

Lear Research Paper

The Enforcement of Imperfect Rules

Paolo Buccirossi

September 2010

The Enforcement of Imperfect Rules*

Paolo Buccirossi[†]

September 2010

Abstract

This paper examines the optimal sanction for rules that are imperfect in that they are either overinclusive, as they prohibit an action that in some circumstances is beneficial, or underinclusive as they allow agents to undertake alternative conducts that are harmful, or both. The paper clarifies why this notion of imperfection differs from the notion of over- and underdeterrence and from that of legal errors. Finally it shows that when rules are imperfect the optimal sanction is lower than the optimal sanction for a perfect rule, both if the rule is overinclusive and if it is underinclusive.

Keywords: *Public enforcement of law, optimal sanction, rules imperfection, legal errors, deterrence.*

JEL Classifications: **K42**

*I wish to thank Tomaso Duso, Elisabetta Iossa and Luca Aguzzoni for their helpful comments. All errors and imperfections remain mine.

[†]Lear - Laboratorio di economia, antitrust, regolamentazione; via del Banco di S. Spirito 42, 00186 Rome, Italy; E.mail: paolo.buccirossi@learlab.com

1 Introduction

Rules are commands that aim to prevent socially undesirable behavior. However, often the design of a rule is imperfect in that it prohibits also conducts that, at least in some circumstances, are not harmful or allows conducts that are harmful, again at least in some circumstances. To pursue its end, a rule must be enforced, that is there must exist a system that looks for breaches of the rule and sanctions the violators. The optimal enforcement system of legal commands is the subject of a vast economic literature, spurred by Becker (1968). The aim of this paper is to study how the optimal enforcement changes when the design of the rule is imperfect. Rules' imperfection is not to be confused with the notions of inappropriate deterrence (overdeterrence and underdeterrence) and of legal errors (false convictions, or type I errors, and false acquittals, or type II errors). The next section provides the definition of rules' imperfection and clarifies in what respects it differs from the other two flaws of a legal system. In this introduction it suffices to say that a rule is imperfect if the prohibition includes conducts that we would rather allow (overinclusion) and/or if the rule permits conducts that we would prefer to prohibit (underinclusion). The main result of the paper is that if the rule is imperfect, the expected sanction need to be lowered to remain optimal, both when the rule is overinclusive and when it is underinclusive. This contrasts with the way to adjust the expected sanction to restore an optimal enforcement system when the level of deterrence is inadequate or when legal errors are possible.

The paper is organized as follows. Section 2 provides a taxonomy of the possible three flaws of the legal system, relates the paper to the relevant literature and describes its motivations. Section 3 presents the model. Section 4 contains the analysis: Section 4.1 characterizes the optimal sanction for an overinclusive rule; Section 4.2. deals with the underinclusive rule; Section 4.3. analyzes a rule that is both overinclusive and underinclusive. Section 5 concludes.

2 Taxonomy, literature and motivations

In this section I will: 1) clarify what I mean by rule's imperfection, 2) explain in what regards this notion differs from other notions of "imperfection" relevant for the enforcement of rules, 3) show that in the existing literature not all the implications of rules' imperfection are adequately examined; and 4) demonstrate that the type of imperfection addressed in this paper is an important legal and social phenomenon.

In order to clarify the notion of rules' imperfection, it is necessary to split the social consequences of a conduct into two components: the "internal effect", that is the gain the agent that undertakes the conduct derives from it, and the "external effect", that is the net welfare effect of that behavior on any other agent. In this paper a rule will be said to be imperfect if the conduct that it prohibits in some circumstances yields a positive external effect (*overinclusion*), or if, in some circumstances, the rule allows some different behavior even if this

alternative conduct yields a negative external effect (*underinclusion*).¹ This type of imperfection was already noted by Ehrlich and Posner (1974), Kennedy (1976), and extensively discussed by Schauer (1991).

The notion of imperfection presented above differs from other types of "imperfections" considered in the literature of the public enforcement of law. First, even if a rule is perfect in the sense just defined, the enforcement of the rule impedes a socially efficient outcome if it prevents a conduct whose positive internal effect (violator's gain) is larger than the negative external effect (victims' harm). In this case the enforcement of a (perfect) rule leads to *overdeterrence*, as it discourages conducts that are harmful *but* socially efficient. The enforcement of a rule can also lead to *underdeterrence*, if it does not deter conducts that are harmful *and* socially inefficient as the private gain is lower than the external harm.

Second, the proposed notion of imperfection differs from the concept of judicial, or legal errors. A judicial error occurs either if someone is convicted for breaching a rule, when in fact he did not (*false conviction* or *type I error*) or if someone is not convicted for breaching a rule, when in fact he did (*false acquittal* or *type II error*). If an agent is convicted when he does undertake conduct *a*, that the rule prohibits, a judicial error does not occur, even if conduct *a* has a positive external effect (overinclusion) or if it is harmful but socially efficient (overdeterrence). Similarly, if an agent is acquitted when she undertakes conduct, *a'*, that the rule permits, and not the conduct *a*, that is prohibited, a judicial error does not occur, also if conduct *a'* has a negative external effect (underinclusion) or if conduct *a'* is harmful and socially inefficient (underdeterrence).

The three possible flaws of the legal system are clearly distinct. Nonetheless, the imperfection of rules retains some of the features of both inappropriate deterrence and of judicial errors and may be confused with both of them if the latter are defined in less precise terms. Indeed, as in a situation of *overdeterrence*, an *overinclusive* rule may prevent conducts that improve social welfare. Analogously, an *underinclusive* rule may allow conducts that are socially inefficient, as in the case of *underdeterrence*. An *overinclusive* rule may lead to the conviction of an agent for undertaking a harmful action when in fact he did not (as in a *type I error*). And an *underinclusive* rule may exonerate an agent when in fact he did commit a harmful action (as in a *type II error*).² Indeed, rules' imperfection, over- and underdeterrence and the legal errors are all defects of

¹I will call "harmful" a conduct with a negative external effect and "beneficial" a conduct with a positive external effect. These names do not indicate the impact of the conduct on total welfare as they neglect the welfare of the agent that undertakes it.

²It may seem that the two notions of imperfection and legal errors coincide when the distinction between legal and illegal conducts is defined by a *standard* rather than a *rule*. The standard may declare illegal an act if and only if it causes an external harm. In this case, however, the legal command is not imperfect by definition and only legal errors are possible. It must be added that very rarely, if ever, a legal command is stated in terms of a pure standard and that, even in these exceptional cases, judges tend to elaborate rules that translate the standard in more precise and concrete prescriptions. The analysis contained in this paper applies also to these judge-made rules.

the legal system. However, the above discussion clarifies that they are different in nature and call for distinct analyses since, as shown also in this paper, they can have very different implications for the design of an optimal enforcement system.

Does the existing literature cover adequately all the three types of flaws? Appropriate deterrence is the central theme of the literature on the optimal enforcement of rules. It has been thoroughly studied in hundreds of articles.³ The present paper is part of this literature. The implications of judicial errors have been extensively examined in the contributions of Png (1986), Polinsky and Shavell (1989), and Kaplow and Shavell (1994).

Rules' imperfection has been addressed, mostly informally, by Ehrlich and Posner (1974) and Kennedy (1976). More recently, there has been an increasing interest about rules' imperfection in the field of antitrust, and especially in the analysis of exclusionary conducts (Beckner III and Salop, 1999; Hylton and Salinger, 2001; Evans and Padilla, 2005). However, the earlier and the latter contributions focus almost exclusively on the optimal design of a rule, taking into account (among other things) the possibility that the legal commands may be overinclusive or underinclusive or both.⁴ The more recent contributions also examine how the optimal adjudication is affected by the inclusion problems and in particular discuss the optimal standard of proof for imperfect rules. However, a formal analysis of how the imperfection of rules (the inclusion problem) alters the optimal enforcement has not been developed so far.⁵

Rules' imperfection is a pervasive phenomenon. Some examples can be drawn from antitrust law. A firm with a significant market power by reducing its price may infringe a (judge-made) rule according to which a price below a pre-defined cost benchmark is predatory and therefore prohibited by antitrust law.⁶ This rule is likely to be overinclusive as, in some circumstances, the decision by a firm to charge a price below the cost benchmark defined by the rule does not have any exclusionary effect and is justified by efficiency reasons. When this occurs the rule prohibits a conduct that would have a positive external effect, as it would improve consumer welfare more than reducing competitors' profits.

³Among them the most noteworthy contributions are Polinsky and Shavell (1979, 1984, 1991, 1992), Shavell (1987a, 1987b, 1991) Kaplow (1990a, 1990b, 1992) and Mookherjee and Png (1994). For a more detailed list of references, see the comprehensive surveys of Garoupa (1997), Mookherjee (1997) and Polinsky and Shavell (2000).

⁴Beckner III and Salop (1999), Hylton and Salinger (2001) and Evans and Padilla (2005) tend to confuse the notion of rules' imperfection with that of judicial or legal errors. For instance, Hylton and Salinger (2001) describe a rule with a two-way matrix in which the cells contain the following combinations: a) illegal/harmful; b) illegal/not harmful; c) legal/harmful; and d) legal/not harmful. Then they refer to b) as a case of "false conviction" (while in fact it is a case of overinclusion) and to c) as a case of "false acquittal" (while, in fact it is a case of underinclusion).

⁵Kaplow (1995) investigates the enforcement of rules concerning behavior that may have different consequences. Yet, he considers the case in which two types of acts cause different harms so that even if a rule does not distinguish between them, it is not overinclusive or underinclusive in the meaning of this paper.

⁶Bolton et al (2000) provide a comprehensive discussion of the cost benchmarks used in assessing alleged predatory behavior. They also point out the inclusion problem of these cost benchmarks.

Even the per se prohibition of hard core cartels is an overinclusive rule. Indeed, several papers show that, in some circumstances, by fixing prices, firms make also consumers better off.⁷ The cost based rule on predatory pricing is likely to be also underinclusive. In some cases, a dominant firm may effectively exclude rivals by charging a price above the predefined cost benchmark. This conduct may harm consumers as, once the rival has been excluded, the dominant firm has the ability to increase the price above the level that would have prevailed otherwise. The prohibition of cartels does not cover some facilitating practices that may allow the members of an oligopoly to reach and sustain a collusive equilibrium.⁸ Hence, even if the rule prohibits overt communication among competitors, it fails to prevent other forms of coordination that, although legal, in some circumstances are apt to yield similar socially harmful consequences. These are just a few examples and many others could be found in any area of law.

What should be done when rules are imperfect? An obvious response is to write better (that is less imperfect) rules. However, this comes at a cost, and, as pointed out by Ehrlich and Posner (1974), and Kennedy (1976) the cost of improving the design of a rule, after a certain point, overcomes the benefits, so that a certain degree of imperfection is optimal, also if one was able to write perfect rules. Moreover, Schauer (1991) has pointed out that rules are inherently imperfect since it is impossible to foresee all the conceivable consequences of a conduct in all possible real world conditions, some of which are necessarily unpredictable when the rule is written.

Since the legislator, or the courts, cannot make rules perfect, their enforcement must take into account that they are imperfect. The enforcement policy is a complex system which encompasses many elements: the level of resources spent to detect possible infringements, the combination of monitoring and investigation activities, the type and level of sanctions for apprehended and convicted violators, and many others. The vast literature on the economics of enforcement already mentioned has satisfactorily addressed the issues related to these policy variables. Therefore to focus on the novelties implied by the inclusion problems, I will consider the simplest setting in which the enforcer has to choose only the expected sanction. It is well known that, if agents are risk neutral, sanctions are costless, and there are no errors, the optimal expected sanction equals the harm caused by the illegal conduct as this choice yields optimal deterrence. Economists have also proved that if judicial errors are possible, both types of error call for a higher expected sanction⁹, because, for any level of the sanction,

⁷For instance, Stiglitz (1989), Fershtman and Pakes (2000), Kranton (2003), and Calzolari and Spagnolo (2005) suggest that in situations where non-contractable quality is very important, restricting price competition may improve the effectiveness of reputational forces, and increase non-contractable quality and consumer welfare. Buccirossi (2003) shows that if consumers face search costs and choose quantity on the basis of the expected price, the average price is above the monopoly level so that a price fixing agreement makes firms *and* consumers better off.

⁸See Buccirossi (2008) for a discussion of practices that may facilitate collusion and their treatment under antitrust law.

⁹See Png (1986), Polinsky and Shavell (1989), and Kaplow and Shavell (1994). Ehrlich

both errors diminish the difference between the payoff the agent obtains if he behaves legally and his payoff if he behaves illegally. Since the deterrence effect of the sanction depends on this difference and not on the sanction absolute level, the expected sanction, if errors are possible, must be increased in order to reestablish the optimal level of deterrence. In the following sections I show that the opposite result is valid when "errors" are due to an imperfect design of the rule. I prove that both if the rule is overinclusive and if it is underinclusive the optimal expected sanction is lower than the sanction for the perfect rule.

3 The model

A representative risk neutral agent¹⁰ can undertake an action, a that is prohibited by a rule. With some probability α action a is beneficial in that it yields a positive external effect, $b > 0$. With probability $1 - \alpha$ action a is harmful as it yields a negative external effect $-h$, with $h > 0$. If the action is beneficial the agent that undertakes it obtains a private gain g_b distributed according to density $m(g_b)$ and cumulative distribution $M(g_b)$ with support $[0, \infty]$. When the action is harmful the private gain is g_h with density $n(g_h)$ and cumulative distribution $N(g_h)$ again with support $[0, \infty]$. While a beneficial action is always socially desirable, a harmful action may be socially efficient or inefficient (i.e. it may be either $g_h > h$ or $g_h < h$).

If the action is harmful, with probability β the agent can pursue the same objective, possibly less efficiently from a private point of view, undertaking a different action, a' that is not prohibited. Action a' yields a negative external effect $-h'$, with $h' > 0$ and a private gain equal to δg_h where δ is a random variable, observed by the agent, with support $[0, 1]$ and distribution $f(\delta)$.¹¹ No further restrictions are imposed. Hence, the alternative action can be more harmful ($h' > h$) or less harmful ($h' < h$), more efficient for society ($\delta g_h - h' > g_h - h$) or less efficient ($\delta g_h - h' < g_h - h$). Moreover both actions may be socially desirable (if $g_h > h$ or if $\delta g_h > h'$) or undesirable (if the opposite inequalities hold).

The rule concerning conduct a may be *overinclusive*, as in some circumstances the prohibited action yields a positive external effect, and *underinclusive* as, in some circumstances, the agent can undertake an alternative conduct that

(1982) and Miceli (1991) take a different view as they assume that wrong punishments and acquittals have a direct and negative effect on social welfare.

¹⁰The agent represents k agents that can undertake the action prohibited by a rule in x_i different circumstances with $i = 1, \dots, k$. The total number of situations in which the rule could be applied is $\sum_{i=1}^k x_i$ and is normalized to 1. In this formulation the various probabilities employed in the model represent the proportion of agents/circumstances that generate a certain outcome. This interpretation is used in the comments to the analytical results with expressions like "there are some agents", "the set of agents" or "in some circumstances".

¹¹The assumption that δ is in the interval $[0, 1]$ does not impose any restriction for the purposes of this paper. Indeed, both if $\delta < 0$ and if $\delta > 1$ the rule and its enforcement are irrelevant. If $\delta < 0$ conduct a' is not adopted independently of the rule as it would lower the agent payoff in any case. If $\delta > 1$ the agent prefers action a' to action a independently of the existence of the rule and its enforcement.

yields a negative external effect but that is not prohibited. The variable α represents the degree of overinclusion and β the degree of underinclusion. Indeed, α indicates the probability that the rule prohibits a conduct that is beneficial and β indicates the probability that the rule does not prohibit a harmful conduct that agents regard as an imperfect substitute of the prohibited conduct.¹²

Social welfare, W , is the sum of the internal and the external effects of the action undertaken by the agent. If the agent does not undertake any action (neither a nor a') $W = 0$. The enforcer problem is to choose the level of the expected (costless) sanction that maximizes the expected $W(s)$, under the constraint $s \geq 0$. If this constraint is binding, we can say that society is better off if the rule is not enforced, or, equivalently, if the rule is abolished. If an internal solution exists we will say that "the rule is enforced". We assume that $W(s)$ is twice differentiable and that $W'' < 0$ which implies that $W(s)$ is concave and that the first order condition (FOC) is necessary and sufficient to identify a maximum.

The model neglects the enforcement and the sanction costs as these issues have been already explored in other contributions and the current model does not alter the previous analyses. Hence s can be thought as the optimal combination of monitoring and policing expenditure and the type and level of the sanction so as to minimize the enforcement cost associated to any value of the expected sanction. Therefore, for the sake of exposition, in what follows I refer to the expected sanction simply as the "sanction".

4 Analysis

I will first examine the two inclusion problems separately in order to understand their specific implications, and then both problems together. In section 4.1 I consider the optimal sanction for an overinclusive rule. In section 4.2 I examine the case of an underinclusive rule. In section 4.3 I analyze a rule that is both overinclusive and underinclusive. The aim of the analysis is to define the optimal sanction, to identify the factors that affect its level, and, finally, to verify if some intuitive conditions are necessary and/or sufficient to make the enforcement of an imperfect rule desirable.

4.1 Overinclusive rules

In this section we assume that $\alpha \geq 0$ and $\beta = 0$. The agent undertakes a if and only if its private gain (g_b, g_h) is above the sanction. Therefore, the expected social welfare is:

¹²The model is framed in terms of a positive action prohibited by the rule. However it can be interpreted so as to describe a rule that prescribes a certain positive action. In this case, a is the conduct of an agent that does not obey the prescription and a' is an alternative action that is considered compliant with the rule even if it causes an external harm. All the other variables in the model keep their meaning as defined in the main text.

$$W = \alpha \int_s^\infty (g_b + b) m(g_b) dg_b + (1 - \alpha) \int_s^\infty (g_h - h) n(g_h) dg_h.$$

The optimal sanction is characterized in the following proposition.

Proposition 1 *If the rule is enforced, the optimal sanction for an overinclusive rule is a weighted average of the negative external effect (harm) produced by the action when it is harmful and its positive external effect (benefit) when it is beneficial. The optimal sanction is non-increasing in the degree of overinclusion of the rule and in the benefit produced by the action when beneficial, and non-decreasing in the harm caused by the action when harmful.*

Proof. Let

$$L = W(s) + \lambda s$$

be the Lagrangian function. If the non-negative constraint is not binding, the FOC is

$$\begin{aligned} \frac{\partial L}{\partial s} &= -(s+b)\alpha m(s) - (s-h)(1-\alpha)n(s) + \lambda = 0 \\ \Rightarrow -s + \frac{(1-\alpha)n(s)}{(\alpha m(s) + (1-\alpha)n(s))}h - \frac{\alpha m(s)}{(\alpha m(s) + (1-\alpha)n(s))}b + \lambda &= 0 \end{aligned}$$

which, when $\lambda = 0$, becomes

$$s^* = (1 - r(s^*))h - r(s^*)b$$

where

$$r(s) = \frac{\alpha m(s)}{\alpha m(s) + (1-\alpha)n(s)}.$$

Since $m(s), n(s) \geq 0$ for any s and $0 \leq \alpha \leq 1$ we have that $0 \leq r(s) \leq 1$ for any s . This proves the characterization of the optimal sanction. Let $G(s, \alpha, b, h)$ be the implicit function defined by the FOC, when $\lambda = 0$. According to the implicit function theorem, since $\partial G/\partial s < 0$ by assumption, the sign of $\partial s^*/\partial x$, where $x = \alpha, b, h$, is the same as the sign of $\partial G/\partial x$. Therefore we have

$$\frac{\partial G}{\partial \alpha} = -\frac{n(s)m(s)(h+b)}{(\alpha m(s) + (1-\alpha)n(s))^2} \leq 0,$$

$$\frac{\partial G}{\partial b} = -r(s) \leq 0,$$

and

$$\frac{\partial G}{\partial h} = 1 - r(s) \geq 0.$$

These derivatives prove the comparative statics statements of proposition 1. ■

The first implication of proposition 1 is that the optimal sanction for an overinclusive rule is lower than the optimal value of the sanction for a perfect

rule. The reduction of the sanction reflects to a certain extent the external social benefit that may stem from the prohibited behavior. The measure of this reduction depends on $r(s)$ as defined in the proof of proposition 1. Therefore, an economic interpretation of $r(s)$ is necessary.

The imposition of a sanction deters both harmful and beneficial actions. For a given s , the probability that a beneficial action is deterred is $\alpha M(s)$, the probability that a harmful action is deterred is $(1 - \alpha) N(s)$ and the probability that any illegal action (either beneficial or harmful) is deterred is $\alpha M(s) + (1 - \alpha) N(s)$. The marginal effect of a variation of s on the probability that an illegal beneficial action is deterred is $\alpha m(s)$, for an illegal harmful action is $(1 - \alpha) n(s)$ and for any illegal action is $\alpha m(s) + (1 - \alpha) n(s)$. Hence $r(s)$ represents, given the marginal deterrence effect of s , the probability that the additional deterred actions are beneficial. For instance, consider a given value, s_0 . If $r(s_0) = .5$, increasing marginally s from s_0 would deter a certain number of additional illegal action, half of which will be beneficial. If $r(s_0) = .25$ only a quarter of the additional deterred action are beneficial.

Remark 1. If the private gains of the beneficial action are distributed over $[h, \infty]$ the optimal sanction equals the harm as in the traditional analysis. Formally, this is proved by the fact that $m(s) = 0$ for any $s \leq h$. Hence, $r(s) = 0$ and $s^* = h$. Intuitively, if the beneficial conduct yields at least a private gain equal to the sanction this behavior is not deterred at all, as all agents that are in the position to undertake such a beneficial action are (weakly) better off if they do so for any $s \leq h$. Note that this does not mean that the rule is perfect, as agents that undertake the beneficial action still face a positive probability to be convicted and bear a sanction.

The other novelty of the characterization of the optimal sanction presented in proposition 1 is that the optimal sanction depends also on the distribution of the private gains and not only on the external effects. The description of the mechanism behind the identification of the optimal sanction for an overinclusive rule explains why the distributions of the private gains matter. The reason is not (as one may think) that the enforcer cares about the contribution of the private gain to the social welfare, but that the distribution of the internal effects influence the way the sanction modifies the composition of the deterred illegal actions. However, a simple example shows that the way the distributions of private gains affects the optimal sanctions is not always self-evident or intuitive.

Example 1. Suppose that for both types of illegal actions the private gains are uniformly distributed. Beneficial actions generates private gains distributed over $[0, \bar{g}_b]$ and the harmful action yields private gains distributed over $[0, \bar{g}_h]$. We have that $m(s) = \frac{1}{\bar{g}_b}$ and $n(s) = \frac{1}{\bar{g}_h}$ for any s . If, ceteris paribus, \bar{g}_b increases, so that the distribution of the internal effect of a beneficial action is distributed according to the new density $m'(g_b)$ such that $m'(s) < m(s)$ for any s , the optimal sanction (if the rule is enforced) increases. On the contrary, if \bar{g}_h increases, so that $n'(s) < n(s)$ for any s , the optimal sanction (if the rule is enforced) decreases. Note that both types of changes in the distribution of the private gains improve the expected total welfare stemming from the illegal

conduct. Moreover, given the value of the sanction, an increase in \bar{g}_b makes more likely that the illegal conduct that will be actually undertaken, given a certain sanction, is beneficial, whereas an increase in \bar{g}_h determines the opposite. The explanation of this results is that what matters for the determination of the optimal sanction is its relative deterrence effects on beneficial and harmful actions. If \bar{g}_b increases, ceteris paribus, a higher sanction has an impact on the decision to undertake a beneficial action less strong than the impact on the decision to undertake a harmful action. Since at the margin the cost of impeding a conduct that yields positive external effects is lower than the benefit of deterring harmful conducts, it is optimal to increase the sanction. The opposite occurs if \bar{g}_h decreases, because in this case an increase in the sanction has a stronger deterrence effect at the margin on beneficial conducts than on those that cause a harm.

The final proposition of this section examines when an overinclusive rule has to be enforced.

Proposition 2 *If the distribution of private gains is the same for harmful and beneficial actions, then the existence of a negative expected external effect in the absence of the rule is a necessary and sufficient condition for preferring the enforcement of the imperfect rule to the abolition of the rule. In all the other cases, the expected external effect of the behavior targeted by the rule does not provide neither a necessary nor a sufficient condition for choosing to keep enforcing the rule.*

Proof. The optimal sanction when $m(s) = n(s)$ for any s , is $s^* = \max\{0, (1 - \alpha)h - \alpha b\}$, so that s^* is positive if and only if the expected net external effect is negative. Two numerical examples prove the second part of the proposition. Suppose that $\alpha = 0.5$, that the private gains of both types of action are uniformly distributed as described in Example 1, that $h = 1$ and that $b = 2$. Without any enforcement the expected net external effect is positive. Therefore it seems inappropriate to prohibit a conduct that produces these consequences. However, simple calculus shows that enforcing this rule, i.e. imposing an expected positive sanction is optimal if $\bar{g}_b > 2\bar{g}_h$. For instance, if $\bar{g}_b = 3$ and $\bar{g}_h = 1$, simple calculus proves that it would be optimal to have a sanction equal to 0.25 and that this would increase total welfare from 1.5 to 1.5208. Now change the previous example with $h = 2$ and $b = 1$. If the rule is not enforced, the expected external effect is -0.5 . If $\bar{g}_b < \bar{g}_h/2$ social welfare is decreasing in s for $s \geq 0$. For instance if $\bar{g}_b = 1$ and $\bar{g}_h = 3$, $\partial W(s)/\partial s = -0.1\bar{6} - 0.\bar{6}s < 0$ for any $s \geq 0$. Therefore, the constraint $s \geq 0$ is binding which means that the rule should be repealed.

■

Proposition 2 shows that it may be desirable to enforce an imperfect rule even if the net expected external effect is positive, (i.e. even if the expected benefits of the beneficial action outweighs the expected harm of the harmful action) and that it may desirable to repeal a rule even if the expected harm of the prohibited conduct is higher than its expected benefit.

4.2 Underinclusive rules

In this section we assume that $\alpha = 0$ and $\beta \geq 0$. For the sake of exposition I remove the subscript h and write $g = g_h$. The agent has to decide whether to undertake action a or a' if this alternative is available. The agent that undertakes a obtains an expected payoff of $g - s$ while the payoff stemming from action a' is δg . Since δg is non-negative, all agents whose private gain is not larger than the expected sanction prefer action a' . An agent with $g > s$ undertakes action a if $\delta < \frac{g-s}{g}$ and action a' if $\delta \geq \frac{g-s}{g}$. With probability $1 - \beta$ the alternative action a' is not available. In this case the agent's choice is between undertaking a or no action. As usual she undertakes a if and only if her private gain is above the sanction. Hence, social welfare can be written as follows:

$$\begin{aligned} W(s) &= \beta \left(\int_0^s \int_0^1 (\delta g - h') f(\delta) d\delta n(g) dg \right. \\ &\quad + \int_s^\infty \int_0^{\frac{g-s}{g}} (g - h) f(\delta) d\delta n(g) dg \\ &\quad + \int_s^\infty \int_{\frac{g-s}{g}}^1 (\delta g - h') f(\delta) d\delta n(g) dg \\ &\quad \left. + (1 - \beta) \int_s^\infty (g - h) n(g) dg \right) \end{aligned}$$

The following proposition characterizes the optimal sanction for an underinclusive rule.

Proposition 3 *The optimal sanction for an underinclusive rule, if the rule is enforced, equals the harm caused by the prohibited action minus a share between 0 and 1 of the harm caused by the alternative but not prohibited action. The optimal sanction is non-increasing in the degree of underinclusion, and in the harm caused by the alternative action and increasing in the harm caused by the prohibited action.*

Proof. When the maximization problem has an internal solution, the FOC is:

$$\begin{aligned} 0 &= \frac{\partial L}{\partial s} = -\beta \int_s^\infty \frac{g-h}{g} n(g) f\left(\frac{g-s}{g}\right) dg \\ &\quad + \beta \int_s^\infty \frac{g-s-h'}{g} n(g) f\left(\frac{g-s}{g}\right) dg \\ &\quad - n(s)s + n(s)h + n(s)\beta s - n(s)\beta h + \lambda. \end{aligned}$$

Let

$$\widehat{\delta}(s) = \frac{g-s}{g}:$$

$$\eta(s) = \int_s^\infty \frac{f(\widehat{\delta}(s))}{g} n(g) dg;$$

and

$$q(s) = \frac{\beta\eta(s)}{\beta\eta(s) + (1-\beta)n(s)}.$$

Then, putting $\lambda = 0$, the FOC can be written as follows

$$s = h - h'q(s)$$

Since $\eta(s)$, $n(s) \geq 0$ and $0 \leq \beta \leq 1$, we have that $0 \leq q(s) \leq 1$. This proves the first part of proposition 3. Again, let $G(s, \alpha, b, h)$ be the implicit function defined by the FOC when $\lambda = 0$. Since $\partial G/\partial s < 0$, by assumption, the sign of $\partial s^*/\partial x$ where $x = \beta, h, h'$, is the same as the sign of $\partial G/\partial x$. Thus, we can compute the following derivatives:

$$\frac{\partial G}{\partial \beta} = -\frac{1}{(1-\beta)^2} \frac{\eta(s^*)}{n(s^*)} \leq 0,$$

$$\frac{\partial G}{\partial h'} = -q(s) \leq 0,$$

and

$$\frac{\partial G}{\partial h} = 1 > 0.$$

The sign of these derivatives proves the comparative statics statements of proposition 3. ■

Proposition 3 proves that also for an underinclusive rule the optimal sanction is lower than the optimal sanction that should be imposed to enforce a perfect rule. The intuition is that the sanction must be set equal to the harm caused by the prohibited conduct with respect to the external effect stemming from the alternative and not prohibited conduct. However, as for the overinclusive rule, the really important factor to take into account is the marginal impact of the sanction on the composition of the set of agents that are deterred from undertaking a . This set is formed by two groups of agents: those that have the possibility to undertake the alternative behavior a' , and those who have no alternatives. If $\beta = 1$ all agents belong to the first group whatever the level of the sanction. Therefore a modification of the sanction has no influence on the composition of the set of agents that are effectively deterred. All of them could have undertaken the alternative conduct a' with harm h' . For any deterred agent the relative harm of action a is $h - h'$. If $\beta < 1$, there exists a fraction of agents that do not have an alternative to action a . Given the value of the sanction, the probability that an agent of the first group (those with an alternative) is deterred is

$$\beta \left(\int_0^s \int_0^1 f(\delta) d\delta n(g) dg + \int_s^\infty \int_{\frac{g-s}{g}}^1 f(\delta) d\delta n(g) dg \right),$$

while the probability of deterring an agent of the second group (those without an alternative) is

$$(1 - \beta) \int_0^s n(g) dg.$$

The derivative of the first probability is $\beta\eta(s)$ and the derivative of the second probability is $(1 - \beta)n(s)$. Hence, $q(s)$ represents, given the marginal deterrence effect of s , the probability that the additional deterred actions could have been undertaken by agents who can choose the alternative conduct, a' .

Remark 2. The traditional result where the optimal sanction is equal to the harm holds in the case in which $f(\widehat{\delta}(s)) = 0$ for $s = h$. This condition holds, for instance, if action a guarantees a minimum gain and the alternative action is so privately inefficient that, even if the sanction equals the harm caused by a , no agent would undertake a' . Formally, if \underline{g} is the lower limit of the support of $n(g)$, $\bar{\delta}$ the upper limit of the support of $f(\delta)$ and $\bar{\delta} \leq 1 - h/\underline{g}$, we have that for any $s \leq h$, $f(\widehat{\delta}(s)) = 0$.

Remark 3. The social efficiency of the two actions does not enter in the determination of the optimal sanction. Hence, the optimal sanction may be positive both if a is more efficient than a' and if the opposite is true, as the following numerical example shows.

Example 2. Suppose that $\beta = 0.5$, that g is uniformly distributed over $[0, 4]$, that $h = 2$ and that δ is uniformly distributed over $[0, 1]$. Conduct a is socially more efficient than a' if $h' > 1$, whereas a' is more efficient if $h' < 1$. Denote with $s^*(h')$ the value of the sanction that solves the maximization problem, we obtain that $s^*(0.5) = 1.776$, $s^*(1.5) = 1.174$, and $s^*(2.2) = 0.527$. These results show that a positive sanction may be optimal also when the alternative conduct is less efficient and that this may be true even in the case in which a' is less efficient both because the private gain is lower *and* because the external harm is higher.

Finally, we can ascertain whether there exist necessary or sufficient conditions for enforcing an underinclusive rule.

Proposition 4 *The existence of a harm caused by the prohibited action higher than the harm caused by the harmful but legal action is a sufficient, but not a necessary condition to make the enforcement of the imperfect rule preferable to the abolition of the rule.*

Proof. Since the optimal sanction is $s^* = \max\{0, h - h'q(s)\}$, and $0 \leq q(s) \leq 1$, it is immediate to see that if $h > h'$, the non-negative constraint is never binding. Hence, $h > h'$ is a sufficient condition to have the rule enforced, but not a necessary condition as shown in Example 2. If $h \leq h'$ the non-negative constraint may become binding, depending on the degree of underinclusion and the distribution of private gains stemming from actions a and a' . ■

4.3 Over and underinclusive rule

In this section we assume that $\alpha, \beta \geq 0$. The expected social welfare is

$$\begin{aligned}
W(s) &= \alpha \int_s^\infty (g_b + b) m(g_b) dg_b + \\
&\quad (1 - \alpha) \beta \left(\int_0^s \int_0^1 (\delta g_h - h') f(\delta) d\delta n(g_h) dg_h \right. \\
&\quad + \int_s^\infty \int_0^{\frac{g-s}{g}} (g_h - h) f(\delta) d\delta n(g_h) dg_h \\
&\quad + \left. \int_s^\infty \int_{\frac{g-s}{g}}^1 (\delta g_h - h') f(\delta) d\delta n(g_h) dg_h \right) \\
&\quad + (1 - \alpha)(1 - \beta) \int_s^\infty (g_h - h) n(g_h) dg_h
\end{aligned}$$

The results of the previous sections hold in this general case, as stated in the following proposition.

Proposition 5 *The optimal sanction for a rule that is both overinclusive and underinclusive, if the rule is enforced, equals the weighted average of the negative external effect (harm) produced by the action when it is harmful and its positive external effect (benefit) when it is beneficial minus a share between 0 and 1 of the harm caused by the alternative but not prohibited action. The optimal sanction is decreasing both in the degree of overinclusion and in the degree of underinclusion of the rule; it is decreasing in the benefit produced by the action when beneficial, and in the harm caused by the alternative action; finally it is increasing in the harm caused by the prohibited action when harmful.*

Proof. Solving the FOC we obtain

$$s^* = (1 - t(s^*))h - t(s^*)b - p(s^*)h'$$

where:

$$t(s) = \frac{\alpha m(s)}{\alpha m(s) + (1 - \alpha)(\beta \eta(s) + (1 - \beta)n(s))},$$

and

$$p(s) = \frac{\beta(1 - \alpha)\eta(s)}{\alpha m(s) + (1 - \alpha)(\beta \eta(s) + (1 - \beta)n(s))}.$$

Since $m(s), n(s), \eta(s) > 0$ and $0 < \alpha < 1$, and $0 < \beta < 1$, we have that $0 < t(s) < 1$ and $0 < p(s) < 1$. The comparative static statements can be proved as in proposition 1 and 3, by using the implicit function theorem. ■

The optimal sanction can be defined in general as equal to the marginal impact of the sanction on the expected incremental net harm. So that the "harm", to which the optimal sanction should be set equal to, is qualified by several expressions. They are "marginal impact", "expected", "incremental" and "net". The first expression clarifies that the harm should be computed at the margin. This is not really a novelty. The use of "expected" refers to

the fact that the marginal impact on the harm depends on the probability of deterring: a) the prohibited harmful conduct, b) the prohibited beneficial conduct or c) the alternative harmful but not prohibited conduct. The relevant harm is "incremental" because the external effect is the specific effect of the prohibited action relative to the negative external effect of any alternative (but legal) conduct. Finally the harm is "net" because the external benefit of the prohibited action has to be taken into account and, to a certain extent, reduces the value of the harm used to set the sanction. This characterization of the optimal sanction has a general validity and applies also to any perfect rule. The optimal sanction for a perfect rule equals the harm caused by the prohibited conduct - without further qualifications - for the simple reason that: 1) the probability that the deterred conduct (if any) is the prohibited harmful conduct is 1 by assumption; 2) the external effect of the alternative legal conducts is zero by assumption; and 3) the benefit of the prohibited action is zero by assumption.

This formal analysis proves also that if there exists an alternative action that is not prohibited (as in the case of an underinclusive rule) and that yields a positive external effect, the optimal sanction might be higher than the harm caused by the prohibited harmful conduct. In this case the rule is not (necessarily) imperfect as the alternative action is not prohibited for a good reason, as it is beneficial and not harmful. Yet, the formal model may cover also this case. It is sufficient to interpret h' as the value of this benefit which enters the welfare function with a positive sign, to see that the optimal sanction might be higher than the harm caused by the prohibited and harmful action.

5 Concluding remarks

Three types of flaws can affect a system composed by a rule and its enforcement: 1) overdeterrence/underdeterrence; 2) type I/type II judicial errors; 3) overinclusion/underinclusion. The first flaw is related to the strength of the enforcement; it occurs when this enforcement is too strong or too weak, so that it either discourages conducts that are socially efficient even if they produce an external harm (overdeterrence), or does not prevent conducts that are harmful and socially inefficient (underdeterrence). The second flaw is related to the quality of the enforcement and/or of the adjudication system. It considers the possibility that an agent is convicted for breaching a rule while in fact he did not undertake the prohibited conduct (type I error) or that the agent is acquitted or not prosecuted even if he had undertaken the prohibited conduct (type II error). The third flaw is related to the design of the rule. It occurs when the rule prohibits a conduct that in some circumstances is in fact beneficial, as it determines a positive external effect (overinclusion) or when the rule permits alternative conducts that are harmful (underinclusion).

The three flaws have very different consequences on the level of enforcement (summarized by the optimal expected sanction) that maximizes social welfare. In case of overdeterrence/underdeterrence, by definition, the sanction must be adjusted so as to compensate the defect. Therefore it must be decreased in

case of overdeterrence and increased in case of underdeterrence. When the flaw derives from the possibility of judicial errors, the optimal policy is to increase the sanction both in response to a higher probability of type I errors and in response to more probable type II errors. This paper completes the picture. It shows that when the design of the rules is imperfect, the optimal way to correct the enforcement is to decrease the sanction both in case of an overinclusive rule and in the case of an underinclusive rule.

References

- [1] **Becker, Gary**, "Crime and Punishment: An Economic Approach," *Journal of Political Economy*, 1968, *76*, 169-217.
- [2] **Bolton, Patrick, Brodley Joseph, H. and Riordan Michael, H.**, "Predatory Pricing: Strategic Theory and Legal Policy," *Georgetown Law Review*, 2000, *88*, 2239-2330.
- [3] **Beckner III, C. Frederick and Salop, Steven C.**, "Decision Theory And Antitrust Rules," *Antitrust Law Journal*, 1999, *67*.
- [4] **Buccirossi, Paolo**, "A Search Model where Consumers Choose Quantity Based on Expected Price," *Journal of Industrial Economics*, 2003, *51*, 427-32.
- [5] **Buccirossi, Paolo**, "Facilitating Practices," in *Handbook of Antitrust Economics*, Buccirossi Paolo, Ed., Cambridge, MIT Press, 2008, 305-51.
- [6] **Calzolari, Giacomo and Spagnolo, Giancarlo**, "Reputation and Collusion in Procurement," *Mimeo*, University of Bologna and Stockholm School of Economics, 2005.
- [7] **Ehlich, Isaac**, "The Optimal Enforcement of Laws and Concept of Justice: A Positive Analysis," *International Review of Law and Economics*, 1982, *2*, 3-27.
- [8] **Ehrlich, Isaac and Posner, Richard, A.**, "An Economic Analysis of Legal Rulemaking," *Journal of Legal Studies*, 1974, *3*, 257-86.
- [9] **Fershtman, Chaim and Pakes, Ariel**, "A Dynamic Game with Collusion and Price Wars," *RAND Journal of Economics*, 2000, *31*, 207-36.
- [10] **Garoupa, Nuno**, "The Theory of Optimal Law Enforcement," *Journal of Economic Surveys*, 1997, *11*, 267-95.
- [11] **Hylton, Keith N. and Salinger, Michael A.**, "Tying Law and Policy: A Decision Theoretic Approach," *Antitrust Law Journal*, 2001, *69*, 486.
- [12] **Kaplow, Luois**, "Optimal Deterrence, Uninformed Individuals and Acquiring Information About Whether Acts are Subject to Sanctions," *Journal of Law, Economics and Organization*, 1990a, *6*, 93-128.
- [13] **Kaplow, Luois**, "A Note on the Optimal Use of Nonmonetary Sanctions," *Journal of Public Economics*, 1990b, *42*, 245-47.
- [14] **Kaplow, Luois**, "The Optimal Probability and Magnitude of Fines for Acts That Definitely Are Undesirable," *International Review of Law and Economics*, 1992, *12*, 3-11.

- [15] **Kaplow, Luois**, "Rules Versus Standards: An Economic Analysis", *Duke Law Journal*, 1992, 42, 557-629.
- [16] **Kaplow, Luois**, "A Model of the Optimal Complexity of Legal Rules," *Journal of Law, Economics and Organization*, 1995, 11, 150-163.
- [17] **Kennedy, Duncan**, "Form and Substance in Private Law Adjudication," *Harvard Law Review*, 1976, 89, 1685-778.
- [18] **Kranton, Rachel E.**, "Competition and the Incentive to Produce High Quality," *Economica*, 2003, 70, 385-404.
- [19] **Miceli Thomas J.**, "Optimal Criminal Procedure: Fairness and Deterrence," *International Review of Law and Economics*, 1991, 11, 3-10.
- [20] **Mookherjee, Dilip**, "The Economics of Enforcement," in *Issues in Economic Theory and Policy, Essays in Honor of Tapas Majumdar*, Amitava Bose, Mihir Rakshit, and Anup Sinha Eds., New Delhi: Oxford University Press, 1997, 202-49.
- [21] **Mookherjee, Dilip and Png I.P.L.**, "Marginal Deterrence in Enforcement of Law," *Journal of Political Economy*, 1994, 102, 1039-66.
- [22] **Padilla, A. Jorge and Evans, David S.**, "Designing Antitrust Rules for Assessing Unilateral Practices: A Neo-Chicago Approach," *University of Chicago Law Review*, 2005, 72, 73-98.
- [23] **Png I.P.L.**, "Optimal Subsidies and Damages in the Presence of Judicial Error," *International Review of Law and Economics*, 1986, 6, 101-5.
- [24] **Polinsky, A. Mitchell and Shavell Steven**, "The Optimal Tradeoff between the Probability and Magnitude of Fines," *American Economic Review*, 1979, 69, 880-91.
- [25] **Polinsky, A. Mitchell and Shavell Steven**, "The Optimal Use of Fines and Imprisonment," *Journal of Public Economics*, 1984, 24, 89-99.
- [26] **Polinsky, A. Mitchell and Shavell Steven**, "Legal Error, Litigation and the Incentive to Obey the Law," *Journal of Law, Economics and Organization*, 1989, 5, 99-108.
- [27] **Polinsky, A. Mitchell and Shavell Steven**, "A Note on Optimal Fines When Wealth Varies Among Individuals," *American Economic Review*, 1991, 81, 618-21.
- [28] **Polinsky, A. Mitchell and Shavell Steven**, "Enforcement Costs and the Optimal Magnitude and Probability of Fines," *Journal of Law and Economics*, 1992, 35, 239-57.

- [29] **Schauer, Frederick**, *Playing By the Rules: A Philosophical Examination of Rule-based Decision Making in Law and Life*, Oxford, Oxford University Press, 1991.
- [30] **Shavell Steven**, "A Model of Optimal Incapacitation," *American Economic Review: Papers and Proceedings*, 1987a, 77, 107-10.
- [31] **Shavell Steven**, "The Optimal Use of Nonmonetary Sanctions as a Deterrent," *American Economic Review*, 1987b, 77, 584-92.
- [32] **Shavell Steven**, "Specific versus General Enforcement of Law," *Journal of Political Economy*, 1991, 99, 1088-108.
- [33] **Stiglitz, Joseph**, "Imperfect Information in the Product Market," in *Handbook of Industrial Organization*, Schmalensee, Richard and Willig, Robert, Eds., Volume 1, Elsevier Science Publishers, Amsterdam, 1989, 771-847.