}) - H_1(\bar{e})\right) \cdot \\
&\quad \cdot \left(1 + \eta \cdot (\lambda - 1) \cdot \left(1 - p^{RD} \cdot \left(2 - H_0\left(\bar{e}^{RD}\right) - H_1\left(\bar{e}^{RD}\right)\right)\right)\right)
\end{aligned}$$

By (7)

$$\frac{\partial L\left(p^{RD},\bar{e}^{RD},F^{RD},\eta,\lambda\right)}{\partial F}>p^{RD}\cdot\left(H_0\left(\bar{e}\right)-H_1\left(\bar{e}\right)\right)\cdot\left(\frac{1}{2}+\frac{\eta\left(\lambda-1\right)}{2}\right)>0$$

Thus, $p^{RD}>0$ implies $F=F^{RD}$. ∎

Thus, even if individuals exhibit reference dependence and loss aversion, the penalty F will be set at the highest possible level ($\bar{F}$) exactly as in our benchmark case. Instead, the next proposition shows how these behavioral biases yield to a change in the standard of proof that favoring type I errors against type II ones.

Proposition 2 *Under reference dependence and loss aversion, the optimal judicial system entails a standard of proof* $\bar{e}^{RD}$ *such that* $h_1\left(\bar{e}^{RD}\right)>h_0\left(\bar{e}^{RD}\right)$*. Thus, by Assumption 1,* $\bar{e}^{RD}>\bar{e}^{RN}$*.*

Proof. By the first order necessary condition, the optimal standard of proof, $\bar{e}^{RD}$, solve:

$$\frac{\partial L\left(p^{RD},\bar{e}^{RD},F^{RD},\eta,\lambda\right)}{\partial\bar{e}}=0, \tag{8}$$

or equivalently:

$$\begin{aligned}&\left(1+\eta\cdot\left(\lambda-1\right)\right)\cdot\left(h_0\left(\bar{e}^{RD}\right)-h_1\left(\bar{e}^{RD}\right)\right)-\\&\quad-\eta\cdot\left(\lambda-1\right)\cdot 2p^{RD}\cdot\left(\left(1-H_0\left(\bar{e}^{RD}\right)\right)\cdot h_0\left(\bar{e}^{RD}\right)-\left(1-H_1\left(\bar{e}^{RD}\right)\right)\cdot h_1\left(\bar{e}^{RD}\right)\right)=0\end{aligned}$$

Now, recall that $h_0\left(\bar{e}^{RN}\right)=h_1\left(\bar{e}^{RN}\right)$ and notice that Assumption 1 implies that

$$\left(1+\eta\cdot\left(\lambda-1\right)\right)\cdot\left(h_0\left(\bar{e}\right)-h_1\left(\bar{e}\right)\right)$$

is decreasing in $\bar{e}$ and equal to 0 when $\bar{e}=e^{RN}$. Furthermore,

$$-\eta\cdot\left(\lambda-1\right)\cdot 2p^{RD}\cdot\left(\left(1-H_0\left(\bar{e}^{RD}\right)\right)\cdot h_0\left(\bar{e}^{RD}\right)-\left(1-H_1\left(\bar{e}^{RD}\right)\right)\cdot h_1\left(\bar{e}^{RD}\right)\right)>0$$

if and only if:

$$\frac{1-H_1\left(\bar{e}\right)}{1-H_0\left(\bar{e}\right)}>\frac{h_0\left(\bar{e}\right)}{h_1\left(\bar{e}\right)} \tag{9}$$

By Assumption 1, the left-hand side of (9) is increasing in $\bar{e}$ and always greater than 1, whereas the right-hand one is decreasing in $\bar{e}$ and equal to 1 at $\bar{e}=\bar{e}^{RN}$. Thus, there exists a threshold $\bar{e}^{*}<e^{RN}$ such that $\frac{1-H_1(\bar{e})}{1-H_0(\bar{e})}>\frac{h_0(\bar{e})}{h_1(\bar{e})}$ whenever $\bar{e}>\bar{e}^{*}$. Furthermore, as

$\bar{e} \to 1$, $\frac{\partial L\left(p^{RD},\bar{e}^{RD},F^{RD},\eta,\lambda\right)}{\partial \bar{e}} < 0$. Thus, there must exists $\bar{e}^{RD} \in \left(\bar{e}^{RN},1\right)$ satisfying (8). Also, by Assumption 1, for every $\bar{e} < \bar{e}^{*}$, the left hand side of (8) is greater or equal than

$$\left(1+\eta\cdot(\lambda-1)\left(1-2p^{RD}\left(1-H_1\left(\bar{e}\right)\right)\right)\right)\cdot\left(h_0\left(\bar{e}\right)-h_1\left(\bar{e}\right)\right)$$

Since $e^{*} < \bar{e}^{RN}$, Assumption 1 implies $\left(h_0\left(\bar{e}\right)-h_1\left(\bar{e}\right)\right) > 0$, whereas (7) implies that:

$$1+\eta\cdot(\lambda-1)\left(1-2p^{RD}\left(1-H_1\left(\bar{e}\right)\right)\right) > 1-\frac{\left(1-H_1\left(\bar{e}\right)\right)}{\left(1-H_0\left(\bar{e}\right)\right)+\left(1-H_1\left(\bar{e}\right)\right)} > 0$$

Thus, no $\bar{e}^{RN} < \bar{e}^{*}$ can satisfy (7). ∎

Thus, if the social planner faces individuals who exhibit reference dependence and loss aversion, he will set up a judicial system for which type II errors are more frequent than type I errors. Indeed, under these assumptions, criminal behavior can be discouraged also by increasing the amount of expected losses that an individual is forced to bear when he chooses $a = 1$ instead of $a = 0$. This will be attained increasing the standard of proof and decreasing (respectively, increasing) the probability of type I (respectively, type II) errors accrodingly.

Now, consider again the optimal level of detection probability p^{RD}. To effectively, compare p^{RD} with p^{RN}, we need to impose a specific functional form on the pdf $g(\cdot)$. In particular, we will assume that $g(\cdot)$ is uniformly distributed over $[0,1]$, so that the first order condition for an internior optimum becomes:

$$\left(H_0\left(\bar{e}^{RD}\right)-H_1\left(\bar{e}^{RD}\right)\right)\cdot\Bigg(1+\eta\cdot(\lambda-1)\cdot \left(1-2p^{RD}\cdot\left(2-H_0\left(\bar{e}^{RD}\right)-H_1\left(\bar{e}^{RD}\right)\right)\right)\Bigg)-\frac{1}{\rho\cdot\bar{F}}\cdot\frac{\partial C_p\left(p^{RD}\right)}{\partial p}=0 \tag{10}$$

Under these assumption, we can prove the following proposition.

Proposition 3 *If $G(\cdot)$ is uniformly distributed in $[0,1]$, then $p^{RD}(\eta,\lambda) > p^{RN}$ for every value of η and λ if and only if $p^{RD} \leq \frac{1}{2\cdot\left(2-H_0\left(\bar{e}^{RD}\right)-H_1\left(\bar{e}^{RD}\right)\right)}$. If instead, $p^{RD} > \frac{1}{2\cdot\left(2-H_0\left(\bar{e}^{RD}\right)-H_1\left(\bar{e}^{RD}\right)\right)}$, then $p^{RD}(\eta,\lambda)$ will be lower than p^{RN} for small values of η and λ and greater than p^{RN} for higher values of the two parameters.*

Proof. By the implicit function thoerem, (10) implicitly defines p^{RD} as a continuous function of η, λ, $p^{RD}(\eta,\lambda)$. From (10), it is easy to show that if $\rho \to \infty$, then $p^{RD}(\eta,\lambda) \to \min\left\{\frac{1+\eta\cdot(\lambda-1)}{2\cdot\left(2-H_0\left(\bar{e}^{RD}\right)-H_1\left(\bar{e}^{RD}\right)\right)},1\right\}$, while if $\rho \to 0$ then $p^{RD}(\eta,\lambda) \to 0$. Intuitively, as the social harm caused gets larger, the social planner will set higher and higher detection probabilities to increase the deterrance power of the judicial system. Furthermore, if either $\eta = 0$ or $\lambda = 1$, condition (10) would become:

$$\left(H_0\left(\bar{e}^{RN}\right) - H_1\left(\bar{e}^{RN}\right)\right) = \frac{1}{\rho \cdot \bar{F}} \cdot \frac{\partial C_p\left(p^{RN}\right)}{\partial p}$$

and $p^{RD}(0,\lambda) = p^{RD}(\eta,1) = p^{RN}$. We will focus on the analysis with respect to η; the one with respect to λ is similar and omitted.If $p^{RD}(\eta) \leq \frac{1}{2\cdot(2-H_0(\bar{e}^{RD})-H_1(\bar{e}^{RD}))}$, then the implicit function theorem implies $\frac{\partial p^{RD}(\eta)}{\partial \eta} > 0$. Indeed, the derivative of the the left hand side of (10) with respect to η can be written as:

$$\left(H_0\left(\bar{e}^{RD}\right) - H_1\left(\bar{e}^{RD}\right)\right) \cdot (\lambda - 1) \cdot \left(1 - 2p^{RD} \cdot \left(2 - H_0\left(\bar{e}^{RD}\right) - H_1\left(\bar{e}^{RD}\right)\right)\right) + \\ + (1 + \eta \cdot (\lambda - 1)) \cdot \left(h_1\left(\bar{e}^{RD}\right) - h_0\left(\bar{e}^{RD}\right)\right) \cdot \frac{\partial \bar{e}^{RD}}{\partial \eta} > 0$$

(to get this expression, we can apply (8) twice). Since $p^{RD}(\eta) \leq \frac{1}{2\cdot(2-H_0(\bar{e}^{RD})-H_1(\bar{e}^{RD}))}$, this expression is positivie. On the contrary, the derivative of the left-with of (10) with respect to p^{RD} is negative. On the contrary, if $p^{RD}(\eta) > \frac{1}{2\cdot(2-H_0(\bar{e}^{RD})-H_1(\bar{e}^{RD}))}$, $\frac{\partial p^{RD}(\eta)}{\partial \eta}$ will be negative for values of η small enough and positive for higher values of η.

Now, suppose that $p^{RN} \leq \frac{1}{2\cdot(2-H_0(\bar{e}^{RN})-H_1(\bar{e}^{RN}))}$. Then, we will show that $p^{RD}(\eta) \leq \frac{1}{2\cdot(2-H_0(\bar{e}^{RD}(\eta))-H_1(\bar{e}^{RD}(\eta)))}$, for every $\eta > 0$. Suppose not. Then, by continuity we can find $\eta^* > 0$ such that $p^{RD}(\eta) < \frac{1}{2\cdot(2-H_0(\bar{e}^{RD}(\eta))-H_1(\bar{e}^{RD}(\eta)))}$ if $\eta \leq \eta^*$ and $p^{RD}(\eta) > \frac{1}{2\cdot(2-H_0(\bar{e}^{RD}(\eta))-H_1(\bar{e}^{RD}(\eta)))}$ for some open interval $\eta \in (\eta^*, \eta')$. By continuity $p^{RD}(\eta^*) = \frac{1}{2\cdot(2-H_0(\bar{e}^{RD}(\eta^*))-H_1(\bar{e}^{RD}(\eta^*)))}$. Then by (10):

$$\left(H_0\left(\bar{e}^{RD}(\eta^*)\right) - H_1\left(\bar{e}^{RD}(\eta^*)\right)\right) = \frac{1}{\rho \cdot \bar{F}} \cdot \frac{\partial C_p\left(p^{RD}(\eta^*)\right)}{\partial p}$$

Observe that if $\eta^* > 0$, our previous results imply both

$$\left(H_0\left(\bar{e}^{RN}\right) - H_1\left(\bar{e}^{RN}\right)\right) > H_0\left(\bar{e}^{RD}(\eta^*)\right) - H_1\left(\bar{e}^{RD}(\eta^*)\right)$$

and

$$\frac{\partial C_p\left(p^{RD}(\eta^*)\right)}{\partial p} > \frac{\partial C_p\left(p^{RN}\right)}{\partial p}$$

Indeed, the former inequality follows from the fact that $\bar{e}^{RN}$ maximizes the accuracy of the judicial system, while the latter one by noticing that $C_p(\cdot)$ is convex and that we have just shown that $p^{RD}(\eta)$ is increasing in η in the interval $[0, \eta^*]$. Thus, if $p^{RN} \leq$

$\frac{1}{2\cdot(2-H_0(\bar{e}^{RN})-H_1(\bar{e}^{RN}))}$, $p^{RD}(\eta) \leq \frac{1}{2\cdot(2-H_0(\bar{e}^{RN}(\eta))-H_1(\bar{e}^{RN}(\eta)))}$ for every $\eta > 0$.

On the contrary, if $p^{RD}(\eta) > \frac{1}{2\cdot(2-H_0(\bar{e}^{RD})-H_1(\bar{e}^{RD}))}$, the first term in the expression

$$\left(H_0\left(\bar{e}^{RD}\right) - H_1\left(\bar{e}^{RD}\right)\right)\cdot(\lambda-1)\cdot\left(1-2p^{RD}(\eta)\cdot\left(2-H_0\left(\bar{e}^{RD}(\eta)\right)-H_1\left(\bar{e}^{RD}(\eta)\right)\right)\right)+ \\ +\left(1+\eta\cdot(\lambda-1)\right)\cdot\left(h_1\left(\bar{e}^{RD}(\eta)\right)-h_0\left(\bar{e}^{RD}(\eta)\right)\right)\cdot\frac{\partial\bar{e}^{RD}(\eta)}{\partial\eta}$$

will be negative, while the second will be positive. Moreover, the first (respectively, second) term will dominate for low (respectively, high) values of η. Furthermore, replicating, the reasoning used in the case $p^{RN} \leq \frac{1}{2\cdot(2-H_0(\bar{e}^{RN})-H_1(\bar{e}^{RN}))}$, we can conclude that if $p^{RD}(\eta) \leq \frac{1}{2\cdot(2-H_0(\bar{e}^{RD}(\eta))-H_1(\bar{e}^{RD}(\eta)))}$, the same inequality will hold for every $\eta' > \eta$. Therefore, by the implicit function theorem, p^{RD} is decreasing (increasing) in η for small (high) values of η. ∎

Thus, if the benefits from engaging in the criminal activity are uniformly distributed, $p^{RD}(\eta,\lambda)$ will be higher than p^{RN} for every value of η and λ if and only if:

$$p^{RN} = p^{RD}(0,\lambda) = p^{RD}(\eta,1) \leq \frac{\frac{1}{2}}{\left(1-H_0\left(\bar{e}^{RN}\right)\right)+\left(1-H_1\left(\bar{e}^{RN}\right)\right)} \tag{11}$$

Intuitively, (11) strengthen (6) by requiring an even lower inspection probability. In particular, (11) holds either if the social harm caused by crime, ρ, is sufficiently low (in which case, p^{RN} will be small) or if the optimal judicial system in the benchmark case entails relatively few convictions ($1-H_0\left(\bar{e}^{RN}\right)+1-H_1\left(\bar{e}^{RN}\right)$ small). Since the standard of proof in the benchmark case, is chosen to maximize the accuracy of the judicial system, the previous condition can be restated by saying that low levels of incriminating evidence are uanble to effectively distinguish between guilty and innocent individuals. Furthermore, (11) is always satisfied if $H_0\left(\bar{e}^{RN}\right)+H_1\left(\bar{e}^{RN}\right) \geq \frac{3}{2}$. Given Assumption 1, a sufficient condition for the previous inequality is $H_1\left(\bar{e}^{RN}\right) > \frac{3}{4}$.**THINK MORE ABOUT THE INTUITION.**

The figures below summarize our analysis. For the special case in which $G(\theta)$ is uniformly distributed in $[0,1]$, $C_p(x) = c_p\cdot\frac{x^2}{2}$, $H_0(e) = \int_0^e 2(1-x)\,dx$, $H_1(e) = \int_0^e 2x dx$, $\bar{F} = 1$ and $\lambda = 2$, we plot p^{RD} and $\bar{e}^{RD}$ as a function of η if (11) holds (Figure 4 and 5) and if (11) fails (Figure 6 and 7).[4] Finally, Figure 8 and 9 compares p^{RN} and $\bar{e}^{RN}$ with p^{RD} and $\bar{e}^{RD}$ for different values of the social harm ρ.[5]

[4] In particular, Figure 4 and 5 are obtained setting $\rho = 0.9$, while Figure 6 and 7 setting $\rho = 1.1$. Since, under these specific functional forms $\left(1-H_0\left(\bar{e}^{RN}\right)\right)+\left(1-H_1\left(\bar{e}^{RN}\right)\right) = 1$, (11) can be written as $p^{RN} \leq \frac{1}{2}$, which, by the analysis in Section 2, holds if and only if $\rho \leq 1$.

[5] In this special case, we set $\eta = 0.5$ and $\lambda = 2$.

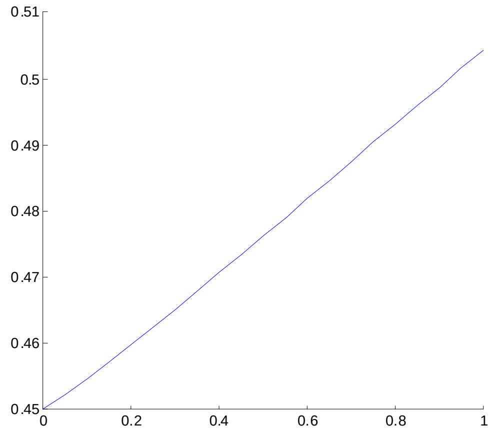


Figure 4: $p^{RD}(\eta)$ when (11) holds.

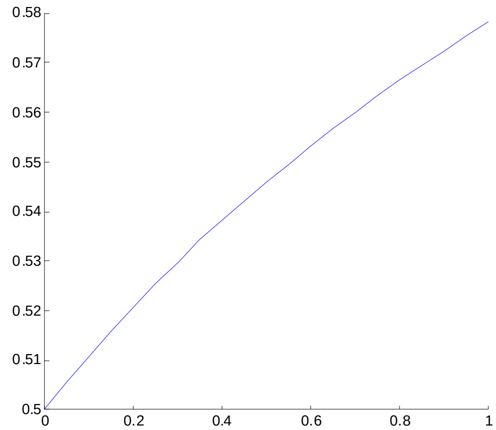


Figure 5: $\bar{e}^{RD}(\eta)$ when (11) holds.

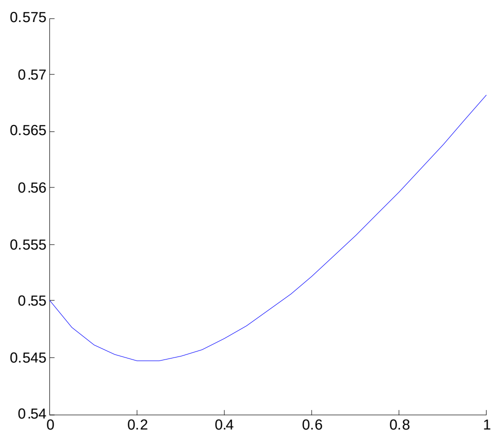


Figure 6: $p^{RD}(\eta)$ when (11) fails.

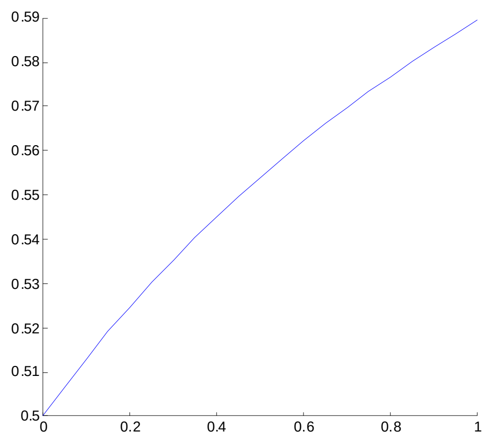
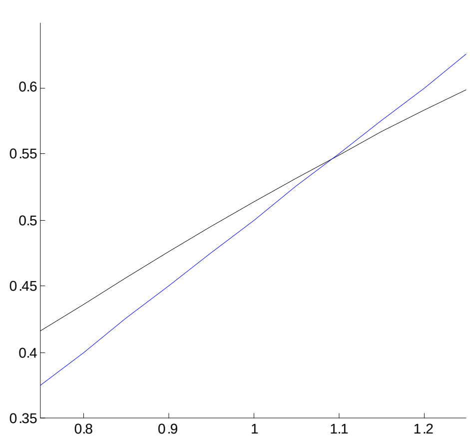
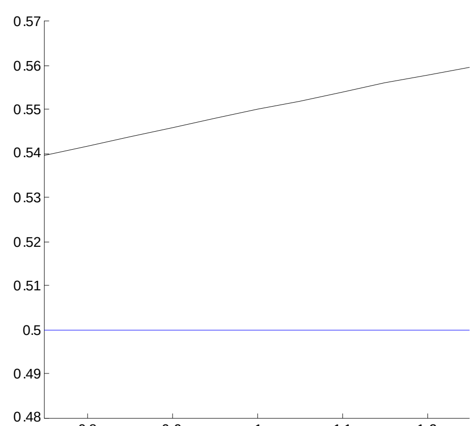


Figure 7: $\bar{e}^{RD}(\eta)$ when (11) fails





p^{RN} (blue) and p^{RD} (black) as a

$\bar{e}^{RN}$ (blue) and $\bar{e}^{RD}$ (black) as a