

# Antitrust, Legal Standards and Investment

**Giovanni Immordino**

*Università di Salerno  
and CSEF*

**Michele Polo**

*Università Bocconi  
and IGIER*

February 2, 2013

**Abstract:** We study the interaction of a firm that invests in research and, if successful, undertakes a practice to exploit the innovation, and an enforcer that sets legal standards, fines and accuracy. In this setting deterrence on actions interacts with deterrence on research. When the practice increases expected welfare the enforcer commits not to intervene by choosing a more rigid per-se legality rule to boost investment, moving to a more flexible discriminating rule combined with type-I accuracy for higher probabilities of social harm. Patent and antitrust policies act as substitutes in our setting; additional room for per-se (illegality) rules emerges when fines are bounded. Our results on optimal legal standards extend from the case of (uncertain) investment in research to the case of (deterministic) investment in physical assets.

**Keywords:** legal standards, accuracy, antitrust, innovative activity, enforcement.

**JEL classification:** D73, K21, K42, L51.

**Acknowledgments:** Giovanni Immordino Università di Salerno and CSEF, 84084 Fisciano (SA), Italy, giimmo@tin.it. Michele Polo, Università Bocconi, Via Sarfatti 25, 20136 Milan, Italy, michele.polo@unibocconi.it. We are indebted to Emanuela Carbonara, Jacques Cremer, Nuno Garoupa, Louis Kaplow, Yannis Katsoulakos, Dilip Mookherjee, Massimo Motta, Marco Ottaviani, Marco Pagano, Patrick Rey, Matteo Rizzolli, John Roberts, Lars-Hendrik Röller, Maarten Schinkel, Giancarlo Spagnolo and David Ulph for helpful discussions and to seminar participants at ZEW 2011 (Mannheim), Cresse 2011 (Rhodes), 2nd Workshop 'Industrial Organization: Theory, Empirics and Experiments' (Otranto), 2011 CESifo Conference on Law and Economics (Munich), Toulouse (Competition Policy workshop), SIDE-ISLE 2011, Intertic Conference, Venice 2012 (Torino) and Universitat Pompeu Fabra (Barcelona). All usual disclaimers apply. This paper supersedes an old version circulated under the title: "Judicial errors and innovative activity."

# 1 Introduction

In recent years, many important antitrust cases on abuse of dominance and monopolization have involved technological market leaders or incumbents owning essential infrastructures. In their investigations, competition agencies have scrutinized a wide range of business strategies that the dominant firms allegedly used to maintain and increase their market power, from rebates to tying, from interoperability to margin squeeze. At the same time those investigations have not directly questioned the investment decisions of the incumbents. Although investments typically fall outside the scope of antitrust intervention, its impact on research or physical investment is one of the relevant issues to judge the desirability of public intervention. This paper studies optimal antitrust intervention, both in terms of legal standards and enforcement tools, in industries where the incumbent's investment plays a fundamental role.

Looking at competition policy in the last decade, many cases have involved dominant firms in high-tech industries, that reached the role of technological market leaders due to successful research investments and innovation. In the American and European cases Microsoft was alleged of foreclosure on a number of practices such as bundling of the operating system and the browser or media player applications, loyalty rebates granted to PC producers and limited access, a mild form of refusal to deal, through a reduction in interoperability of the servers' and clients' operating systems. The record fine to Intel in the case before the European Commission was motivated, among other conducts, by foreclosure through loyalty rebates. In the last years the focus of antitrust enforcement seems to be moving towards new technological leaders as Google and Apple. Although the enforcers investigated particular business practices, the debate in competition policy has then raised new issues on the impact of antitrust enforcement on the innovative activity that characterizes these industries. For instance the commitments imposed in the EC v. Microsoft decisions to disclose the API codes of the server operating system to competitors, have been commented not only in their ability to restore competition, but also in their indirect adverse effects on the incentives to innovate.

Other antitrust proceedings have focussed on the business practices of dominant firms in network industries, as electricity, gas or telecommunications, where the liberalization has led, through a combination of antitrust and regulatory interventions, to opening the access of networks to competitors. The primary goal of competition agencies, then, has been to prevent the incumbent from using business practices to restrict

the ability of newcomers to compete. Technical conditions of access, including interoperability, and price abuses in the form of predation, margin squeeze and aggressive rebates, have been at the core of cases involving the telecom incumbents in Europe, including Deutsche Telekom, British Telecom, France Telecom, Telecom Italia and Telefonica. Antitrust and regulatory policies have strongly supported the interests of competitors, burdening the incumbents with strict monitoring and commitments to open the access to the existing local loop. However, more recently, the need to create proper incentives to invest in new network infrastructures<sup>1</sup> has added to the debate, and several pro-incumbent measures, as the weakening of cost oriented access pricing, or the temporary lifting of regulatory or antitrust commitments, have been proposed. Hence, in many important cases we find the interaction between the short run control of business practices – the core activity of antitrust agencies – and the medium term impact on investment in research or physical assets. Although many commentators have argued that these issues should be recognized by antitrust enforcers, a detailed analysis of their implications on the desirable antitrust approach is still lacking, and the debate on competition policy has addressed other themes.

Indeed, the role played by economics in improving the analysis of anticompetitive conducts and its full recognition in the antitrust practice have been at the center of the discussion. In Europe, following the important reforms on cartel cases (article 101) and merger control, in 2009 the DG Competition of the European Commission has reshaped the enforcement of article 82 (now 102), pursuing an approach that rests on a deeper and more intelligent use of the new findings of economic analysis in the enforcement against unilateral practices.<sup>2</sup> A common view has emerged, labelled “effect-based” (or “more economic”) as opposed to the traditional form-based approach. The novelty of these proposals refers to identifying anticompetitive practices through a careful analysis of the foreclosure effects of the conducts, beyond their formal description.

The debate on monopolization practices, as unilateral conducts are defined in the US, has developed also on the other side of the Atlantic. Conflicting views emerged, with a report (Department of Justice, 2008) on enforcement policies under Section 2 of the Sherman Act that the Federal Trade Commission judged as a “blueprint for radically weakened enforcement” and the new Obama administration withdrew once

---

<sup>1</sup>The development of broadband services requires a new wave of investments in networks especially for the ultrabroad optical fiber infrastructures.

<sup>2</sup>See Gual et al. (2005) and DG Competition (2005) and (2008).

in office, announcing a more aggressive approach to the enforcement of monopolization issues.

This brief summary highlights some major aspects of the recent discussion. First, industrial organization offers different and diverging results on the impact on welfare of several business practices at the center of important antitrust cases. While some economists argue that dominant firms adopt socially harmful practices to maintain their market power, others consider this possibility skeptically, stressing instead the pursuit of superior efficiency as the driving force explaining the emergence of market leaders.<sup>3</sup> Hence, there is no general consensus today that certain business practices always produce desirable or negative welfare effects.

Secondly, the debate between different schools has extended from the economic arguments to be adopted in antitrust cases to the legal standards that the investigations should follow.<sup>4</sup> A wide range of proposals emerged, that can be roughly grouped into two sets: per-se rules that define legality or unlawfulness with reference to the conduct undertaken, and discriminating or effect-based rules that instead base the legal treatment of a certain practice on its anti-competitive or efficiency-enhancing effects.<sup>5</sup>

To sum up, several landmark cases have posed the issue of the indirect effects of antitrust enforcement on the incentives to invest. The rich debate on competition policy has further examined the different components of antitrust intervention, that require to choose appropriate legal standards as well as enforcement tools. We argue that time is ripe to put together these ingredients, analyzing how legal standards and enforcement policies should be shaped to take into account the impact of short run monitoring and control of business practices on long run investment.

This paper studies the optimal legal standards and the enforcement policy to reg-

---

<sup>3</sup>See Evans and Padilla (2005) for a brilliant summary of the evolution of economic thinking in antitrust from the traditional view to the Chicago critique to the post-Chicago approaches.

<sup>4</sup>For instance, Kovacic and Shapiro (2000), taking into account the modern contributions of the post-Chicago literature, observe that "some types of conducts (..) could deter entry and entrench dominance, but they also could generate efficiencies. The only way to tell in a given case appeared to be for the antitrust agencies and the courts to conduct a full-scale rule of reason inquiry".

<sup>5</sup>Discriminating rules, in turn, range from a case-by-case evaluation of the pro and anti-competitive effects, the so called rule of reason, to more structured rules that try to indirectly evaluate the effects by considering a set of factors that should affect the welfare impact of a certain practice.

ulate certain business practices of a dominant firm which invests in research or in physical capital. Initially we focus our analysis on an innovative environment, where a firm invests in research, moving then to physical assets. If research is successful, the firm gains market power, the kind of winner-takes-all competition that we often observe in high-tech industries. Then, the fresh incumbent becomes subject to antitrust scrutiny when undertaking commercial practices. Its expected profits, therefore, reflect the stricter or laxer enforcement by the competition agency on the practices adopted by the innovator. While the practice, applied to the new technology, is always privately profitable, its social effects may be positive or negative depending on the market conditions when the firm undertakes it, something that is inherently uncertain at the time the investment is sunk. Antitrust intervention operates within the boundaries set by the legal standard, that specify when a practice is unlawful. Once the investment is chosen, enforcement affects how the practice is adopted and the profits realized (*ex post* deterrence); however, enforcement also influences the initial decision to invest, that is driven by expected profits (*ex ante* deterrence). These two effects determine the choice of the optimal legal standards and enforcement policies. In this paper we consider per-se and discriminating rules, deriving the optimal enforcement policies under each regime and then identifying the optimal legal standard for given expectations of the enforcer on the effects of the practice.<sup>6</sup>

Our main results are the following. First, we show how the optimal legal standard and enforcement policy vary when the enforcer's presumptions on the effects of the practice become more and more pessimistic. Specifically, a more rigid per-se legality rule prevails on the more flexible discriminating legal standard for low probability of social harm: per-se legality acts as a commitment not to intervene ex-post, in the (unlikely) event the practice is socially harmful, boosting this way the innovative investment. When the harmful effect becomes more likely, the enforcer moves to the discriminating rule and improves type-I accuracy to sustain investment. Second, we show that while *ex ante* it is optimal, when negative social effects are unlikely, to commit to a more rigid per-se legality rule, once the investment is sunk a flexible discriminating rule would be preferred. Hence, there is a time-consistency issue that may require to use commitment tools (regulations, guidelines, precedents). Third, we extend the model to include a positive effect of the new technology on profits and welfare

---

<sup>6</sup>The welfare effect of the practice depends on the magnitude of its social benefits and harms and the likelihood of these effects, what we can call the "economic model" of the enforcer, or, in the words of Judge Frank Easterbrook, her presumptions (see Easterbrook, 1984). These presumptions express the view of the enforcer on the expected effects of a certain business practice.

even when the practice is not adopted, adding an additional motive to invest. In this framework, the baseline profits may be thought as guaranteed by patent protection, while the additional profits that can be obtained through the practice are affected by the antitrust policy. This way, we can consider in a simple setting the interaction of patent and competition policies. We show that, when the degree of patent protection is reduced, the region where antitrust policy opts for per-se legality, an extreme form of innovation-friendly antitrust intervention, becomes larger. In other words, patent and antitrust policies act as substitutes in our setting. Fourth, some additional room for per-se rules emerges, as a cost saving solution to enforcement, when fines are capped at some upper bound: per-se legality is adopted for low probability of social damages, then replaced by a discriminating rule with more and more symmetric accuracy, with per-se illegality as the optimal legal standard when the new technology is very likely to be socially harmful. Finally, we show that, with minor differences, our results on the optimal legal standard extend from the case of (uncertain) investment in research to the case of (deterministic) investment in physical assets, establishing a more general result on antitrust legal standards when investment matters.

We contribute to the literature on antitrust and regulatory intervention in industries. Immordino, Pagano and Polo (2011, hereafter IPP) propose an analytical framework similar to this paper. They focus on the choice between *ex post* law enforcement and *ex ante* authorization, identifying when each policy is optimal. In this paper, instead, we go more in depth into the selection of optimal legal standards, comparing per-se and discriminating rules, within the ex-post law enforcement regimes. Hence, the results of the two papers can be read as complementary.

Another model that comes close to ours is that of Schwartzstein and Shleifer (2012, hereafter SS). They consider a setting where safe and unsafe firms decide whether to produce and may take precautions. Firms face uncertainty as to the liability for damages that will apply to them, due to possible judicial errors: a judge may mistake a safe firm for an unsafe one, which creates a disincentive effect for safe firms. Similarly to us both IPP and SS find that regulation should be softer when social harm is unlikely (our first result). But our analysis differs in three main directions. First, we focus on antitrust policies and its interaction with patent policy. Second, we enlarge the enforcer set of instruments to include the optimal choice of accuracy. Third, differently from SS, in our setting uncertainty comes from the unpredictability of market conditions at the time the investment is sunk, and not only from judicial errors. As such, we argue it captures a more general phenomenon than frictions in

enforcement.

The impact of antitrust enforcement in innovative industries is analyzed also in a paper by Segal and Whinston (2007). Considering a sequence of innovations, the authors analyze the trade-off between protecting the incumbents, increasing this way the rents of the winner and the incentives to invest in innovation, and protecting the innovative entrants, that increases the rate of technical progress. They derive conditions under which the latter effect is the dominant one. While the previous paper offers interesting results on law enforcement when innovative activity is a crucial component, it does not consider the choice among different legal standards that represents the focus of this paper.

In Katsoulakos and Ulph (2009) a welfare analysis of legal standard is developed, which compares per-se rules and discriminating (effect based) rules. The authors identify some key elements that can help deciding the more appropriate legal standard and the cases in which type-I or type-II accuracy are more desirable. However, the impact of enforcement on investment, that is key in our paper, is not addressed.

Moreover, our results, although motivated with reference to competition policy and framed in terms of antitrust intervention, give useful insights in the more general debate on legal standards and accuracy in the law and economics literature<sup>7</sup>.

The remainder of this paper is organized as follows. Section 2 presents the model. Section 3 focus on antitrust intervention in innovative industries. In Section 4 we study the interaction of patent policy and antitrust policy, the effect on legal standards of a cap on fines and the case of deterministic investment in physical capital. All proofs not following immediately from the main text are relegated to the Appendix.

## 2 The model

In this section we describe in detail how we model the interaction of antitrust intervention and research investment. A firm sinks resources in research, discovering with a certain probability a new technology and developing, in a second stage, business

---

<sup>7</sup>Judicial errors and their reduction, i.e. accuracy, are a central concern in law enforcement: they have been analyzed in the standard model of law enforcement proposed by Kaplow (1994), Kaplow and Shavell (1994, 1996), Polinsky and Shavell (2000) and Png (1986) among others, which focuses on the (negative) impact of such errors on marginal deterrence. On legal standards see Evans and Padilla (2005).

strategies to obtain profits from the innovation.<sup>8</sup> The larger the initial investment, the larger, *ceteris paribus*, the expected profits, since the probability of discovering the new technology increases in the investment itself. At the same time, the *ex post* profits depend on how the antitrust policy deals with the business practices that the firm applies to extract profits from the investment. The laxer (stricter) is competition policy, the more (less) profitable opportunities are opened if research is successful, boosting (reducing) the *ex ante* incentives to invest.

We first analyze the investment and the practice undertaken by the firm; then we introduce the legal standard adopted by an antitrust authority to evaluate the practice and the enforcement tools used to influence the firm's choices.

*Investment and practices.* We consider an industry that is initially competitive and characterized by fragmentation and symmetry among firms, none of which has market power. By investing in research a firm can discover a new technology that generates a strong competitive advantage and creates market power, the winner-takes-all dynamics that we observe in many high-tech industries. For instance, the firm invests to design a new operating system and applications for pc's that significantly improve over the existing packages. The innovating firm, if research is successful, becomes dominant and subject to antitrust scrutiny. The **investment**  $I$  determines the chances of success in the research process<sup>9</sup>: for simplicity, the firm's probability of innovating  $p(I)$  is assumed to be linear in  $I$ , i.e.  $p(I) = I$  and  $I \in [0, 1]$ . The cost of learning is increasing and convex in the firm's investment and is assumed to be  $c(I) = \frac{I^2}{2}$ .

The firm can exploit the new technology by adopting particular **business practices** that allow to extract profits from the investment. A practice can be undertaken at a different intensity by choosing an action  $a$ , making the design of business strategies a matter of degree rather than a yes/no decision. For instance, the firm, rather than simply offering an innovative operating system bundled (or unbundled) with the applications, can implement different levels of interoperability of its new operating system with its, and the competitors', applications, controlling this way the value of the joint use of these packages. The set of actions is  $A = [0, 1]$ , where the lower bound  $a = 0$  can be interpreted as not undertaking the practice at all.

---

<sup>8</sup>In Section 4.3 we study the case of a deterministic investment in physical assets.

<sup>9</sup>We do not model competition in research and patent races, but rather adopt the approach first proposed by Arrow (1962) to study the incentives to invest in research. We further discuss this issue in section 4.1.



*Private and social effects.* When the dominant firm undertakes the practice, this latter affects profits and welfare according to the intensity measured by the action undertaken. More precisely, the practice and associated actions yield **profits**  $\Pi(a) = \pi a \geq 0$  which are normalized to zero when the practice is not adopted ( $a = 0$ ) and correspond to the returns from “business as usual”.

While the private effects of the practice are always positive, its **social impact** may be positive or negative. Indeed, the way a practice affects social welfare once the new technology is introduced depends on the occurrence *ex post* of a set of circumstances (market structure, conditions of entry, products offered by the competitors, state of demand, etc.). This set of factual elements makes the equilibrium of the market game welfare enhancing or detrimental compared with the initial situation.

More precisely, the effects of the practice are described by two states of the world. In the bad state  $s = b$ , when the firm exploits the new technology through the business practice, social welfare is reduced compared to the benchmark level according to the expression  $W^b(a) = -w^b a \leq 0$  with  $w^b > 0$ . In the bad state, private incentives conflict with social welfare, that is, when the firm increases the intensity of the practice, social welfare falls. For instance, limiting interoperability of competitors’ applications with the innovative operating system marketed by the firm restricts the rivals’ ability to compete, with a stronger effect the less compatible are the products.

In the good state  $s = g$ , instead, social welfare increases when the firm undertakes the practice:  $W^g(a) = w^g a \geq 0$  with  $w^g > 0$ . In this case, there is no conflict between private and social incentives since the practice increases both the profits of the firm and social welfare. Examples are when alternative operating systems are marketed, offering additional opportunities to the competitors’ applications, while the integrated package released by the firm allows a more user-friendly usage of the software.<sup>10</sup>

*Information.* We assume that neither the firm nor the enforcer know the social effects of the practice at the time the policy is set and the investment is sunk, and they both assign a probability  $\beta$  to the realization of the bad state. Later, if the research activity has successfully led to a new technology, the firm, that has a better

---

<sup>10</sup>In the benchmark model we assume that the new technology produces private and social effects only if combined with the practice, while profits and welfare do not change with respect to the competitive scenario if the practice is not adopted ( $a = 0$ ). We choose this modeling strategy to focus on the impact of antitrust intervention (that affects the adoption of the practice), on the incentives to invest in research. In Section 4.1 we extend the model by considering a positive fixed effect of the new technology on profits and welfare, that adds to the effect of the practice described in the benchmark model.

knowledge of market conditions, perfectly observes the effects of the practice (state of the world  $s = b$  or  $g$ ), while the enforcer imperfectly assesses them.

Following this approach, we assume that the enforcer perfectly recognizes the action chosen by the firm, i.e. any  $a \in A$ . Yet, the information regarding the effects of the practice (state of the world) is less accurate and the enforcer can commit **errors**. Specifically, the enforcer receives a noisy signal  $\sigma$  on the state of the world, that is whether the incumbent's practice is welfare enhancing or decreasing. The enforcer interprets the signal as follows: if  $\sigma > x$  then she concludes that  $s = b$ , where the threshold  $x$  in the legal literature is called the burden of proof.<sup>11</sup> With probability  $\varepsilon^I$  the signal is incorrect in the good state: when the new action indeed is socially beneficial the enforcer considers it as socially harmful, a type-I error. Conversely, with probability  $\varepsilon^{II}$  the signal is incorrect when the true state is the bad one: in this case the enforcer fails to identify  $A$  as socially damaging, committing a type-II error. Hence,

$$\varepsilon^I = \Pr(\sigma > x | s = g) \quad \text{and} \quad \varepsilon^{II} = \Pr(\sigma \leq x | s = b).$$

We assume that the signals are informative, i.e.  $\varepsilon^i \leq \bar{\varepsilon} < \frac{1}{2}$ ,  $i = I, II$ .

The economic model implicitly adopted by the enforcer when considering a certain practice and its implementation through the actions, what we can consider as her presumptions, is summarized in the terms  $\{w^g, w^b, \pi, \beta\}$ . In the remaining part of the paper we show that the optimal legal standards and enforcement policies for a certain practice depend, given the feasible policy instruments, on these parameters of the enforcer's economic model.

We impose the following restrictions on the parameters:

$$w^g < 1, \tag{1}$$

that ensures an internal solution for the investment in all regimes, and

$$w^g - w^b - \pi > 0, \tag{2}$$

which implies that the welfare effect of the practice in the good state is sufficiently large.

---

<sup>11</sup>On the burden of proof see, for instance, Kaplow (2011a, 2011b, 2012) and Demougin and Fluet (2008). In this paper we maintain, within each legal standard, the burden of proof fixed while allowing the enforcer to improve the accuracy.

*Public policies: legal standards, fines and accuracy.* The focus of antitrust intervention are the practices undertaken by (dominant) firms. Specifically, the enforcer designs the public policies to contain the potential hazards posed by certain practices and collects information according to the legal standards in place, to properly implement the enforcement policy. Each legal standard specifically defines under which circumstances (if any) the practice is considered unlawful, and requires to specify a minimum amount of evidence to convict the firm. A richer definition of unlawfulness in general requires a more complex set of information, which is more costly to collect and may lead more frequently to errors.

The enforcer can choose among different **legal standards**: we consider *per-se rules* based on the action undertaken and *discriminating rules* that depend on the effect of this action. Per-se rules can be further distinguished in:

*L per-se legality*: any action  $a \in A$  is always legal no matter which signal the enforcer receives;

*IL per-se illegality*: any action  $a \in A$  is always illegal no matter which signal the enforcer receives.

It should be stressed that per-se legality and per-se illegality differ in the power of the enforcer to fine the firm when the practice is undertaken, and not in the fact that the practice is adopted or not in equilibrium. Indeed, we shall see that even under per-se illegality it may be optimal to have the firm undertake the practice at some degree (and pay a positive fine).

Alternatively, the enforcer can adopt a discriminating legal standard (or effect-based rule) that links the unlawfulness of a practice to its social consequences:

*D discriminating*: any action  $a \in A$  is legal unless the enforcer receives a signal  $\sigma > x$ .

Since in our setting errors occur only in the assessment of the social effects and not when recognizing the action undertaken, they are an issue only under a discriminating rule, while per-se rules do not lead to errors. This is a simple way to introduce the distinction between per-se rules, based on a narrower set of elements but less prone to errors, and discriminating rules, that use a wider set of information but are potentially less accurate.

Given the legal standard the enforcer designs her policies through a set of **enforcement policy tools**, that is controlling the level of errors, and setting the fine schedule. The enforcer can reduce the level of type- $i$  error by committing resources to refine the assessment of the effects, what is usually called **accuracy**. In other words, the enforcer can collect additional evidence, reducing this way the variance of the conditional distribution of the signal and estimating more precisely whether the practice increases or reduces welfare. We assume that the cost of reducing a type- $i$  error is increasing and convex, and that if no resources are devoted to this goal the error committed is equal to  $\bar{\varepsilon}$ .<sup>12</sup> More precisely, the cost of implementing an error probability  $\varepsilon^i$  is  $g(\varepsilon^i) = \frac{\gamma}{2}(\bar{\varepsilon} - \varepsilon^i)^2$ .

Besides the level of type-I and type-II errors, the enforcer controls a third policy variable: a non decreasing **fine schedule**  $f(a) \in [f, F]$ . The fine may be levied on the practice, since the antitrust law applies to business conducts, while it cannot be related to the investment activity, that typically is outside the scope of competition policy.<sup>13</sup> Since the profit function  $\pi a$  is increasing and linear in  $a$ , we can use with no loss of generality, within the set of non-decreasing fine schedules, the stepwise function

$$f(a) = \begin{cases} 0 & \text{if } a = 0 \\ \underline{f} \geq f & \text{if } 0 < a \leq \tilde{a} \\ \bar{f} \leq F & \text{if } a > \tilde{a}. \end{cases} \quad (3)$$

Notice that, under any rule, the enforcer cannot fine a firm when it does not undertake the practice ( $a = 0$ ). In the benchmark model the feasible set of fines includes full amnesty ( $f = 0$ ) and an upper bound sufficiently high not to bind the enforcer on the desired fine. We discuss the case when the minimum fine is positive ( $f > 0$ ) after Lemma 2, and the case when the maximum fine  $F$  is capped in Section 4.2.

*Timing.* The timing of the model is as follows. At time 0 nature chooses the state of the world  $s = \{g, b\}$ . At time 1, the enforcer commits to a certain legal standard  $i \in \{L, IL, D\}$  and sets the fine schedule  $f(a)$  and the level of the errors  $\varepsilon^I$  and  $\varepsilon^{II}$  (accuracies). At time 2, having observed the legal standard and the enforcement policy set by the enforcer, the firm chooses the research investment  $I$ , innovates with probability  $p(I) = I$  and in this case also learns the state of the world  $s = b, g$ . At

---

<sup>12</sup>In this case the decision is based on a small set of evidence easy and inexpensive to collect.

<sup>13</sup>Beyond this institutional argument, moreover, we could argue that private investment effort is hardly observable and/or verifiable by third parties, that therefore cannot condition the fines to  $I$ .

time 3, the firm chooses an action, conditional on what it learnt in the previous stage. Finally, at time 4 the action undertaken determines the private profits and the social welfare; the enforcer receives a signal  $\sigma$  that is incorrect with probability  $\varepsilon^I$  in the good state and  $\varepsilon^{II}$  in the bad state and levies a fine (if any) consistently with the legal standard and enforcement policy adopted.

### 3 Optimal legal standards and enforcement policies

To evaluate the benefits of public intervention we start by identifying the first-best outcome ( $FB$ ), which would be obtained if the enforcer could directly control the firm's action and investment.

Let us denote by  $a^s$  the action chosen in state  $s = b, g$ . The welfare maximizing actions are clearly  $a^b = 0$  (do not undertake the practice when socially harmful) and  $a^g = 1$  (undertake the practice at the highest degree when welfare enhancing). The associated expected welfare is therefore  $EW_{FB}(\beta, I) = I(1 - \beta)w^g - \frac{I^2}{2}$ , that yields the optimal investment level

$$I_{FB} = (1 - \beta)w^g. \quad (4)$$

The first best investment  $I_{FB}$  is increasing in the likelihood of the good state  $(1 - \beta)$  and in the welfare gain  $w^g$ . Since under a first best policy the practice is undertaken only when it is welfare improving, the investment always has a positive expected impact on welfare, and it is therefore always positive and increasing in the probability of social gains.<sup>14</sup>

In what follows, the policy maker is assumed not to control firm's choices directly, but to influence them via penalties. More precisely, the enforcer observes the actions  $a$ , and can condition the penalties to them, whenever they can be levied according to the legal standard in place, but cannot base the fine on the level of investment. We start with per-se rules, identifying the optimal enforcement in this setting, and then move to discriminating rules and the associated optimal enforcement policy. Finally, we compare the two legal standards, evaluated at the corresponding optimal enforcement policy, and select, for different values of the prior on social harm  $\beta$ , the overall optimal solution.

---

<sup>14</sup>The expected welfare, evaluated at the first-best policies, is  $EW_{FB}(\beta) = \frac{[(1-\beta)w^g]^2}{2}$ .

### 3.1 Per-se rules

The very nature of per-se rules is to treat the practice at any degree  $a \in A$  as legal ( $L$ -rule) or unlawful ( $IL$ -rule) irrespective of the effects (signal  $\sigma$  received). We analyze the optimal enforcement starting from stage 3, when the firm chooses the action, that is the level of intensity of the practice. Since the practice is equally profitable in both states of the world and per-se rules treat the practice irrespective of its effects, the incumbent undertakes the same profit maximizing action in both states, no matter if it is welfare enhancing or socially harmful. The specific action undertaken, however, depends on the fine schedule  $f(a)$  designed by the enforcer. If the research investment has been successful, the profits at time 3 when the action is selected are  $\pi a - f(a)$  and the firm chooses  $\tilde{a} = \arg \max_a \pi a - f(a)$ . Given the fine schedule (3), the incentive compatibility constraint can be written as  $\pi \tilde{a} - \underline{f} \geq \pi - \bar{f}$ . The undertake constraint, instead, ensures that the firm (weakly) prefers to adopt the practice ( $a > 0$ ) rather than keeping on with "business as usual" ( $a = 0$ ), and it is relevant as long as  $\tilde{a} > 0$ :  $\pi \tilde{a} - \underline{f} \geq 0$ .

Then, the expected profits at time 2 under per-se rules (subscript  $PS$ ) are  $E\Pi_{PS} = I(\pi \tilde{a} - \underline{f}) - \frac{I^2}{2}$  and the profit maximizing investment is

$$I_{PS} = \pi \tilde{a} - \underline{f}. \quad (5)$$

Hence, although the fine is not conditioned on the level of investment, it affects the firm's research effort  $I$ . Indeed, the firm realizes that it will pay  $\underline{f}$  only if research is successful. Then, a higher fine  $\underline{f}$  increases the expected fine, reducing the marginal benefit from research and the associated investment.

We can write the expected welfare under per-se rules as:

$$EW_{PS}(\beta) = I_{PS} [(1 - \beta)w^g - \beta w^b] \tilde{a} - \frac{(I_{PS})^2}{2} = I_{PS} Ew(\beta) \tilde{a} - \frac{(I_{PS})^2}{2}, \quad (6)$$

where  $Ew(\beta) = (1 - \beta)w^g - \beta w^b$  is the expected marginal welfare of an increase in the intensity of the practice. The design of the optimal fine schedule is equivalent to (indirectly) implementing, among the profit-maximizing actions  $\tilde{a}$ , the one that maximizes welfare – which we denote  $\hat{a}$  – that is the action that the firm is willing to choose according to the incentive compatibility and undertake constraints, and that is socially optimal. The enforcer therefore maximizes the expected welfare setting  $\tilde{a}$ ,  $\underline{f}$  and  $\bar{f}$ , subject to the incentive compatibility and undertake constraints, given the investment  $I_{PS}$ .

Notice that although antitrust policy intervenes only on the practice (actions), deterrence works through two different channels: *ex post* deterrence on actions, once the investment is sunk and has been successful (marginal deterrence);<sup>15</sup> and *ex ante* deterrence on investment. This latter effect works through the impact of the action implemented on *ex post* profits and through the direct effect of the fine on the investment itself.

In the following lemma we derive the optimal enforcement policy under per-se rules. It's worth noting that by studying the optimal fines we can implicitly identify whether per-se legality or per-se illegality is the desirable legal standard. Indeed, if the optimal enforcement policy prescribes to set  $\hat{a} = 1$  and  $\underline{f} = 0$ , it is optimal not to fine the practice at any degree  $a \in A$ . Then, the corresponding legal standard is per-se legality. If, instead,  $\hat{a} < 1$  and  $\underline{f} > 0$ , the practice is always sanctioned, possibly with different levels of the fine, and, therefore, the enforcer is applying a per-se illegality rule.

Before describing the optimal legal standards and enforcement policies under per-se rules, it is convenient to introduce the following thresholds

$$\beta_1 \equiv \frac{w^g - \pi}{w^g + w^b} < \beta_2 \equiv \frac{w^g}{w^g + w^b}.$$

**Lemma 1 (Optimal enforcement policy under per-se rules)** *Assume the minimum fine is zero and the maximum fine sufficiently high, i.e.  $f = 0$  and  $F > \pi$ . The optimal legal standard and enforcement policy under per-se rules are:*

1. for  $\beta \in [0, \beta_1]$ , the optimal legal standard is per-se legality and the optimal enforcement implements  $a^g = a^b = 1$  and  $I_{PS} = \pi$ , by setting  $\hat{a} = 1$ ,  $\underline{f} = 0$ . The expected welfare is  $EW_{PS}(\beta) = \frac{\pi}{2} [2Ew(\beta) - \pi]$  and is decreasing and linear in  $\beta$  with  $EW_{PS}(\beta_1) = \frac{[Ew(\beta_1)]^2}{2}$
2. for  $\beta \in (\beta_1, \beta_2)$ , the optimal legal standard turns to per-se illegality and the optimal enforcement implements  $a^g = a^b = 1$  and  $I_{PS} = Ew(\beta)$ , decreasing in  $\beta$ , by setting  $\hat{a} = 1$  and  $\underline{f} = [\pi - Ew(\beta)]$ . The expected welfare is  $EW_{PS}(\beta) = \frac{[Ew(\beta)]^2}{2}$  and is decreasing and concave in  $\beta$ , with  $EW_{PS}(\beta_2) = 0$ .

---

<sup>15</sup>For the standard marginal deterrence problem in law enforcement see for instance Mookherjee and Png (1994).

3. for  $\beta \in [\beta_2, 1]$ , the optimal legal standard is still per-se illegality and the optimal enforcement implements  $a^g = a^b = 0$  and  $I = 0$ , by setting  $\hat{a}^R = 0$  and any  $\bar{f} \geq \pi$ . The expected welfare is  $EW_{PS}^R(\beta) = 0$ .

Lemma 1 shows that the optimal legal standard and enforcement policy vary with the likelihood of social harm. The enforcement policy allows to implement the action  $\hat{a}$  by properly setting the fines. The optimal policy discourages the action when it is welfare detrimental and implements the practice (at the highest degree  $a = 1$ ) otherwise. In this latter case, turning to the optimal legal standards, per-se legality is adopted ( $\beta < \beta_1$ ), while it is replaced by per-se illegality when the practice is socially harmful ( $\beta > \beta_2$ ). Since the investment is influenced by the fine  $\underline{f}$ , a third outcome arises for intermediate values of the parameter  $\beta$ . When  $\beta \in (\beta_1, \beta_2)$  the enforcer adopts a per-se illegality regime, but focuses enforcement on progressively reducing the investment by raising the fine  $\underline{f}$ , rather than discouraging the practice. In other words, in this region the enforcer intervenes through *ex ante* rather than *ex post* deterrence. Finally, it is worth noting that the expected welfare is continuous and decreasing in  $\beta$ .

### 3.2 Discriminating rules

A discriminating rule is based both on the observed actions and on the signal. An action  $a \in A$  is illegal if the enforcer receives a signal  $\sigma > x$ . Although the signal may be incorrect, we have assumed it to be informative. The enforcer, then, can indeed implement – in contrast with per-se rules – different actions in different states of the world. Since the discriminating legal standard does not allow the enforcer to levy any fine if the signal is  $\sigma \leq x$ , the fine schedule  $f(a)$  applies only when the signal of the bad state is received. Due to judicial errors, this occurs with probability  $1 - \varepsilon^{II}$  when indeed the practice is socially harmful, and with probability  $\varepsilon^I$  when instead it is welfare enhancing.

When the practice is socially harmful, given the fine schedule  $f(a)$ , the incentive compatibility and undertake constraints give the following inequalities:  $\pi \tilde{a}^b - (1 - \varepsilon^{II})\underline{f} \geq 0 \geq \pi - (1 - \varepsilon^{II})\bar{f}$ . Although the incentive compatibility constraint to implement  $\tilde{a}^b$  puts only a lower bound on the maximum fine  $\bar{f}$ , when we turn to the good state, type-I errors are committed, and an excessively high  $\bar{f}$  might induce the



firm to undertake  $a^g = \tilde{a}^b$  rather than  $a^g = 1$ .<sup>16</sup> Hence, we have to further impose the following incentive compatibility and undertake constraints for the good state. Taken together, they give the following inequalities:  $\pi\tilde{a}^b - \varepsilon^I \underline{f} \leq 0 \leq \pi - \varepsilon^I \bar{f}$ . These constraints define the interval in which the fines must be chosen in order to implement  $a^b = \tilde{a}^b$  and  $a^g = 1$ , i.e.,

$$\bar{f} \in \left[ \underline{f} + \frac{\pi(1 - \tilde{a}^b)}{1 - \varepsilon^{II}}, \underline{f} + \frac{\pi(1 - \tilde{a}^b)}{\varepsilon^I} \right]. \quad (7)$$

At stage 2, the firm decides the level of investment that maximizes the expected profits under discriminating rules (subscript  $D$ )

$$E\Pi_D = I \left\{ (1 - \beta) [\pi - \varepsilon^I \bar{f}] + \beta [\pi\tilde{a}^b - (1 - \varepsilon^{II})\underline{f}] \right\} - \frac{I^2}{2}.$$

The innovative investment in the discriminating regime is therefore

$$I_D = (1 - \beta) [\pi - \varepsilon^I \bar{f}] + \beta [\pi\tilde{a}^b - (1 - \varepsilon^{II})\underline{f}] \geq 0. \quad (8)$$

Notice that errors play an opposite role on the investment: when type-I errors increase, over-deterrence reduces the investment while a higher probability of type-II errors, inducing under-deterrence, boosts the research effort.

The expected welfare under the discriminating rule is

$$EW_D = I_D \left[ \Delta W_D - \frac{I_D}{2} \right] - \frac{\gamma}{2} (\bar{\varepsilon} - \varepsilon^I)^2 - \frac{\gamma}{2} (\bar{\varepsilon} - \varepsilon^{II})^2, \quad (9)$$

where  $\Delta W_D = (1 - \beta)w^g - \beta w^b \tilde{a}^b$ . The optimal policy requires therefore to set the fine schedule  $(\underline{f}, \bar{f}, \tilde{a}^b)$  and the errors  $\varepsilon^I$  and  $\varepsilon^{II}$  to maximize the expected welfare under the above constraints. As before, we denote as  $\hat{a}^b$  the action that solves this program (in the bad state). Finally, let us define the following threshold

$$\beta_0 \equiv \frac{w^g - w^b - \pi}{w^g + w^b}.$$

In the following lemma we identify the optimal enforcement policy.

**Lemma 2 (Optimal enforcement policy under discriminating rules)** *Assume the minimum fine is zero and the maximum fine sufficiently high, i.e.  $f = 0$  and  $F > \pi$ . The optimal legal standard and enforcement policy under the discriminating regime are:*

---

<sup>16</sup>This is what Kaplow (2011a) defines as the chilling effect of fines on desirable actions.

1. for  $\beta \in [0, \beta_0]$ , the optimal policy implements  $a^g = a^b = 1$  and  $I_D = \pi$  by setting  $\hat{a}^b = 1$ ,  $\underline{f} = 0$  and the minimum level of accuracy ( $\varepsilon^I = \varepsilon^{II} = \bar{\varepsilon}$ ). The optimal policy makes the discriminating regime equivalent to a per-se legality rule. The expected welfare is  $EW_D(\beta) = \pi [Ew(\beta) - \frac{\pi}{2}]$  and is decreasing and linear in  $\beta$ .
2. for  $\beta \in (\beta_0, 1]$  if  $\gamma$  is sufficiently large the optimal policy implements the actions  $a^b = \hat{a}^b < 1$ ,  $a^g = 1$  and investment  $I_D < \pi$  by improving type-I accuracy ( $\varepsilon^I < \bar{\varepsilon}$ ,  $\varepsilon^{II} = \bar{\varepsilon}$ ) and by setting  $\hat{a}^b < 1$ ,  $\underline{f} = 0$ , and  $\bar{f} = \frac{\pi(1-\hat{a}^b)}{(1-\bar{\varepsilon})}$ . Moreover,  $\hat{a}^b$  is decreasing in  $\beta$  with  $\hat{a}^b \rightarrow 1$  for  $\beta \rightarrow \beta_0$  and  $\hat{a}^b \rightarrow 0$  for  $\beta \rightarrow 1$ . Finally, the expected welfare  $EW_D(\beta)$  is decreasing in  $\beta$  and tends to 0 when  $\beta \rightarrow 1$ .

The optimal enforcement policy under the discriminating rule is shaped by the interaction of *ex post* (marginal) deterrence, focussed on the control of the action, and *ex ante* deterrence related to investment. Compared to per-se regimes, the discriminating rule allows implementing different actions in the two states, the welfare maximizing action  $a^g = 1$  in the good state and an action  $\hat{a}^b \in (0, 1)$  in the bad state. While *ex post* deterrence always requires a lower  $\hat{a}^b$ , *ex ante* deterrence prescribes a high  $\hat{a}^b$  to increase expected profits and investment when the expected welfare increases with the practice.<sup>17</sup>

When social harm is unlikely ( $\beta < \beta_0$ ), *ex ante* deterrence prevails and calls for a lax enforcement, implementing  $\hat{a}^b = 1$ , an outcome equivalent to a per-se legality rule.<sup>18</sup> Above this threshold, the enforcer implements  $\hat{a}^b < 1$  by properly setting the fine schedule and errors according to the incentive compatibility constraints. By lowering  $\hat{a}^b$ , the enforcer reduces the negative impact of the practice on welfare, counterbalancing the higher probability of social harm, and at the same time progressively lowers the investment. The optimal policy also commands a reduction in type-I errors, that make the firm sanctioned in the good state, softening over-deterrence and boosting the innovative investment. This goal cannot be pursued only through a reduction in the fine  $\bar{f}$  since the incentive compatibility constraint requires a sufficiently high

---

<sup>17</sup>Taking the first partial derivative of the expected welfare with respect to the implemented action  $\hat{a}^b$ ,  $\frac{\partial EW_D}{\partial \hat{a}^b} = [\Delta W_D - I_D] \beta \pi - \beta w^b I_D$ , the *ex ante* deterrence effect corresponds to the first term, and it is positive as long as  $\Delta W_D - I_D > 0$ , while *ex post* deterrence refers to the second (negative) term.

<sup>18</sup>Notice that this occurs in an interval  $[0, \beta_0]$  in which the per-se rule also opted for generalized acquittal, since  $\beta_0 < \beta_1$ .

fine to induce the firm to choose  $\hat{a}^b < 1$  instead of 1 in the bad state. Then,  $\varepsilon^I$ , that acts as a substitute to the fine in affecting the investment, is reduced.

Lemma 2 shows that for a low probability of social harm ( $\beta < \beta_0$ ) the discriminating rule replicates a per-se legality regime. This result is due to our assumption that the range of feasible fines includes full amnesty ( $f = 0$ ). If, instead, the minimum fine that can be levied in case of a bad signal is positive ( $f > 0$ ), for low  $\beta$  the enforcer would still implement the action at the highest level,  $\hat{a}^b = 1$  and apply the lowest admissible fine, i.e.  $\underline{f} = f > 0$ . However, in this case the outcome under a discriminating rule would no longer encompass the per-se legality regime, since the investment and the expected welfare would be lower under the discriminating rule compared with the per-se legality regime.

### 3.3 Optimal legal standards

We are now equipped to find the optimal regime, by comparing the expected welfare, evaluated at the corresponding optimal enforcement policies, under the per-se and discriminating rules. The following Proposition, proved in the Appendix, establishes the result.

**Proposition 1 (Optimal legal standards)** *The optimal legal standard is a per-se legality rule for  $\beta \leq \beta_0$  and a discriminating rule for higher  $\beta$ .*

The choice of the legal standard depends on the ability of the different regimes to ensure both *ex post* deterrence, implementing the practice at the welfare maximizing level, and *ex ante* deterrence, inducing the desired level of investment in research. When  $\beta$  is low, *ex ante* and *ex post* deterrence may require opposite policies and legal standards. Indeed, *ex post* deterrence requires to discourage the practice whenever it is socially harmful; then, a discriminating rule is more flexible and effective under this concern, allowing to be lenient when the practice is welfare enhancing and severe when welfare detrimental. Hence, concerning *ex post* deterrence, a discriminating rule is superior. *Ex ante* deterrence, instead, requires to discourage the investment only if it is expected to reduce welfare, and to boost it otherwise. In this latter case, that occurs when the social harm is unlikely, a discriminating rule may become less appealing. Under a discriminating regime, indeed, the enforcer cannot be lenient when a negative signal is received, and a fine must be levied reducing the investment. In this case, a

rigid rule (per-se legality) may dominate a flexible one (discriminating) for its ability to commit not to intervene *ex post* on the practice when socially harmful, boosting the research investment at most.<sup>19</sup> In other words, when the probability of social harm is sufficiently low, the enforcer sustains the desirable research investment by opting for a more rigid per-se legality rule, as a way to commit not to fine the firm. When, instead, social harm is more likely, that is for  $\beta > \beta_0$ , the more flexible discriminating rule dominates, allowing to better combine *ex ante* and *ex post* deterrence.<sup>20</sup>

### 3.4 Sunk investment

Since the impact of legal standards and enforcement policies on the investment played a key role in our previous analysis, it is interesting to discuss a different environment where the enforcer selects the legal standard, the fines and the level of accuracy once the investment has been sunk by the firm. This case may shed some light on two different issues: first, whether the initial commitment to a certain policy, assumed in the benchmark model, matters, compared to a case where the enforcer does not bind her hands before the investment is decided; secondly, which is the optimal legal standard in industries where new research investments are not a major element of the picture.

In the alternative environment we are discussing, the level of investment is given at the time legal standard and policy tools are chosen. Hence, the enforcer designs them considering only their impact on the action  $a$ . In other words, if the investment is sunk before the policy is chosen, this latter is designed to maximize welfare for a given investment. Drawing from our previous discussion, it is evident that in this alternative environment *ex ante* deterrence does not bite, and the policy is entirely driven by the *ex post* concern for the action chosen, that is the marginal deterrence issue.

Per-se rules, in this case, appear to be inferior, as they treat an action in the same way no matter if it increases or reduces welfare. Conversely, a discriminating rule, by appropriately setting the fines and the threshold  $\hat{a}^b$ , can implement the first

---

<sup>19</sup>This difference between per-se legality and a discriminating rule is particularly evident in the case, discussed above, when the legal norm does not include full amnesty in the range of feasible fines, that is  $f > 0$ . In this case, the discriminating rule charges  $f$  when implementing the action  $\hat{a}^b$  and does not succeed to replicate the per-se legality regime.

<sup>20</sup>The role of commitment and flexibility of a legal system in affecting growth has been recently studied by Anderlini et.al. (2013).

best course of actions  $a^g = 1$  and  $a^b = 0$ . Hence, if the enforcer selects the legal standard for a given investment, a discriminating rule dominates for any value of the probability  $\beta$ .

Notice that the first best course of actions is also feasible in the benchmark model, but it is not optimal, being replaced by  $a^g = 1$  and  $a^b > 0$ . Indeed, this way the enforcer reduces the negative impact on  $I$ . When, instead, the policy is chosen once the investment has been sunk, there is no reason to distort  $a^b$  upwards. We conclude that potentially there is a time inconsistency issue that the enforcer can solve by committing to the policy and legal standard before the investment is chosen, for instance by adopting regulations or guidelines, or through precedents. This result is summarized in the following proposition:<sup>21</sup>

**Proposition 2 (Sunk investment)** *Although ex post a flexible discriminating rule would be preferred, ex ante it may be optimal to commit to a more rigid per-se legality rule.*

## 4 Patent policy, limited fines and physical capital

In this section we extend the baseline model in three directions. First, we include a positive effect of the new technology on profits and welfare independently of the practice adopted. In this framework, these extra profits may be considered the result of patent protection, while additional profits can be obtained through the practice and are affected by the antitrust policy. This setting allows us to study whether patent policy and antitrust intervention play a complementary role or act as substitutes in the policy design. Secondly, we investigate how the choice of the optimal legal regime is affected by a cap on fines, in the form of a limited liability constraint. Finally, we show that our results on the optimal legal standard extend from the case of (uncertain) investment in research to the case of (deterministic) investment in physical assets.

### 4.1 Antitrust policy v. patent policy

In this section we extend the baseline model to include a fixed and positive effect of innovation on profits ( $\Pi$ ) and welfare ( $W$ ), that adds up to the impact of the practice

---

<sup>21</sup>To save space we omit to prove the result, that should be evident from the discussion and the previous results. A formal proof is available upon request.

on private and social payoffs. Formally, if the research investment is successful, the firm's profits are  $\Pi(a) = \Pi + \pi a$ , while welfare in the good and the bad state is, respectively,  $W^g(a) = W + w^g a$  and  $W^b(a) = W - w^b a$ . In this setting, we can interpret the consumers' surplus  $W - \Pi$ , as an inverse measure of the degree of protection granted to the innovative firm by the patent policy. The case  $W = \Pi$  corresponds to full protection, when the innovator does not fear any imitation by competitors and fully appropriates the benefits without transferring any surplus to consumers or rivals. Conversely, when  $W > \Pi = 0$  all the benefits accrue to consumers while the innovating firm is unable to retain any rent from the new technology, being immediately free raided by the rivals. This simple extension allows to study in a unified way the interaction between patent policy (affecting the fixed effects) and antitrust intervention ( influencing the variable part that depends on the practice).

We impose the following restrictions on the parameters:<sup>22</sup>

$$\Pi + \pi > W > \Pi > 0.$$

Apart from the fixed effects, the model remains the same as in the benchmark case. Hence, we briefly sketch the differences in the analysis.<sup>23</sup>

Under per-se rules, the optimal investment and expected welfare are

$$I_{PS} = \Pi + \pi \tilde{a} - \underline{f}, \quad (10)$$

and

$$EW_{PS}(\beta) = I_{PS}(W + Ew(\beta)\tilde{a}) - \frac{I_{PS}^2}{2}. \quad (11)$$

When, instead, a discriminating rule applies, the investment is

$$I_D = \Pi + (1 - \beta) [\pi - \varepsilon^I \underline{f}] + \beta [\pi \tilde{a}^b - (1 - \varepsilon^{II}) \underline{f}],$$

while the expected welfare becomes

$$EW_D = I \left[ W + \Delta W_D - \frac{I_D}{2} \right] - \frac{\gamma}{2} (\bar{\varepsilon} - \varepsilon^I)^2 - \frac{\gamma}{2} (\bar{\varepsilon} - \varepsilon^{II})^2. \quad (12)$$

---

<sup>22</sup>Since a complete analysis of all possible parameter regions is beyond the scope of this section, we concentrate on the most interesting case where the benefits from innovation accrue both to consumers and to the firm ( $W > \Pi > 0$ ) and antitrust policy is relevant ( $\Pi + \pi > W$ ). Moreover, this case is consistent with the benchmark model ( $\pi > 0$ ) when  $W$  and  $\Pi$  converge to 0.

<sup>23</sup>The expected welfare at the first best course of action is  $EW_{FB}(\beta, I) = I(1 - \beta)(W + w^g) - \frac{I^2}{2}$  that yields the optimal investment level  $I_{FB} = (1 - \beta)(W + w^g)$ . The expected welfare, evaluated at the first-best policies, is then  $EW_{FB}(\beta) = \frac{[(1 - \beta)(W + w^g)]^2}{2}$ .

We can now establish in the following proposition the optimal legal standards for different values of the likelihood of social harm,  $\beta$ .

**Proposition 3 (Optimal legal standards with fixed effects of the innovation)** *The optimal legal standard is a per-se legality rule for  $\beta \leq \beta'_0$  and a discriminating rule for higher  $\beta$ , where*

$$\beta'_0 = \beta_0 + \frac{(W - \Pi) - w^b \frac{\Pi}{\pi}}{w^g + w^b}.$$

Proposition 2 shows that, qualitatively, the results on optimal legal standards are the same as in the benchmark model. Per-se legality initially dominates, and is then replaced, for higher  $\beta$ , by an effect-based rule.<sup>24</sup>

When the degree of patent protection is reduced, i.e.  $W - \Pi$  is increased, the threshold  $\beta'_0$  shifts to the right and we observe an expansion of the region where antitrust policy opts for per-se legality, an extreme form of innovation-friendly antitrust intervention. In other words, patent and antitrust policies act as substitutes in our setting. This result is reported in the following corollary:

**Corollary 1 (Antitrust versus patent policy)** *Antitrust and patent policy are substitutes.*

## 4.2 Limited fines and the cost of flexible rules

So far we have assumed that the enforcer can use unlimited fines so as to save on costly accuracy. In this case, the potential weakness of discriminating rules, which more frequently lead to errors and may require to invest in accuracy, does not play a major role in the determination of the optimal legal standard. However, if fines are capped at some upper level, the enforcer, under a discriminating rule might be forced to change the mix of instruments, using more accuracy, with an increase in enforcement costs. In this section we explore how limited liability affects the optimal trade-off between per-se and discriminating rules.

According to Lemma 2 and Proposition 1, the optimal enforcement for  $\beta > \beta_0$  is a discriminating rule that progressively reduces the socially harmful practice  $\hat{a}^b$  and increases the fine  $\bar{f} = \frac{\pi(1-\hat{a}^b)}{(1-\bar{\varepsilon})}$  as  $\beta$  increases. At the same time, type-I accuracy is

---

<sup>24</sup>It is immediate to see that for  $W = \Pi = 0$ , the threshold and all the equilibrium expressions in Proposition 2 converge to the ones in Proposition 1.

improved to reduce the negative effect of the increasing fine on the investment in the good state. Let us now suppose that fines are subject to a limited liability constraint,  $F = \pi$ . When social harm is unlikely,  $\hat{a}^b$  is close to 1 and the fine  $\bar{f}$  is low. In this case, the limited liability constraint does not bind and the policy problem is equivalent to the one analyzed in Lemma 2. However, for  $\beta$  sufficiently large,  $\bar{f}$  cannot be set at the level required to implement the action in the unconstrained solution. More precisely, there will exist a  $\beta_3 > \beta_0$  such that  $\bar{f} = \pi$  and the limited liability constraint starts binding. For  $\beta > \beta_3$ ,  $\hat{a}^b$  becomes a function of the type-II error  $\varepsilon^{II}$ , as can be seen setting  $\underline{f} = 0$  in the lower bound of (7) to get<sup>25</sup>

$$\hat{a}^b = \varepsilon^{II}. \quad (13)$$

By reducing  $\varepsilon^{II}$  (collecting evidence on the variables that help to better estimate the signal in the bad state), the enforcer is able to implement a lower (less damaging) action  $\hat{a}^b$ , improving marginal deterrence. The following lemma states the optimal policy under discriminating rule and limited liability.

**Lemma 3 (Optimal enforcement policy under discriminating rule and limited liability)** *Under a discriminating rule, there exists a  $\beta_3 > \beta_0$  such that the limited liability constraint  $\bar{f} \leq \pi$  does not bind for  $\beta \in [0, \beta_3]$  when  $\bar{f}$  is optimally set. In this interval the optimal policy is the one described in Lemma 2. Instead, for  $\beta \in (\beta_3, 1]$  and  $\gamma$  is sufficiently large the optimal policy entails more symmetric accuracies ( $\varepsilon^I < \bar{\varepsilon}$  and  $\varepsilon^{II} < \bar{\varepsilon}$ ). The actions implemented are  $\hat{a}^b = \varepsilon^{II}$  and  $a^g = 1$ . The expected welfare  $EW_D(\beta)$  is decreasing in  $\beta$  and negative for  $\beta \rightarrow 1$ .*

It is interesting to observe that when the limited liability constraint binds, the enforcer implements a balanced reduction in both errors, a lower type-I error to sustain the investment softening *ex ante* deterrence on innovative effort in the good state, and more type-II accuracy to improve *ex post* deterrence on actions in the bad state.

In the following proposition we summarize the optimal legal standards.

**Proposition 4 (Optimal legal standards under limited liability)** *When fines are capped by limited liability, the optimal legal standard for increasing values of  $\beta$  is:*

---

<sup>25</sup>The same qualitative argument applies for any  $F \in \left(\pi, \frac{\pi}{(1-\bar{\varepsilon})}\right)$ . When  $F$  is capped in the interval above, the implementable action in the bad state is  $\hat{a}^b = 1 - (1 - \varepsilon^{II})\frac{F}{\pi}$ .



- i) for  $\beta \in [0, \beta_0)$  per-se legality ;*
- ii) for  $\beta \in [\beta_0, \beta_3)$  the discriminating rule with type-I accuracy;*
- iii) for  $\beta \in [\beta_3, \beta_4)$  the discriminating rule with the limited liability constraint binding and more balanced accuracy on both errors;*
- iv) for  $\beta \in [\beta_4, 1]$  per-se illegality.*

Up to the threshold  $\beta_3$  the limited liability constraint does not bind, and the results correspond to the case in Proposition 1. When the likelihood of social harm increases above  $\beta_3$ , the (unconstrained) optimal fine  $\bar{f}$  exceeds the admitted threshold  $\pi$ . In this region, the dominant legal standard is still initially the discriminating rule, realized combining the maximum fine admitted with a reduction in both errors. When the social loss is very likely ( $\beta > \beta_4$ ), the expected welfare becomes negative under a discriminating rule due to the high accuracy costs, and the more rigid per-se illegality rule replaces the discriminating rule, saving on accuracy cost although discouraging the practice in the (unlikely) good state.

In the previous proposition we identify two different reasons for a rigid per-se rule to dominate an effect-based regime. The first, observed in the baseline model, refers to more effective *ex ante* incentives to sustain investment, that make per-se legality more attractive than a discriminating rule when the likelihood of social harm is low. In this case, a more rigid rule allows the enforcer to commit to be *ex post* lenient when the practice is socially harmful, to the benefit of *ex ante* investment.

The second reason rests on a cost saving argument: a discriminating rule better adapts to *ex post* effects, but it requires more information and is therefore more prone to errors than a simpler, per-se rule. When fines are unlimited, this potential weakness plays a minor role, since fines act as substitutes to accuracy. When, however, fines are capped, the mix of policy instruments under a discriminating rule requires to further refine accuracy, making this regime more costly. When the practice is very likely to be harmful, then, a per-se illegality regime that completely deters it, destroying also the *ex ante* incentives to invest, dominates a discriminating rule.

An interesting feature of our results refers to accuracy. We have seen that type-II accuracy can improve deterrence on actions, while the reduction of type-I error may sustain innovative investments. The possibility of refining type-I or type-II accuracy rests on the following argument. A practice may be welfare enhancing (good state) or detrimental (bad state). Each of the two possibilities can be analyzed within

an appropriate model, and their empirical predictions suggest a set of observables. As long as the two sets of predictions are, at least in part, distinct, we can obtain identifying restrictions that allow to validate either of the two explanations.<sup>26</sup> Then, the enforcer can collect a minimum of information – facing the default probabilities of errors ( $\bar{\varepsilon}$ ) – or enrich the set of evidence. As long as the enforcer collects information on the (empirical) predictions of the competitive model, she is able to refine the assessment of the efficiency-enhancing effects, reducing the probability of condemning an innocent firm, that is a type-I error. This corresponds to reducing the variance of the probability distribution of the signal conditional on the good state. Conversely, additional evidence of the anti-competitive explanation implements a better type-II accuracy, and reduces the variance of the probability distribution of the signal conditional on the bad state. Finally, collecting evidence on both sets of observables symmetrically improves the accuracy on both errors.<sup>27</sup>

### 4.3 Investment in physical capital

In the benchmark model the firm invests in research activity, the outcome of the investment is uncertain, and leads to a new discovery with a probability proportional to the investment itself. In this section, instead, we explore a different type of investment, where the outcome is deterministic and the size of the investment is chosen by the firm. The most natural reference are investments in physical capital, as for instance building a broadband network. The firm, in this setting, decides the size of the investment  $I$  and the gross profits are proportional to the size of the investment itself. The profits from the broadband services are indeed increasing in the size of the network installed, that determines the number of (potential) clients and the range of services that can be offered. We maintain the assumption that profits are concave in the investment (decreasing returns) by assuming, as in the benchmark model, that the investment costs are increasing in its size.

---

<sup>26</sup>See Polo (2010) for an application to selective price cuts.

<sup>27</sup>Our analysis of the optimal enforcement policy has focussed on the choice of type-I and type-II accuracy, that can be chosen independently by the enforcer, while maintaining fixed the burden of proof (the threshold  $x$  of the signal  $\sigma$ ). Kaplow (2011b) instead analyzes the case when the enforcer controls the minimum strength of evidence  $x$  required to sanction a firm. In this case the enforcer faces a trade-off between a higher (lower) probability of type-I error and a lower (higher) probability of type-II errors. In other words, while setting accuracies gives the enforcer the possibility of choosing, at least to a certain extent, type-I and type-II errors independently, changing the burden of proof allows for a specific, inversely related, combination of type-I and type-II errors.

Moreover, as before, the firm exploits the potential profits of the investment by designing business strategies, that is choosing the action  $a \in A$ . For instance, the firm can impose specific restrictions on the access of competitors to the broadband network, either in terms of technical access or to access pricing and margin squeeze, including an extreme form of refusal to deal.

The profits, net of the investment costs, are therefore<sup>28</sup>  $\Pi(a, I) = I\pi a - \frac{I^2}{2}$ .

The social effects of the practice may be positive (good state) or negative, depending on the market conditions at the time the practice is undertaken, and are proportional to the investment size:  $W^b(a, I) = -Iw^b a \leq 0$  when the practice reduces welfare and  $W^g(a, I) = Iw^g a \geq 0$  when it is welfare enhancing. A more extended broadband network has larger positive or negative welfare effects, depending on market conditions. The assumptions regarding information, legal standards, policy tools and the timing remain the same as in the benchmark model.

Although so far the case of physical capital may seem just a reinterpretation of the benchmark model, once we solve for the optimal investment, an important difference arises. When the firm is involved in physical investment, whose outcome is deterministic, its *ex post* realized profits depend on the size of the investment ( $I\pi a$ ), contrary to the case of research investment, where the *ex ante* (gross profits) are  $I\pi a$  but the *ex post* profits in case of successful innovation are given by  $\pi a$ .

Consider first the per-se rules, where the enforcer implements the same action  $\tilde{a}$  in both states. The incentive compatibility and undertake constraints, taken together, give the inequalities:  $I\pi\tilde{a} - \underline{f} \geq 0 \geq I\pi - \bar{f}$ . The net profits at time 2 are therefore  $E\Pi_{PS} = I\pi\tilde{a} - \underline{f} - \frac{I^2}{2}$  and the firm chooses the profit maximizing investment

$$I_{PS} = \pi\tilde{a}. \quad (14)$$

Analogously, under a discriminating rule, the enforcer implements  $a^g = 1$  and  $a^b = \tilde{a}^b$  in the two states. Moreover the incentive compatibility and undertake constraints give the following inequalities:  $I\pi\tilde{a}^b - (1 - \varepsilon^{II})\underline{f} \geq 0 \geq I\pi - (1 - \varepsilon^{II})\bar{f}$  in the bad state and  $\pi\tilde{a}^b - \varepsilon^I \underline{f} \leq 0 \leq \pi - \varepsilon^I \bar{f}$  in the good state, leading to the following restrictions on the fines:

$$\bar{f} \in \left[ \underline{f} + \frac{I\pi(1 - \tilde{a}^b)}{1 - \varepsilon^{II}}, \underline{f} + \frac{I\pi(1 - \tilde{a}^b)}{\varepsilon^I} \right]. \quad (15)$$

---

<sup>28</sup>Notice that this expression corresponds, in the benchmark model, to the profits, gross of any fine, evaluated at the time the investment  $I$  is sunk.

At time 2 the expected profits for a firm that chooses  $a^g = 1$  and  $a^b = \tilde{a}^b$  are

$$E\Pi_D = (1 - \beta) [I\pi - \varepsilon^I \underline{f}] + \beta [I\pi\tilde{a}^b - (1 - \varepsilon^I)\underline{f}] - \frac{I^2}{2}$$

and the optimal investment in physical assets is therefore

$$I_D = (1 - \beta)\pi + \beta\pi\tilde{a}^b. \quad (16)$$

We can notice that, both for per-se and discriminating rules, when investment leads to a deterministic outcome (physical assets), it does not directly depend on fines and errors, contrary to the case of investment with a random outcome (research). However, the indirect effect of enforcement on investment, that takes place through the control of the implemented action  $\tilde{a}^b$ , continues to work in the case of physical capital.

The difference between research and physical investment comes from the different nature of the investment activity, whose outcome is uncertain in case of research while it is deterministic in case of physical assets. In both cases, the optimal choice requires to equate the marginal cost of investment and its marginal benefit. This latter term, in case of research investment, includes the fines, that instead have no marginal effect when investing in physical capital. Indeed, in the case of research activity, the firm realizes that it will pay  $\underline{f}$  only if research is successful. Then, a higher investment increase the probability of paying the fine, reducing the marginal benefit of the investment. When physical investment is involved, instead, the firm anticipates that it will pay the same fine  $\underline{f}$  in any case and for any positive level of  $I$ , with no marginal effect on the incentives to invest.

Finally, the expected welfare both for per-se and discriminating regimes has the same expression as in the benchmark case. Although the optimal enforcement policies are slightly different, the result in terms of optimal legal standards is identical to Proposition 1.

**Proposition 5 (Optimal legal standards in case of physical investment)**

*When the investment is deterministic (physical investment), the optimal legal standard is a per-se legality rule for  $\beta \leq \beta_0$  and a discriminating rule for higher  $\beta$ .*

Hence, our result obtained in the case of (uncertain) investment in research extends to the case of (deterministic) investment in physical assets. In both cases, when the expected welfare effects of a practice are sufficiently positive the enforcer prefers

to commit to a rigid per-se legality rule as a tool not to intervene ex-post in the unlikely case that the practice is harmful, thereby sustaining the (research or physical) investment. A more flexible discriminating rule, instead, is preferred when the effects of the practice are more mixed, and a combination of control on the practice and on the investment is required.

## 5 Conclusions

We have shown in this paper that the optimal legal standards and enforcement policies in antitrust intervention depend on the parameters that summarize the economic model, or the presumptions, of the enforcer. In this sense, legal standard, level of accuracy and fine schedule all depend on the priors of the enforcer regarding the economic effects of the practices, i.e. on the parameters  $(\beta, w^g, w^b, \pi)$ . Under this respect, our results recall the debate briefly summarized in the introduction. Economic approaches that have stressed the efficiency enhancing effects of many business practices (a low  $\beta$ ), as those proposed by the Chicago school, have also campaigned for per-se legality rules, while a more articulated reconstruction of the competitive and anticompetitive effects of those practices (a higher  $\beta$ ), usually associated to the post-Chicago scholars, has represented the background for the effect-based approach to unilateral practices.

## Appendix

**Proof of Lemma 1.** We solve our problem by omitting the undertake constraint and verifying it *ex post*. The maximization program is solved by the following first-order conditions

$$\frac{\partial EW_{PS}}{\partial \tilde{a}} = [Ew(\beta)\tilde{a} - I_{PS}] \pi + Ew(\beta)I_{PS} + \lambda \geq 0, \quad (17)$$

$$\frac{\partial EW_{PS}}{\partial \underline{f}} = -[Ew(\beta)\tilde{a} - I_{PS}] - \frac{\lambda}{\pi} \leq 0, \quad (18)$$

$$\frac{\partial EW_{PS}}{\partial \bar{f}} = \frac{\lambda}{\pi} \geq 0, \quad (19)$$

Finally, the complementary slackness condition is

$$\lambda \left( \tilde{a} - 1 + \frac{\bar{f} - \underline{f}}{\pi} \right) = 0. \quad (20)$$

First of all, notice that the incentive compatibility constraint does not bind, so that  $\lambda = 0$ . In fact, if it were  $\lambda > 0$ , then  $\bar{f} = F$  and  $\lambda$  should be zero to satisfy the complementary slackness condition, leading to a contradiction. Since  $\lambda = 0$ , the high fine  $\bar{f}$  can be any value satisfying the incentive compatibility constraint.

Notice that  $Ew(\beta) - \pi \geq 0$  if  $\beta \leq \beta_1 = \frac{w^g - \pi}{w^g + w^b}$  and  $Ew(\beta) \geq 0$  if  $\beta \leq \beta_2 = \frac{w^g}{w^g + w^b}$ . Then we have three possible cases:

(i) For  $\beta \in [0, \beta_1]$  we have  $Ew(\beta) > Ew(\beta) - \pi \geq 0$ . Then, if we set  $\underline{f} = 0$ , the investment is  $I_{PS} = \pi \tilde{a}$  and, substituting in the first order conditions, we get  $\frac{\partial EW_{PS}}{\partial \underline{f}} = -[Ew(\beta) - \pi] \tilde{a} < 0$  and setting  $\underline{f} = 0$  is optimal. Moreover,  $\frac{\partial EW_{PS}}{\partial \tilde{a}} = \{[Ew(\beta) - \pi] + Ew(\beta)\} \pi \tilde{a} > 0$  and  $\hat{a} = 1$ .

(ii) For  $\beta \in (\beta_1, \beta_2)$ ,  $Ew(\beta) > 0 > Ew(\beta) - \pi$  and the first order condition  $\partial EW_{PS} / \partial \underline{f} = 0$  holds for  $Ew(\beta) \tilde{a} - I_{PS} = 0$ . Then  $\frac{\partial EW_{PS}}{\partial \tilde{a}} = Ew(\beta) I_{PS} > 0$  and  $\hat{a} = 1$ . Substituting in  $\partial EW_{PS} / \partial \underline{f} = 0$  and solving we get  $\underline{f} = \pi - Ew(\beta) > 0$ . Substituting  $\underline{f}$  in the expression of the optimal investment we obtain  $I_{PS} = Ew(\beta) > 0$  that is decreasing in  $\beta$  and equal to 0 when  $\beta = \beta_2$ .

(iii) For  $\beta \in [\beta_2, 1]$ ,  $0 \geq Ew(\beta) > Ew(\beta) - \pi$  implying that  $\partial EW_{PS} / \partial \tilde{a} = \partial EW_{PS} / \partial \underline{f} = 0$ . It is immediate to see that the only values of the action and low fine that satisfy both equalities are  $\hat{a} = 0$  and  $\underline{f} = 0$ . Moreover, the incentive compatibility constraint is satisfied for any  $\bar{f} \geq \pi$ .

It is immediate to see that in all three cases the undertake constraint is satisfied.

■

**Proof of Lemma 2.** The proof is organized as follows. First, we identify the equilibrium value of the policy variables; then we analyze the comparative statics of  $\hat{a}^b$  and  $EW_D$  with respect to  $\beta$ . We solve our problem by omitting the incentive compatibility constraints (7) and the undertake constraints and verifying them *ex*

post. The first order conditions are the following

$$\begin{aligned}
\frac{\partial EW_D}{\partial \hat{a}^b} &= [\Delta W_D - I_D] \beta \pi - \beta w^b I_D \geq 0 \\
\frac{\partial EW_D}{\partial \underline{f}} &= -[\Delta W_D - I_D] \beta (1 - \varepsilon^{II}) < 0 \\
\frac{\partial EW_D}{\partial \bar{f}} &= -[\Delta W_D - I_D] (1 - \beta) \varepsilon^I < 0 \\
\frac{\partial EW_D}{\partial \varepsilon^I} &= -[\Delta W_D - I_D] (1 - \beta) \bar{f} + \gamma (\bar{\varepsilon} - \varepsilon^I) \geq 0 \\
\frac{\partial EW_D}{\partial \varepsilon^{II}} &= [\Delta W_D - I_D] \beta \underline{f} + \gamma (\bar{\varepsilon} - \varepsilon^{II}) \geq 0,
\end{aligned}$$

where  $\Delta W_D = (1 - \beta)w^g - \beta w^b \hat{a}^b$  and  $I_D$  is given by (8).

Let us consider the following candidate solution and check in which interval of  $\beta$  it holds:  $\underline{f} = \bar{f} = 0$  and  $\hat{a}^b = 1$ . Substituting we have  $I_D = \pi$  and  $\Delta W_D - I_D = [w^g - \pi - \beta(w^g + w^b)] > 0$  for  $\beta < \frac{w^g - \pi}{w^g + w^b} = \beta_1$ . Moreover, for  $\beta < \frac{w^g - w^b - \pi}{w^g + w^b} = \beta_0 < \beta_1$ ,  $\frac{\partial EW_D}{\partial \hat{a}^b} = \beta \pi [w^g - w^b - \pi - \beta(w^g + w^b)] > 0$ . Hence, for  $\beta < \beta_0$ ,  $\frac{\partial EW_D}{\partial \hat{a}^b} > 0$ ,  $\frac{\partial EW_D}{\partial \underline{f}} < 0$  and  $\frac{\partial EW_D}{\partial \bar{f}} < 0$  at  $\varepsilon^I = \varepsilon^{II} = \bar{\varepsilon}$ . Finally, the incentive compatibility constraints (7) and the undertake constraints are clearly satisfied. The expected welfare is therefore  $\pi [Ew(\beta) - \frac{\pi}{2}]$ . Notice that this outcome is equivalent to the one under per-se legality.

Consider next the case  $\beta > \beta_0$ . We set  $\hat{a}^b < 1$  to obtain  $\frac{\partial EW_D}{\partial \hat{a}^b} = 0$ , implying that  $\Delta W_D - I_D > 0$ . Then  $\frac{\partial EW_D}{\partial \underline{f}} < 0$  and we get  $\underline{f} = 0$ . Since  $\underline{f} = 0$  we have  $\frac{\partial EW_D}{\partial \varepsilon^{II}} = \gamma (\bar{\varepsilon} - \varepsilon^{II}) = 0$  at  $\varepsilon^{II} = \bar{\varepsilon}$ . Moreover,  $\frac{\partial EW_D}{\partial \varepsilon^I} = 0$  for  $\varepsilon^I < \bar{\varepsilon}$ . Finally,  $\frac{\partial EW_D}{\partial \bar{f}} < 0$  implies that  $\bar{f}$  is determined by the lower bound of the constraint (7), that is,  $\bar{f} = \frac{\pi(1 - \hat{a}^b)}{(1 - \bar{\varepsilon})}$ . Finally, notice that the undertake constraints are satisfied since  $\pi \hat{a}^b - 0 \geq 0$  and  $\pi - \varepsilon^I \bar{f} = \pi - \varepsilon^I \frac{\pi(1 - \hat{a}^b)}{(1 - \bar{\varepsilon})} > \pi \left[ \frac{1 - \bar{\varepsilon} - \varepsilon^I}{1 - \bar{\varepsilon}} \right] > 0$ .

To check the second order condition, notice that only  $\hat{a}^b$  and  $\varepsilon^I$  are set at an internal solution. Hence,

$$\begin{aligned}
\frac{\partial^2 EW_D}{\partial \hat{a}^{b2}} &= -\beta^2 \pi (2w^b + \pi) < 0 \\
\frac{\partial^2 EW_D}{\partial \varepsilon^{I2}} &= -(1 - \beta)^2 \bar{f}^2 - \gamma < 0 \\
H_{\hat{a}^b \varepsilon^I} &= \beta^2 \left[ -2w^{b2} (1 - \beta)^2 \bar{f}^2 + \pi (2w^b + \pi) \gamma \right] > 0
\end{aligned}$$

for  $\gamma$  sufficiently large.

Let us now turn to the comparative statics of  $\hat{a}^b$  with respect to  $\beta$ . For  $\beta > \beta_0$ , rearranging from the first order conditions we get the following expressions of the implemented action and investment as a function of the optimal type-I error

$$\hat{a}^b = \frac{(1 - \beta) [(1 - \bar{\varepsilon})w^g - ((1 - \bar{\varepsilon} - \varepsilon^I)(w^b + \pi))]}{(1 - \beta)\varepsilon^I(w^b + \pi) + \beta(1 - \bar{\varepsilon})(2w^b + \pi)}$$

and

$$I_D = \pi \frac{(1 - \beta) [\beta((1 - \bar{\varepsilon} - \varepsilon^I)(w^g + w^b) + \varepsilon^I w^g)]}{(1 - \beta)\varepsilon^I(w^b + \pi) + \beta(1 - \bar{\varepsilon})(2w^b + \pi)}.$$

with  $\hat{a}^b \rightarrow 1$  and  $I_D \rightarrow \pi$  for  $\beta \rightarrow \beta_0$  and  $\hat{a}^b \rightarrow 0$  and  $I_D \rightarrow 0$  for  $\beta \rightarrow 1$ . Moreover, it is easy to check that  $\hat{a}^b$  is increasing in  $\bar{\varepsilon}$  and  $\varepsilon^I$  for  $\beta > \beta_0$ . Therefore, the expected welfare tends to 0 when  $\beta \rightarrow 1$ . Notice that the expressions above are not the equilibrium value since they both depend on the equilibrium level of type-I error  $\varepsilon^I$ , and they can be evaluated only at the extremes of the interval. To further analyze the effect of  $\beta$  on the equilibrium value of  $\hat{a}^b$  we can differentiate the first order conditions with respect to  $\hat{a}^b$ ,  $\varepsilon^I$  and  $\beta$ . Then, we have that  $\text{sign} \frac{d\hat{a}^b}{d\beta} = \text{sign}(-\frac{\partial^2 EW_D}{\partial \beta \partial \hat{a}^b} \frac{\partial^2 EW_D}{\partial \varepsilon^{I2}} + \frac{\partial^2 EW_D}{\partial \varepsilon^I \partial \hat{a}^b} \frac{\partial^2 EW_D}{\partial \varepsilon^I \partial \beta})$  where  $\frac{\partial^2 EW_D}{\partial \varepsilon^I \partial \hat{a}^b} = \frac{\partial I_D}{\partial \varepsilon^I} [-\beta w^b - \frac{1}{2} \frac{\partial I_D}{\partial \hat{a}^b}] + \frac{\partial^2 I_D}{\partial \varepsilon^I \partial \hat{a}^b} [\Delta W_D - I_D] > 0$ , since  $\frac{\partial^2 I_D}{\partial \varepsilon^I \partial \hat{a}^b} = \frac{(1 - \beta)\pi}{(1 - \bar{\varepsilon})} > 0$ ,  $\frac{\partial I_D}{\partial \hat{a}^b} > 0$  and  $\frac{\partial I_D}{\partial \varepsilon^I} < 0$ .  $\frac{\partial^2 EW_D}{\partial \varepsilon^I \partial \beta} = \frac{\pi(1 - \hat{a}^b)}{(1 - \bar{\varepsilon})} [\Delta W_D - I_D] - \frac{(1 - \beta)\pi(1 - \hat{a}^b)}{(1 - \bar{\varepsilon})} [-w^g - w^b \hat{a}^b - \frac{1}{2} \frac{\partial I_D}{\partial \beta}] > 0$ .

Finally,

$$\frac{\partial^2 EW_D}{\partial \hat{a}^b \partial \beta} = \pi [\Delta W_D - I_D] - w^b I_D - \beta w^b \frac{\partial I_D}{\partial \beta} + \beta \pi \frac{\partial [\Delta W_D - I_D]}{\partial \beta}.$$

Multiplying the previous expression by  $\beta$  we notice that

$$\beta \frac{\partial^2 EW_D}{\partial \hat{a}^b \partial \beta} = \frac{\partial EW_D}{\partial \hat{a}^b} + \beta^2 \left[ -w^b \frac{\partial I_D}{\partial \beta} + \pi \frac{\partial [\Delta W_D - I_D]}{\partial \beta} \right],$$

where the first term is zero (envelope theorem). The term in square brackets can then be rewritten as

$$\beta^2 \pi \left[ - (w^b + \pi) (\hat{a}^b - 1) \left( \frac{1 - \bar{\varepsilon} - \varepsilon^I}{1 - \bar{\varepsilon}} \right) - (w^g + w^b \hat{a}^b) \right],$$

or equivalently as

$$\beta^2 \pi \left[ - (w^b + \pi) \hat{a}^b \left( \frac{1 - \bar{\varepsilon} - \varepsilon^I}{1 - \bar{\varepsilon}} \right) - w^b \hat{a}^b - \left( w^g - (w^b + \pi) \left( \frac{1 - \bar{\varepsilon} - \varepsilon^I}{1 - \bar{\varepsilon}} \right) \right) \right] < 0,$$



since  $\left(\frac{1-\bar{\varepsilon}-\varepsilon^I}{1-\bar{\varepsilon}}\right)$  is smaller than one and  $w^g > w^b + \pi$ . Then,  $\frac{\partial^2 EW_D}{\partial \hat{a}^b \partial \beta} < 0$  and  $\frac{d\hat{a}^b}{d\beta} < 0$  when  $\gamma$  (that is in the expression for  $\frac{\partial^2 EW_D}{\partial \varepsilon^{I2}}$ ) is sufficiently large. Hence,  $\hat{a}^b$  decreases from 1 to 0 as  $\beta$  varies from  $\beta_0$  to 1.

Finally, differentiating with respect to  $\beta$  the expected welfare we get

$$\frac{dEW_D}{d\beta} = \frac{\partial EW_D}{\partial \beta} + \frac{\partial EW_D}{\partial \varepsilon^I} \frac{\partial \varepsilon^I}{\partial \beta} + \frac{\partial EW_D}{\partial \hat{a}^{bR}} \frac{\partial \hat{a}^b}{\partial \beta},$$

where the first term (direct effect) is negative and the last two terms are zero due to the FOC (envelope theorem). Indeed,

$$\frac{\partial EW_D}{\partial \beta} = \frac{\partial I_D}{\partial \beta} [(1-\beta)w^g - \beta w^b \hat{a}^b - I_D/2] + I_D \left[ -w^g - w^b \hat{a}^b - \frac{1}{2} \frac{\partial I_D}{\partial \beta} \right] < 0,$$

is negative because  $\frac{\partial I_D}{\partial \beta} = -\frac{1-\varepsilon^I-\bar{\varepsilon}}{(1-\bar{\varepsilon})} \pi(1-\hat{a}^b)$  is negative and the same is true for the term in the second square bracket. Hence,  $EW_D(\beta)$  is decreasing in  $\beta$ . ■

**Proof of Proposition 1.** Given Lemma 1, the per-se rules give

$$\begin{aligned} & \pi \left[ Ew(\beta) - \frac{\pi}{2} \right] \text{ for } \beta \in [0, \beta_1], \\ & \frac{[Ew(\beta)]^2}{2} \text{ for } \beta \in (\beta_1, \beta_2), \\ & 0 \text{ for } \beta \in [\beta_2, 1]. \end{aligned}$$

Instead, the discriminating rule (Lemma 2) gives

$$\begin{aligned} & \pi \left[ Ew(\beta) - \frac{\pi}{2} \right] \text{ for } \beta \in [0, \beta_0] \\ & EW_D(\beta, \varepsilon^I(\beta), \hat{a}^b(\beta)) \text{ for } \beta \in (\beta_0, 1]. \end{aligned}$$

Let us compare the expected welfare in the different regimes for increasing values of  $\beta$ . For  $\beta \in [0, \beta_0]$ , both  $D$  and  $PS$  are equivalent to the per-se legality regime. In the interval  $(\beta_0, \beta_1]$  the discriminating rule, although it may still implement the per-se legality outcome, chooses a different policy, implying that  $EW_D(\beta, \varepsilon^I(\beta), \hat{a}^b(\beta)) > EW_{PS}(\beta)$ .

For  $\beta \in (\beta_1, \beta_2)$ , per-se rule implements  $a^g = a^b = 1$  and  $I = Ew(\beta)$  by setting  $\hat{a} = 1$  and,  $\underline{f} = [\pi - Ew(\beta)]$ . The same allocation can be implemented also under a discriminating rule by setting  $\varepsilon^I = \varepsilon^{II} = \bar{\varepsilon}$ ,  $\hat{a}^b = 1$  and  $\underline{f} = [\pi - Ew(\beta)] / [(1-\beta)\bar{\varepsilon} + \beta(1-\bar{\varepsilon})]$ , adjusting the fine with respect to the  $PS$  regime to take into account the errors. Although implementable, this allocation is

not optimal under a discriminating rule, and therefore  $EW_D(\beta) > EW_{PS}(\beta)$  in this interval. Finally, for  $\beta \in [\beta_2, 1]$ ,  $EW_{PS}(\beta) = 0$  while  $EW_D(\beta)$  is decreasing and equal to zero only at  $\beta = 1$ . ■

**Proof of Proposition 3.** Assume  $\Pi + \pi > W > \Pi$ . The proof of the statement requires first to derive the optimal enforcement policy under per-se (step 1) and discriminating (step 2) rules and then the selection of the optimal legal standard (step 3). Since the proof follows closely the ones in Section 3 we will only underline the main differences.

**Step 1.** First of all, we find the **optimal policy under per-se rules**. While the logic of the proof is as the one of Lemma 1 we find one more region and the thresholds are different. Therefore, we find it useful to go through it. Recall from the text the expressions for the optimal investment  $I_{PS} = \Pi + \pi\tilde{a} - \underline{f}$  and for the expected welfare  $EW_{PS}(\beta) = I_{PS}(W + Ew(\beta)\tilde{a}) - \frac{I_{PS}^2}{2}$ . Then, the maximization program is solved by the following first-order conditions

$$\begin{aligned}\frac{\partial EW_{PS}}{\partial \tilde{a}} &= [W + Ew(\beta)\tilde{a} - I_{PS}]\pi + Ew(\beta)I_{PS} + \lambda \geq 0, \\ \frac{\partial EW_{PS}}{\partial \underline{f}} &= -[W + Ew(\beta)\tilde{a} - I_{PS}] - \frac{\lambda}{\pi} \leq 0, \\ \frac{\partial EW_{PS}}{\partial \bar{f}} &= \frac{\lambda}{\pi} \geq 0,\end{aligned}\tag{21}$$

Finally, the complementary slackness condition is

$$\lambda \left( \tilde{a} - 1 + \frac{\bar{f} - \underline{f}}{\pi} \right) = 0.\tag{22}$$

First of all, notice that the incentive compatibility constraint does not bind, so that  $\lambda = 0$ . In fact, if it were  $\lambda > 0$ , then  $\bar{f} = F$  and  $\lambda$  should be zero to satisfy the complementary slackness condition, leading to a contradiction. Since  $\lambda = 0$ , the high fine  $\bar{f}$  can be any value satisfying the incentive compatibility constraint. we have four possible subcases:

(i) For  $\beta \in [0, \beta'_1]$  we have  $W + Ew(\beta) - \Pi - \pi \geq 0$  where  $\beta'_1 = \frac{w^g - \pi + W - \Pi}{w^g + w^b}$ . Moreover,  $W < \Pi + \pi$  implies that  $\beta'_1 < \beta_2$  so that  $Ew(\beta) \geq 0$  for  $\beta \in [0, \beta'_1]$ . Then, if we set  $\underline{f} = 0$  and  $\hat{a} = 1$ , the investment is  $I_{PS} = \Pi + \pi$  and, substituting in the first order conditions, we get  $\frac{\partial EW_{PS}}{\partial \underline{f}} = -[W + Ew(\beta) - \Pi - \pi] < 0$  and  $\frac{\partial EW_{PS}}{\partial \tilde{a}} = [W + Ew(\beta) - \Pi - \pi]\pi + Ew(\beta)\Pi > 0$  then setting  $\underline{f} = 0$  and

$\hat{a} = 1$  is optimal. Hence,  $\bar{f}$  is not needed to define the fine schedule. Finally,  $EW_{PS} = (\Pi + \pi) \left[ W + Ew(\beta) - \frac{\Pi + \pi}{2} \right]$ .

(ii) For  $\beta \in (\beta'_1, \beta_2)$ ,  $Ew(\beta) > 0 > W + Ew(\beta) - \Pi - \pi$  and the first order condition  $\partial EW_{PS} / \partial \underline{f} = 0$  holds for  $W + Ew(\beta)\tilde{a} - I_{PS} = 0$ . Then  $\frac{\partial EW_{PS}}{\partial \tilde{a}} = Ew(\beta)I_{PS} > 0$  and  $\hat{a} = 1$ . Substituting in  $\partial EW_{PS} / \partial \underline{f} = 0$  and solving we get  $\underline{f} = -W - Ew(\beta) + \Pi + \pi > 0$ . Substituting  $\underline{f}$  in the expression of the optimal investment we obtain  $I_{PS} = W + Ew(\beta) > 0$  that is decreasing in  $\beta$  and equal to  $W$  when  $\beta = \beta_2$ . Finally,  $EW_{PS} = \frac{(W + Ew(\beta))^2}{2}$ .

(iii) For  $\beta \in [\beta_2, \beta'_2)$ ,  $(W - \Pi)\pi + Ew(\beta)\Pi > 0 \geq Ew(\beta)$  where  $\beta'_2 = \frac{w^g + (W - \Pi)\pi}{w^g + w^b}$  and  $\beta'_2 > \beta'_1$  (using both  $W < \Pi + \pi$  and  $W > \Pi$ ). Then, if we set  $\frac{\partial EW_{PS}}{\partial \tilde{a}} = [W + Ew(\beta)\tilde{a} - \Pi - \pi\tilde{a}]\pi + Ew(\beta)(\Pi + \pi\tilde{a}) = 0$  we have that  $\underline{f} = 0$  since  $\frac{\partial EW_{PS}}{\partial \underline{f}} < 0$ ,  $\hat{a}$  is interior and equal to  $\frac{W - \Pi}{\pi - 2Ew(\beta)} + \frac{Ew(\beta)\Pi}{(\pi - 2Ew(\beta))\pi}$  and  $I_{PS} = \frac{\pi W - Ew(\beta)\Pi}{\pi - 2Ew(\beta)}$ . Finally,  $EW_{PS} = I_{PS} \left[ W + Ew(\beta)\hat{a} - \frac{I_{PS}}{2} \right]$ . Substituting  $\hat{a}$  and  $I_{PS}$  and rearranging we get  $EW_{PS} = \frac{[\pi W - Ew(\beta)\Pi]^2}{2\pi[\pi - 2Ew(\beta)]}$ . Differentiating w.r.t.  $\beta$  we get:

$$\begin{aligned} \frac{\partial EW_{PS}}{\partial \beta} &= \frac{2(w^g + w^b)(\pi W - Ew(\beta)\Pi)(\Pi - 2W)}{(\pi - 2Ew(\beta))} < 0 \\ \frac{\partial EW_{PS}^2}{\partial \beta^2} &= \frac{\pi(\Pi - 2W)^2}{(\pi - 2Ew(\beta))^2} > 0. \end{aligned}$$

Hence, in this region  $EW_{PS}$  is decreasing and convex in  $\beta$ .

(iv) For  $\beta \in [\beta'_2, 1]$ ,  $0 \geq (W - \Pi)\pi + Ew(\beta)\Pi > Ew(\beta)$  implying  $\partial EW_{PS} / \partial \tilde{a} < 0$  and  $\partial EW_{PS} / \partial \underline{f} < 0$ . So that  $\hat{a} = 0$ ,  $\underline{f} = 0$ . Substituting  $\hat{a}$  and  $\underline{f}$  in the expression for the optimal investment and for the expected welfare we obtain  $I_{PS} = \Pi$  and  $EW_{PS} = \Pi \left( W - \frac{\Pi}{2} \right) > 0$ . Moreover, the incentive compatibility constraint is satisfied for any  $\bar{f} \geq \pi$ . It is immediate to see that in all cases the undertake constraint is satisfied.

**Step 2.** Second, we find the **optimal policy under discriminating rules**. Recall from the text the expressions for the innovative investment  $I_D = \Pi + (1 - \beta) [\pi - \varepsilon^I \bar{f}] + \beta [\pi \tilde{a}^b - (1 - \varepsilon^{II}) \underline{f}]$  and for the expected welfare  $EW_D = I \left[ W + \Delta W_D - \frac{I_D}{2} \right] - \frac{\gamma}{2} (\bar{\varepsilon} - \varepsilon^I)^2 - \frac{\gamma}{2} (\bar{\varepsilon} - \varepsilon^{II})^2$ . The proof follows closely the one in Lemma 2. We only underline three differences: first, going through the same steps in Lemma 2 from  $\frac{\partial EW_D}{\partial \tilde{a}^b} = 0$  we find a new threshold  $\beta'_0 = \frac{w^g - w^b - \pi - w^b \frac{\Pi}{\pi} + (W - \Pi)}{w^g + w^b}$  (instead of  $\beta_0$ ) such that for  $\beta \leq \beta'_0$ :  $\hat{a}^b = 1$ ,  $\bar{f} = \underline{f} = 0$ ,  $I_D = \Pi + \pi$  and the expected welfare is  $(\Pi + \pi) \left[ W + Ew(\beta) - \frac{\Pi + \pi}{2} \right]$ . Notice that once again this outcome is equivalent to the one under per-se legality.

Second, differently from the proof of Lemma 2, showing that  $EW_D$  is decreasing in  $\beta$  is not enough to completely characterize the optimal legal standard. Indeed, we also need to show that for  $\beta > \beta'_0$  the expected welfare is concave in  $\beta$ .

Therefore, differentiating two times with respect to  $\beta$  the expected welfare we get

$$\frac{d^2 EW_D}{d\beta^2} = \frac{\partial^2 EW_D}{\partial \beta^2} + \frac{\partial^2 EW_D}{\partial \varepsilon^{I^2}} \left( \frac{\partial \varepsilon^I}{\partial \beta} \right)^2 + \frac{\partial^2 EW_D}{\partial \widehat{a}^{b^2}} \left( \frac{\partial \widehat{a}^b}{\partial \beta} \right)^2 + 2 \frac{\partial^2 EW_D}{\partial \varepsilon^I \partial \widehat{a}^b} \left( \frac{\partial \varepsilon^I}{\partial \beta} \frac{\partial \widehat{a}^b}{\partial \beta} \right),$$

where  $\frac{\partial^2 EW_D}{\partial \beta^2} = 2 \frac{\partial I_D}{\partial \beta} \left[ -w^g - w^b \widetilde{a}_D - \frac{1}{2} \frac{\partial I_D}{\partial \beta} \right] > 0$ ,  $\frac{\partial^2 EW_D}{\partial \varepsilon^{I^2}} < 0$ ,  $\frac{\partial^2 EW_D}{\partial \widehat{a}^{b^2}} < 0$  while  $\frac{\partial^2 EW_D}{\partial \varepsilon^I \partial \widehat{a}^b} \left( \frac{\partial \varepsilon^I}{\partial \beta} \frac{\partial \widehat{a}^b}{\partial \beta} \right)$  is ambiguous in sign. The expected welfare is then always decreasing in  $\beta$  and concave when  $\gamma$  (that is in the expression for  $\frac{\partial^2 EW_D}{\partial \varepsilon^{I^2}}$ ) is sufficiently large.

Third, differently from before the innovative investment and the expected welfare do not tend to zero when  $\beta$  goes to 1. Rather, they tend to the level prevailing under per-se regime, i.e.,  $I_D = \Pi$  and  $EW_D(1) = \Pi \left( W - \frac{\Pi}{2} \right)$ .

**Step 3.** We are now able to select the **optimal legal standard** by comparing the per-se and the discriminating rule, very much like in Proposition 1. Indeed the proof is the same except for the new region with  $\beta \in [\beta_2, \beta'_2)$ . To compare the regimes in this interval we need three pieces of information: i) First, remind that in this region  $EW_{PS}$  is decreasing and convex in  $\beta$ . ii) Second, as in the proof of Proposition 1 it is still true that for  $\beta \in (\beta'_0, \beta_2)$ , (in Proposition 1 the interval was  $\beta \in (\beta_0, \beta_2)$ ) under a discriminating rule the regulator could replicate the choice implemented by the per-se rule ( $a^g = a^b = 1$  and  $I = W + Ew(\beta)$ ). Although implementable, this allocation is not optimal under a discriminating rule, and therefore  $EW_D(\beta) > EW_{PS}(\beta)$  in this interval. iii) Moreover, for  $\beta \in [\beta'_2, 1]$ ,  $EW_{PS}(\beta) = \Pi \left( W - \frac{\Pi}{2} \right)$  while  $EW_D(\beta)$  is decreasing and equal to  $\Pi \left( W - \frac{\Pi}{2} \right)$  only for  $\beta = 1$ . Summing up,  $EW_D(\beta)$  lies above  $EW_{PS}(\beta)$  both at  $\beta = \beta_2$  and at  $\beta = \beta'_2$ , it is decreasing and concave (as shown in Step 2), while  $EW_{PS}(\beta)$  is decreasing and convex. Then, we can conclude that  $EW_D(\beta) > EW_{PS}(\beta)$  also in this interval. ■

**Proof of Lemma 3.** Combining the incentive compatibility and limited liability constraints by setting  $\bar{f} = \pi$  and  $\underline{f} = 0$  in (7) we obtain

$$\widehat{a}^b = \varepsilon^{II}$$

increasing in type-II error  $\varepsilon^{II}$ . Then, substituting the implementable actions in the expression of the investment we get

$$I_D = \pi \left[ 1 - \varepsilon^I - \beta(1 - \varepsilon^I - \varepsilon^{II}) \right].$$

with  $\frac{\partial I_D}{\partial \varepsilon^I} = -\pi(1 - \beta) < 0$  and  $\frac{\partial I_D}{\partial \varepsilon^{II}} = \pi\beta > 0$ . To find the optimal errors, we substitute the expressions for the action and the investment in the expected welfare.

The first order conditions are

$$\begin{aligned}\frac{\partial EW_D}{\partial \varepsilon^I} &= [\Delta W_D - I_D] \frac{\partial I_D}{\partial \varepsilon^I} + \gamma(\bar{\varepsilon} - \varepsilon^I) \geq 0 \\ \frac{\partial EW_D}{\partial \varepsilon^{II}} &= [\Delta W_D - I_D] \frac{\partial I_D}{\partial \varepsilon^{II}} - \beta w^b \frac{\partial \tilde{a}_D^b}{\partial \varepsilon^{II}} + \gamma(\bar{\varepsilon} - \varepsilon^{II}) \geq 0\end{aligned}$$

that hold as equalities with internal solutions  $\varepsilon^I < \bar{\varepsilon}$  and  $\varepsilon^{II} < \bar{\varepsilon}$ . Notice that for  $\bar{f} = \pi$ ,  $\underline{f} = 0$ ,  $\hat{a}^b = \varepsilon^{II}$  and  $a^g = 1$  the undertake constraints are also satisfied.

Finally, the second order conditions hold, since

$$\begin{aligned}\frac{\partial^2 EW_D}{\partial \varepsilon^{I2}} &= -\left(\frac{\partial I_D}{\partial \varepsilon^I}\right)^2 - \gamma < 0 \\ \frac{\partial^2 EW_D}{\partial \varepsilon^{II2}} &= -\left(\frac{\partial I_D}{\partial \varepsilon^{II}}\right)^2 - \gamma < 0 \\ H_{\varepsilon^I \varepsilon^{II}} &= \gamma \left[ \left(\frac{\partial I_D}{\partial \varepsilon^I}\right)^2 + \left(\frac{\partial I_D}{\partial \varepsilon^{II}}\right)^2 \right] + \gamma^2 > 0.\end{aligned}$$

Differentiating with respect to  $\beta$  the expected welfare we get

$$\frac{dEW_D}{d\beta} = \frac{\partial EW_D}{\partial \beta} + \frac{\partial EW_D}{\partial \varepsilon^I} \frac{\partial \varepsilon^I}{\partial \beta} + \frac{\partial EW_D}{\partial \varepsilon^{II}} \frac{\partial \varepsilon^{II}}{\partial \beta},$$

where the first term (direct effect) is negative and the last two terms are zero due to the FOC (envelope theorem). Indeed,

$$\frac{\partial EW_D}{\partial \beta} = \frac{\partial I_D}{\partial \beta} [(1 - \beta)w^g - \beta w^b \varepsilon^{II} - I_D/2] + I_D \left[ -w^g - w^b \varepsilon^{II} - \frac{1}{2} \frac{\partial I_D}{\partial \beta} \right] < 0,$$

is negative because  $\frac{\partial I_D}{\partial \beta} = -\pi(1 - \varepsilon^I - \varepsilon^{II})$  is negative and the same is true for the term in the second square bracket. Finally, evaluating the expected welfare at  $\beta = 1$  we obtain  $EW_D(1) = -\varepsilon^{II2} \pi (w^b + \frac{\pi}{2}) < 0$ . ■

**Proof of Proposition 5.** We first derive the optimal policies under per-se rules and discriminating rules, and then select the optimal legal standards.

**Per-se rules:** since the optimal investment is  $I_{PS} = \pi \tilde{a}$ , the first order conditions are now, after rearranging:

$$\frac{\partial EW_{PS}}{\partial \tilde{a}} = [2Ew(\beta) - \pi] \pi \tilde{a} + \lambda \geq 0, \quad (23)$$

$$\frac{\partial EW_{PS}}{\partial \underline{f}} = -\frac{\lambda}{\pi} \leq 0, \quad (24)$$

$$\frac{\partial EW_{PS}}{\partial \bar{f}} = \frac{\lambda}{\pi} \geq 0, \quad (25)$$

while the complementary slackness conditions is

$$\lambda \left( \tilde{a} - 1 + \frac{\bar{f} - \underline{f}}{I\pi} \right) = 0.$$

From the second and the third FOC's it's immediate to see that  $\lambda = 0$ . Then,  $\underline{f}$  is determined by the undertake constraint, i.e.  $\underline{f} \leq (\pi\tilde{a})^2/2$ , since the optimal investment is  $I_{PS} = \pi\tilde{a}$ . This condition holds for sure when  $\underline{f} = 0$ . Since  $2Ew(\beta) - \pi \geq 0$  for  $\beta < \beta'_1 = \frac{2w^g - \pi}{2(w^g + w^b)}$ , we have  $\frac{\partial EW_{PS}}{\partial \tilde{a}} > 0$  in this interval and  $\hat{a} = 1$ . Then, a per-se legality rule applies. Conversely, for  $\beta > \beta'_1$  the first derivative is negative and the enforcer sets  $\hat{a} = 0$  and  $\bar{f} \geq \pi$ , adopting a per se illegality regime and discouraging the practice and the investment.

**Discriminating rule:** We already noticed that the investment  $I_D$  does not depend on the fines  $\underline{f}$  and  $\bar{f}$  nor on the errors  $\varepsilon^I$  and  $\varepsilon^{II}$ . Hence, we have  $\frac{\partial EW_D}{\partial \underline{f}} = \frac{\partial EW_D}{\partial \bar{f}} = 0$  and  $\frac{\partial EW_D}{\partial \varepsilon^i} = \gamma(\bar{\varepsilon} - \varepsilon^i) = 0$  for  $i = I, II$ . When  $\beta < \beta_0$  welfare is increasing in the action  $\hat{a}^b$ , that is  $\frac{\partial EW_D}{\partial \hat{a}^b} = [\Delta W_D - I_D] \beta \pi - \beta w^b I_D \geq 0$ , where  $\Delta W_D = (1 - \beta)w^g - \beta w^b \hat{a}^b$ , for the same argument developed in the case of research investment,  $\hat{a}^b = 1$  and  $I_D = \pi$ . When  $\beta > \beta_0$  the enforcer chooses an internal solution  $\hat{a}^b < 1$ . The fines are set to meet the incentive compatibility constraint. For instance, the pair  $\underline{f} = 0$  and  $\bar{f} = \frac{I\pi(1 - \hat{a}^b)}{(1 - \bar{\varepsilon})}$  satisfies the constraint (and the undertake constraint as well). Finally, there is no need to spend resources in costly accuracy since errors do not affect the investment and fines can be set to adjust the constraints.

The comparative statics in case of physical investment is much simpler than in the research case, since we can solve explicitly for the equilibrium action  $\hat{a}^b$ . Solving as an equality  $\frac{\partial EW_D}{\partial \hat{a}^b} = 0$  we get the equilibrium action

$$\hat{a}^b = \frac{(1 - \beta)(w^g - w^b - \pi)}{\beta(2w^b + \pi)}$$

that is lower than 1 for  $\beta > \beta_0$  and decreasing in  $\beta$ . Notice that, in case of research activity, the action  $\hat{a}^b$  evaluated at  $\bar{\varepsilon} = \varepsilon^I = 0$  gives the expression above. Moreover, being increasing in  $\bar{\varepsilon}$  and  $\varepsilon^I$  for  $\beta > \beta_0$  it follows that in case of research investment the enforcer implements a higher action in the bad state than in case of physical capital. Substituting in the optimal investment we obtain:

$$I_D(\beta) = \frac{\pi(1 - \beta)(w^g + w^b)}{2w^b + \pi}$$

which is lower than  $\pi$  for  $\beta > \beta_0$  and decreasing in  $\beta$ . Finally, substituting  $\hat{a}^b$  and  $I_D$  in  $EW_D(\beta) = I_D(\beta) [\Delta W_D(\beta) - I_D(\beta)/2]$  we have  $\partial EW_D / \partial \beta \simeq -2(1 - \beta)\pi(w^g + w^b)^2 / 2(2w^b + \pi) < 0$ .

**Optimal legal standard:** the per-se rule gives

$$\begin{aligned} & \pi \left[ Ew(\beta) - \frac{\pi}{2} \right] \text{ for } \beta \in [0, \beta'_1] \\ & 0 \text{ for } \beta \in (\beta'_1, 1]. \end{aligned}$$

The discriminating rule gives

$$\begin{aligned} & \pi \left[ Ew(\beta) - \frac{\pi}{2} \right] \text{ for } \beta \in [0, \beta_0], \\ & (1 - \beta)^2 \frac{\pi(w^g + w^b)^2}{2(2w^b + \pi)} \text{ for } \beta \in (\beta_0, 1] \end{aligned}$$

For  $\beta \in [0, \beta_0]$ , both  $D$  and  $PS$  are equivalent to the per-se legality regime. For  $\beta \in (\beta_0, \beta'_1]$ , although the enforcer may still implement  $\hat{a}^b = 1$ , it is optimal to set  $\hat{a}^b < 1$ . Hence,  $EW_D(\beta) > EW_{PS}(\beta)$ . Finally, for  $\beta \in (\beta'_1, 1)$ ,  $EW_D(\beta) > 0 = EW_{PS}(\beta)$  and  $EW_D(1) = 0 = EW_{PS}(1)$ . ■

## Bibliography

- Anderlini, L., Felli L., Immordino, G. and A. Riboni, (2013), “Legal institutions, innovation and growth,” forthcoming *International Economic Review*.
- Arrow, K.J., (1962), “Economic Welfare and the Allocation of Resources for invention,” in: Nelson, R.R. (Ed.), *The Rate and Direction of Inventive Activity: Economic and Social Factors*, Princeton University Press, Princeton, pp. 609-626.
- Dasgupta, P., and Stiglitz, J., (1980), “Uncertainty, Industrial Structure, and the Speed of R&D,” *The Bell Journal of Economics* 11, 1-28.
- Demougin D. and C. Fluet, (2008), “Rules of Proof, Courts and Incentives,” *Rand Journal of Economics*, 39, 20-40.
- Department of Justice (2008), “Competition and Monopoly: Single-Firm Conduct under Section 2 of the Sherman Act.”
- DG Competition (2005), “Discussion Paper on the Application of Article 82 of the Treaty to Exclusionary Abuses.”
- DG Competition (2008), “Draft Guidance on the Commission’s Enforcement Priorities in Applying Article 82 EC Treaty to Abusive Exclusionary Conducts by Dominant Undertakings.”

- Easterbrook F., (1984), “The Limits of Antitrust,” *Texas Law Review*, 63, 9-14.
- David E. and J. Padilla, (2005), “Designing Antitrust Rules for Assessing Unilateral Practices: a Neo.Chicago Approach,” *The University of Chicago Law Review*, 72, 73-98.
- Gual, J., Hellwig, M., Perrot, A., Polo, M., Rey, P., Schmidt, K. and R. Stenbacka, (2005), “An Economic Approach to Article 82,” Report for the DG Competition, European Commission, published on *Competition Policy International*, 2, 111-156.
- Immordino, G., Pagano, M. and M. Polo, (2011), “Incentives to Innovate and Social Harm: Laissez-faire, Authorization or Penalties?,” *Journal of Public Economics*, 95, 864-876.
- Immordino, G. and M. Polo, (2008), “Judicial Errors and Innovative Activity,” CSEF working paper n. 196.
- Kaplow, L., (1994), “The Value of Accuracy in Adjudication: an Economic Analysis,” *Journal of Legal Studies*, 15, 371-385.
- Kaplow, L., (1995), “A Model of the Optimal Complexity of Legal Rules,” *The Journal of Law, Economics and Organization*, 11, 150-163.
- Kaplow, L., (2011a), “Optimal Proof Burdens, Deterrence and the Chilling of Desirable Behavior,” *American Economic Review: Papers & Proceedings*, 101, 277–80.
- Kaplow, L., (2011b), “On the Optimal Burden of Proof,” *Journal of Political Economy*, 119, 1104-1140.
- Kaplow, L., (2012), “Burden of Proof,” *Yale Law Journal*, 121, 738–859.
- Kaplow, L. and S. Shavell, (1994), “Accuracy in the Determination of Liability,” *Journal of Law and Economics*, 37, 1-15.
- Kaplow, L. and S. Shavell, (1996), “Accuracy in the Assessment of Damages,” *Journal of Law and Economics*, 39, 191-210.
- Yannis, K. and D. Ulph, (2009), “On Optimal Legal Standards for Competition Policy: a General Welfare-Based Analysis,” *Journal of Industrial Economics*, LVII: 410-437.
- Lando, H. (2006), “Does wrongful conviction lower deterrence?” *Journal of Legal Studies*, 35, 327-337.
- Lear (2006), “The Cost of Inappropriate Intervention and Non Intervention Under Article 82,” Report for the Office of Fair Trade, London, UK.



- Lee, T., and Wilde, L.L., (1980), "Market Structure and Innovation: A Reformulation," *Quarterly Journal of Economics* 94, 429-36.
- Loury, G., (1979), "Market Structure and Innovation," *Quarterly Journal of Economics* 93, 395-410.
- Mookherjee, D. and I. Png, (1994), "Marginal Deterrence in the Enforcement of Law," *Journal of Political Economy* 102, 1039-66.
- Png I.P.L., (1986), "Optimal Subsidies and Damages in the Presence of Judicial Error," *International Review of Law and Economics*, 6, 101-105.
- Polinsky, M.A. and S. Shavell, (2000), "The Economic Theory of Public Enforcement of Law," *Journal of Economic Literature*, 38, 45-76.
- Polo M., (2010), "Anticompetitive vs Competitive Explanations of Unilateral Practices: the Identification Problem," *Journal of Competition Law and Economics*, 6, 457-476.
- Posner, R.A., (1992), "The Economic Analysis of Law," 4th Edition. Boston: Little Brown.
- Schwartzstein, J., Shleifer, A., (2012) "An Activity-Generating Theory of Regulation," forthcoming, *Journal of Law and Economics*.
- Segal I., and M. Whinston, (2007), "Antitrust in Innovative Industries," *American Economic Review*, 97, 1703-1730.
- Shavell, S., (1980), "Strict Liability versus Negligence," *The Journal of Legal Studies*, 9, 1-25.
- Shavell, S., (1992), "Liability and the Incentive to Obtain Information about Risk," *Journal of Legal Studies*, 21, 259-270.
- Shavell, S., (2007), "Do Excessive Legal Standards Discourage Desirable Activity?" *Economics Letters*, 95, 394-397.
- Stigler, G.J., (1970), "The Optimum Enforcement of Laws," *Journal of Political Economy*, 78, 526-536.