International Review of Law and Economics 40 (2014) 36–50

Contents lists available at ScienceDirect

International Review of Law and Economics

Antitrust, legal standards and investment

Giovanni Immordino[a,*], Michele Polo[b]

[a] Università di Salerno and CSEF, 84084 Fisciano, SA, Italy
[b] Università Bocconi, IEFE and IGIER, Via Sarfatti 25, 20136 Milan, Italy

ARTICLE INFO

Article history:
Received 22 March 2014
Received in revised form 2 August 2014
Accepted 19 August 2014
Available online 28 August 2014

JEL classification:
D73
K21
K42
L51

Keywords:
Legal standards
Accuracy
Antitrust
Innovative activity
Enforcement

ABSTRACT

We study the interaction of a firm that invests in research and, if successful, undertakes a business 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.

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. Although antitrust intervention typically focuses on incumbent's practices and does not consider its research investment decisions, these latter have a strong impact on the evolution of high-tech industries and social welfare.[1] When business strategies are a tool to further extract profits from the innovation, the incumbent's research efforts depend jointly on the degree of patent protection and on the antitrust treatment of the practices that the innovator might undertake upon discovery.

This paper studies the optimal antitrust intervention – both in terms of legal standards and enforcement tools – for given intellectual property rights protection to condition the adoption of certain business practices 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 with 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 its servers' and clients' operating systems with the competitors' server 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. In parallel with those complex cases, the debate in competition policy has also raised questions about the impact of antitrust enforcement on the innovative activity characterizing these industries. For instance the commitments imposed in the EC v. Microsoft decisions to disclose

* Corresponding author.
E-mail addresses: giimmo@tin.it (G. Immordino), michele.polo@unibocconi.it (M. Polo).

[1] In this paper we do not consider antitrust issues related to joint research programs or licensing, that typically involve the threat of collusion, but rather analyze the case of individual firms research programs and business strategies that may lead to foreclosure. On the first topic, see Chang (1995), Green and Scotchmer (1995) and Erkal (2004).

http://dx.doi.org/10.1016/j.irle.2014.08.003
0144-8188/

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 adverse effects on the incentives to innovate.

The impact of antitrust policies on investment has recently emerged as an important theme also in network industries where the investment is primarily in physical capital. The early stages of public utilities liberalization in Europe have focused on granting the competitors non-discriminatory access to the existing infrastructures. In recent years, the need of huge investments in new electricity, gas or telecom networks has urged the policy debate to combine the promotion of competition and the incentives to invest.[2]

To sum up, several landmark cases have raised the issue of the effects of antitrust enforcement on the incentives to invest. The debate on competition policy has further examined the different components of antitrust intervention, that require to choose appropriate legal standards and 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 enforcement policy to regulate certain business practices of a dominant firm which invests in research or in physical capital. We include in the model a positive effect of a new technology on profits and welfare when the practice is not adopted, which derives directly from the innovation. In this framework, the baseline profits are 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.

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 not only the degree of patent protection but also the stricter or laxer enforcement of 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 present at the time the firm undertakes it. Something that is inherently uncertain at the time the investment is sunk.

A key feature of our approach is that the effects of the practices, when applied to the new technologies, are unknown at the time the investment is sunk and the policy is set. This may be due to the interaction of the innovation, whose properties may have been controlled and planned by the firm with sufficient confidence, with the economic or social environment at the time the innovation will be introduced. The features of this environment at this later stage, in turn, will depend on the decisions of other agents and cannot be assessed *ex ante* with certainty. To illustrate, consider the example of a dominant software company that may invest in research to tie a new software application into a new personal computer operating system. Beyond the initial intent of the company, the efficiency and foreclosure effects of tying this new software packages will depend, at the time of its commercial introduction, on the alternative packages and applications available from competitors, which may be only imperfectly foreseen at the time of the research investment.

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 intervention.

We consider two aspects of the antitrust intervention. The first is the selection of the optimal legal standard,[3] that establishes under which conditions a practice is unlawful and therefore specifies when a firm can be convicted, together with the evidence needed to prove it guilty. The second concerns the enforcement policy, that is the sanctioning rule and the accuracy in collecting evidence. We show that the optimal legal standard and enforcement policies depend on the expected social effects of a certain business practice, what we can call the "economic model" of the enforcer, or, in the words of Judge Frank Easterbrook, her presumptions (see Easterbrook, 1984). Easterbrook, for instance, quotes Donald Turner on the inhospitality tradition in antitrust: "the tradition is that judges view each business practice with suspicion, always wondering how firms are using it to harm consumers." This specific a priori corresponds, in our model, to a configuration of parameters that assigns a high probability to a negative and large impact of the practice on social welfare. Given these presumptions, then, we predict the kind of legal standard and enforcement policy that the judge will apply (for instance, a per-se illegality rule).

Our main results are the following. First, we fully characterize the optimal policies under per-se or discriminating rules for any expectation of the social effects of the practices. A general feature of the optimal enforcement policy – for any legal standard – refers to balancing the need to deter the practice when unlawful, a concern that is relevant *ex post*, and the attempt to sustain the investment by adopting a more lenient policy when, *ex ante*, the practice is expected to improve welfare. For instance, under a per-se rule, one may think that the practice should be allowed when welfare enhancing or completely discouraged when socially harmful. We show that the optimal enforcement policy is much richer than that. It may be optimal to allow the firm adopting the practice, giving up *ex post* deterrence, still fining the practice (that is per-se illegal) to reduce the incentives to invest; or, for more pessimistic expectations, it is optimal to implement the practice at an intermediate level, to balance the two dimensions of deterrence.

Secondly, we identify the optimal legal standards, that 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 rules out the possibility of sanctioning the firm that undertakes the practice in the (unlikely) event that the practice is socially harmful in order to boost 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. Third, the design of the optimal antitrust intervention adapts to the degree of patent protection, choosing a laxer (stricter) approach when intellectual property rights are weakly (strongly) protected. Hence, competition and patent policies act as substitutes. 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 per-se illegality as the optimal legal standard when the new technology is very likely to be socially harmful. Fifth, we show that, although the two cases are not equivalent, our results on the optimal legal standard

[2] See, for instance, EC (2013) on the recent European debate on the different access price regimes to be adopted in the legacy copper telecom network and in the new optical fiber network to be constructed.

[3] The debate on the appropriate legal standards for foreclosure or monopolization practices has developed in recent years in Europe and the US. See DG Competition (2005), DG Competition (2008), Gual et al. (2005) and Department of Justice (2008).

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 intervention when investment matters. Finally, 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, if the enforcer would be allowed to switch to a different legal standard once the investment is in place, she would move to a discriminating rule. Thus, there is a time-consistency issue that may require to use commitment tools (regulations, guidelines, precedents). For the same argument, industries where the investment issue is not relevant should opt for a generalized discriminating rule and a more interventionist approach.

Contributions to the literature. We contribute to the literature on antitrust and regulatory intervention in investment-intensive industries. Immordino et al. (2011, hereafter IPP) propose an analytical framework similar to this paper. They focus on the choice between *ex post* liability and *ex ante* authorization of innovative products as genetically modified (GM) organisms, or new drugs. They identify when each policy is optimal. In this paper, instead, we compare per-se and discriminating legal standards, within the *ex post* law enforcement regime, referring to antitrust intervention against abuse of dominance. Hence, the two papers can be read as complementary.

Another model that comes close to ours is that of Schwartzstein and Shleifer (2012, hereafter SS), where firms may take precautions but face uncertainty due to possible judicial errors. Similarly to us both IPP and SS find that regulation should be softer when social harm is unlikely. 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.

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 a 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 and the interaction between antitrust and patent policy, that are key in our paper, are not addressed.

Chang (1995), Green and Scotchmer (1995) and Erkal (2004) consider the interaction between antitrust and patent policy in a setting of sequential innovations. They show that different licensing agreements may have an impact on the incentives to further discover the second innovation as well as on market coordination. Then, the focus in these papers is on the antitrust treatment of licensing and collusive agreements rather than on foreclosure strategies by an innovative incumbent, as in our paper.

Finally, 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.[4]

The remainder of this paper is organized as follows. Section 2 presents the model. Section 3 focuses on antitrust intervention in innovative industries and its interaction with patent policy. Section 4 studies the effect on legal standards of a cap on fines, Section 5 analyzes the case of deterministic investment in physical capital and Section 6 deals with the case of sunk investment. All proofs not following immediately from the main text are in 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 and discovers with a certain probability a new technology affecting both social welfare and profits. The innovator, thanks to its superior technology becomes dominant; moreover, using the new technology he undertakes business strategies that allow to further extract profits from the innovation. Hence, the profits in case of successful innovation come from two sources: the new discovery itself, that creates a competitive advantage that the firm can partially appropriate according to the degree of patent protection; and the business strategy, that is subject to antitrust scrutiny (due to the dominant position acquired by the innovator). The larger the initial investment in research, 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 patent policy is effective in protecting the innovator, and 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 describe the investment and the business practice undertaken by the firm; then we introduce the legal standard adopted by the 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 might invest in 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[5]: 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 $c(I)=I^2/2$.

The innovation has a positive effect on welfare (W) and on the profits of the innovator (Π). In this simple setting, we can interpret the level of profits Π appropriated by the firm, for given welfare W, as a measure of the effectiveness of patent protection. A higher Π, indeed, can be interpreted as a more effective protection of the

[4] 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), Kaplow and Shavell (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).

[5] 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.

patent holder from the imitation of competitors, while weak intellectual property rights correspond to the case when most of the welfare W generated by the innovation is distributed to consumers and/or competitors, with a low Π for the innovating firm.

The firm can further 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 $a \in [0, 1]$, where the lower bound $a = 0$ can be interpreted as not undertaking the practice at all.

Private and social effects. When the dominant firm undertakes the practice, this latter affects profits and welfare, adding to the fixed effects described above, according to the intensity measured by the action a undertaken. More precisely, the overall **profits** when the innovation is successful and the practice is adopted are $\Pi(a) = \Pi + \pi a$ with $\Pi \geqslant 0$ and $\pi > 0$.

While the private effects of the practice are always positive, its **social impact** may be positive or negative. More precisely, while the pure effect of the innovation is always positive ($W > 0$), 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 case when the practice is not undertaken.

We model the effects of the practice as described by two states of the world. In the bad state $s = b$, when the firm exploits the new technology through the business practice a, social welfare is reduced according to the expression $W^b(a) = W - w^b a$ where $w^b > 0$. In the bad state, private incentives conflict with social welfare. 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: $W^g(a) = W + w^g a$ 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 and avoiding foreclosure, while the integrated package released by the firm allows a more user-friendly usage of the software.

Information. We assume that both the firm and the enforcer know the private and social effects of the innovation (W and Π) but do not observe the social effects of the practice at the time the policy is set and the investment is sunk. At this stage they both assign a probability β 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 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 a chosen by the firm. 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 σ on the state of the world, that is whether the incumbent's practice is welfare enhancing or decreasing. We interpret the signal as the evidence collected on the effects of the practice, i.e. the state of the world. 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.[6] With probability ε^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 ε^{II} the signal is incorrect when the true state is the bad one: in this case the enforcer fails to identify the practice as socially damaging, committing a type-II error. Hence,

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

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

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 $\{W, \Pi, w^g, w^b, \pi, \beta\}$. 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:

$$W > \Pi \tag{1}$$

$$w^g > \pi \tag{2}$$

$$W + w^g < 1 \tag{3}$$

that ensures that the innovation as well as the practice in the good state generates an increase in consumer surplus and that the optimal investment, corresponding to the probability to innovate, is an internal solution.

Antitrust policies: legal standards, fines and accuracy. Antitrust intervention focuses on the practices undertaken by (dominant) firms. Specifically, the enforcer designs competition policy 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 defines under which circumstances (if any) the practice is considered unlawful. Hence, the legal standard identifies the conditions that give the enforcer the right to sanction the firm, and the kind of evidence that must be produced to prove it guilty. 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[7]:

L per-se legality: any action $a \in (0, 1]$ is always legal no matter which signal the enforcer receives;
IL per-se illegality: any action $a \in (0, 1]$ is always illegal no matter which signal the enforcer receives.

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

[6] On the burden of proof see, for instance, Kaplow (2011a), Kaplow (2011b), Kaplow (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.

[7] 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 in some cases it may be optimal to have the firm undertake the practice at some degree (and pay a positive fine).

D discriminating: any action $a \in (0, 1]$ is legal unless the enforcer receives a signal $\sigma > x$.

Notice that in the per-se illegality regime the enforcer has to prove that the practice has been undertaken, no matter what the effects are. Conversely, in the discriminating regime the enforcer cannot convicts the firm only on the basis of the action undertaken, if she is unable to prove that the effects are socially harmful. 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 intervention 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 in the good and/or the bad state and reducing the probability of the error accordingly. 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 $\bar{\varepsilon}$.[8] More precisely, the cost of implementing an error probability ε^i is $g(\varepsilon^i) = (\gamma/2)(\bar{\varepsilon} - \varepsilon^i)^2$.

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.[9] Moreover, the right to convict the firm according to the fine schedule depends on the ability of the enforcer to prove that the condition established in the legal standard in place are met. For instance, if a discriminating rule applies, the enforcer has to produce evidence of social harm (the negative signal $\sigma > x$), that entitles her to set a fine, that may possibly vary with the level of the practice (action a). Since the profit function 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

$$f(a) = \begin{cases} 0 & \text{if } a = 0 \\ \underline{f} \geqslant f & \text{if } 0 < a \leqslant \widetilde{a} \\ \bar{f} \leqslant F & \text{if } a > \widetilde{a}. \end{cases} \tag{4}$$

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.

Timing. The timing of the game 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 ε^I and ε^{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 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 σ at the effects that is incorrect with probability ε^I in the good state and ε^{II} in the bad state and levies a fine (if any) consistently with the legal standard and enforcement policy adopted.

[8] In this case the decision is based on a small set of evidence easy and inexpensive to collect.

[9] Besides 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.

3. Optimal legal standards and enforcement policies

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

3.1. First best

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 $EW_{FB}(\beta, I) = I(W + (1 - \beta)w^g) - (I^2/2)$, that yields the optimal investment level

$$I_{FB} = W + (1 - \beta)w^g, \tag{5}$$

that is lower than 1 given our assumptions, increasing in the welfare gain generated by the innovation (W) and by the practice ($(1 - \beta)w^g$).

In what follows, for given effectiveness of the patent policy, summarized in the terms W and Π, the antitrust enforcer 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 by fully characterizing the optimal enforcement policies in the per-se and discriminating regimes given the economic model ($W, \Pi, w^g, w^b, \pi, \beta$) that the enforcer adopts, an issue of independent interest. Then, we move to the analysis of the optimal legal standards, evaluating each regime at its optimal enforcement policy. This way, we are able to identify how the economic model of the enforcer determines the selection of the optimal legal standard and enforcement policy.

3.2. 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 σ received). We analyze the optimal antitrust 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. Specifically, if the research investment has been successful the firm at time 3 chooses the action that maximizes profits $\Pi + \pi a - f(a)$. Given the fine schedule (4), the incentive compatibility constraint can be written as

$$\pi\widetilde{a} - \underline{f} \geqslant \pi - \bar{f}.$$

The undertake constraint, instead, ensures that the firm (weakly) prefers to adopt the practice ($a > 0$) rather than keeping on with "business as usual" ($a = 0$).[10] Let us define $\widetilde{A}$ as the set of actions that satisfy the incentive compatibility and undertake constraint given π, f and F. The enforcer, then, by properly defining the fine schedule, can implement any action $\widetilde{a} \in \widetilde{A}$.

The expected profits at time 2 under per-se rules (subscript PS) are $E\Pi_{PS} = I(\Pi + \pi\widetilde{a} - \underline{f}) - (I^2/2)$ and the profit maximizing investment is

$$I_{PS} = \Pi + \pi\widetilde{a} - \underline{f} \tag{6}$$

where $\widetilde{a} \in \widetilde{A}$.

This latter expression shows that, although the fine is not conditioned on the level of investment, it affects the firm's research effort I through the level of the *ex post* net profits.

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

$$EW_{PS}(\beta) = I_{PS}(W + Ew(\beta)\widetilde{a}) - \frac{(I_{PS})^2}{2}, \tag{7}$$

where

$$Ew(\beta) = (1 - \beta)w^g - \beta w^b$$

is the expected marginal welfare of the practice. The enforcer, given the investment I_{PS}, selects through the fine schedule, among the actions $\widetilde{a} \in \widetilde{A}$, the one that maximizes welfare – which we denote $\widehat{a}$ – that is the action that the firm is willing to choose according to the incentive compatibility and undertake constraints (being in $\widetilde{A}$), and that is socially optimal.

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)[11]; and *ex ante* deterrence on investment: *ex post* profits, indeed, depend on the action implemented and on the fines levied, affecting this way the *ex ante* investment.

In the following lemma we derive the optimal enforcement policy under per-se rules. It is 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 $\widehat{a} = 1$ and $\underline{f} = 0$, it is optimal not to fine the practice at any degree a. Then, the corresponding legal standard is per-se legality. If, instead, $\widehat{a} < 1$, the practice is 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:

$$\beta_2 \equiv \frac{w^g}{w^g + w^b} > 0$$

$$\beta_1 \equiv \beta_2 + \frac{W - \Pi - \pi}{w^g + w^b} > 0$$

$$\beta_1' \equiv \beta_2 + \frac{(W - \Pi - \pi)\pi}{(w^g + w^b)(\Pi + 2\pi)}$$

$$\beta_3 \equiv \beta_2 + \frac{(W - \Pi)\pi}{\Pi(w^g + w^b)} > 0.$$

Notice that β_3 may be larger or smaller than 1, and that $\beta_2 > \beta_1' > \beta_1$ if $W - \Pi - \pi < 0$, while $\beta_2 < \beta_1' < \beta_1$ if $W - \Pi - \pi > 0$. The following lemma completely characterizes the optimal enforcement policies and per-se rules.

Lemma 1 (Optimal enforcement policy under per-se rules). *Assume that the minimum fine is zero and the maximum fine is sufficiently high, i.e.* $\underline{f} = 0$, $F > \pi$. *Further define case (a) if* $W - \Pi - \pi < 0$ *and case (b) if* $W - \Pi - \pi \geqslant 0$. *The optimal legal standard and enforcement policy under per-se rules are:*

1 *For* $\beta \in [0, \beta_1]$ *in case (a) and for* $\beta \in [0, \beta_1']$ *in case (b), the optimal legal standard is per-se legality and the optimal enforcement implements* $a^g = a^b = 1$ *and* $I_{PS} = \pi$, *by setting* $\widehat{a} = 1, \underline{f} = 0$.
2 *For* $\beta \in (\beta_1, \beta_2)$ *in case (a), the optimal legal standard is per-se illegality and the optimal enforcement implements* $a^g = a^b = 1$ *and* $I_{PS} = W + Ew(\beta)$, *decreasing in* β, *by setting* $\widehat{a} = 1$ *and* $\underline{f} = [-W - Ew(\beta) + \Pi + \pi]$.
3 *For* $\beta \in [\beta_2, \min\{\beta_3, 1\}]$ *in case (a) and for* $\beta \in [\beta_1', \min\{\beta_3, 1\}]$ *in case (b), the optimal legal standard is per-se illegality and the optimal enforcement implements* $a^g = a^b = \widehat{a}$ *and* $I_{PS} = \frac{\pi W - Ew(\beta)\Pi}{\pi - 2Ew(\beta)}$ *by setting* $\widehat{a} = \frac{(W-\Pi)\pi + Ew(\beta)\Pi}{(\pi - 2Ew(\beta))\pi}$ *decreasing in* β, $\underline{f} = 0$ *and* $\bar{f} \geqslant \widehat{a}(1 - \pi)$.
4 *If* $\beta_3 < 1$, *for* $\beta \in [\beta_3, 1]$ *in both case (a) and (b), the optimal legal standard is still per-se illegality and the optimal enforcement implements* $a^g = a^b = 0$ *and* $I_{PS} = \Pi$, *by setting* $\widehat{a} = 0, \underline{f} = 0$ *and any* $\bar{f} \geqslant \pi$.

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 $\widehat{a}$ 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 ($\beta < \beta_1$ in case (a) and $\beta < \beta_1'$ in case (b)). Notice that a less effective patent protection, corresponding to a lower Π for given W, shifts both thresholds to the right expanding the region where per-se legality is selected.

This regime, in turn, is replaced by per-se illegality when β, the probability of social harm increases. One may wonder that, in this case, since the practice is illegal, the enforcer would try to deter it completely ($\widehat{a} = 0$). This is the case only when the probability of social harm is very high ($\beta > \beta_3$). In the other cases, the enforcement policy is richer, implementing a reduction in the investment through fines and/or through the level of the practice. In case (a), since the investment is influenced by the fine $\underline{f}$, when $\beta \in (\beta_1, \beta_2)$ the enforcer adopts a per-se illegality regime, but focuses enforcement on progressively reducing the investment (and the probability of undertaking the practice) by raising the fine $\underline{f}$, rather than discouraging directly the practice, that is still implemented at the highest level ($\widehat{a} = 1$). In other words, in this region the enforcer intervenes through *ex ante* rather than *ex post* deterrence. Conversely, for $\beta \geqslant \beta_2$ in case (a) and for $\beta > \beta_1'$ in case (b), the enforcer intervenes by progressively reducing the practice (and the expected profits), thereby indirectly decreasing the investment. In this region, therefore, the policy works through *ex post* deterrence ($\widehat{a} < 1$), affecting indirectly also the investment. Finally, it is worth noting that the expected welfare is continuous and decreasing in β.

3.3. Discriminating rules

In a discriminating regime, the enforcer has the right to sanction the firm if she collects evidence of a socially harmful effect of the practice (the informative although noisy signal $\sigma > x$). In this case, the fine schedule is applied, possibly levying different fines depending on the level of the action. This feature of the discriminating regime has two implications. First, the same action can be treated differently according to the signal received. Therefore, the enforcer, in contrast with per-se rules, can implement different levels of the practice in different states of the world. Secondly, the evidence on

[10] The constraint is relevant as long as $\widetilde{a} > 0$ and implies $\pi\widetilde{a} - \underline{f} \geqslant 0$.

[11] For the standard marginal deterrence problem in law enforcement see for instance Stigler (1970) and Mookherjee and Png (1994).

the effects (the signal σ) is the crucial element to prove the firm guilty, and cannot be replaced by any inference based on the level of the practice itself. We shall come back to these features when commenting the optimal enforcement policy.

Since the discriminating legal standard does not allow the enforcer to levy any fine if the signal is $\sigma \leqslant x$, 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 ε^{I} when instead it is welfare enhancing.

In the bad state, given the fine schedule $f(a)$, the incentive compatibility and undertake constraints give the following inequalities:

$$\Pi + \pi\widetilde{a}^b - (1-\varepsilon^{II})\underline{f} \geqslant \max\{\Pi + \pi - (1-\varepsilon^{II})\bar{f}, 0\}.$$

Although the incentive compatibility constraint to implement $\widetilde{a}^b$ 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 $a^g = \widetilde{a}^b$ rather than $a^g = 1$.[12] Hence, we have to further impose the incentive compatibility and undertake constraints for the good state:

$$\Pi + \pi - \varepsilon^{I}\bar{f} \geqslant \max\{\Pi + \pi\widetilde{a}^b - \varepsilon^{I}\underline{f}, 0\}.$$

These constraints define the interval in which the fines must be chosen in order to implement $a^b = \widetilde{a}^b$ and $a^g = 1$, i.e.,

$$\bar{f} \in \left[\underline{f} + \frac{\pi(1-\widetilde{a}^b)}{1-\varepsilon^{II}}, \underline{f} + \frac{\pi(1-\widetilde{a}^b)}{\varepsilon^{I}}\right]. \tag{8}$$

Let us define with $\widetilde{A}^b$ the set of implementable actions when the practice is welfare detrimental.

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

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

The innovative investment in the discriminating regime is therefore

$$I_D = \Pi + (1-\beta)[\pi - \varepsilon^{I}\bar{f}] + \beta[\pi\widetilde{a}^b - (1-\varepsilon^{II})\underline{f}] \geqslant 0. \tag{9}$$

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.

Finally, the expected welfare under the discriminating rule is

$$EW_D = I_D\left[W + \Delta W_D - \frac{I_D}{2}\right] - \frac{\gamma}{2}(\bar{\varepsilon} - \varepsilon^{I})^2 - \frac{\gamma}{2}(\bar{\varepsilon} - \varepsilon^{II})^2, \tag{10}$$

where $\Delta W_D = (1-\beta)w^g - \beta w^b\widetilde{a}^b$. The optimal policy requires therefore to set the fine schedule $(\underline{f}, \bar{f}, \widetilde{a}^b)$ and the errors ε^{I} and ε^{II} to maximize the expected welfare under the above constraints. As before, the enforcer will select through the fine schedule and the level of accuracy the action, among those that are implementable ($\widetilde{A}^b$), that maximizes welfare: we denote with $\widehat{a}^b$ the action that solves this program (in the bad state). Finally, let us define the following threshold

$$\beta_0 \equiv \beta_2 + \frac{W - \Pi - \pi - w^b((\Pi+\pi)/\pi)}{w^g + w^b}.$$

Notice that β_0 may be larger or lower than 0 according to our assumptions. In the following lemma we identify the optimal enforcement policy in the two cases.

Lemma 2. *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:*

1 *If* $\beta_0 > 0$, *for* $\beta \in [0, \beta_0)$, *the optimal policy implements* $a^g = a^b = 1$ *and* $I_D = \Pi + \pi$ *by setting* $\widehat{a}^b = 1, \underline{f} = 0$ *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.*
2 *For* $\beta \in [\max\{\beta_0, 0\}, 1]$ *if* γ *is sufficiently large the optimal policy implements the actions* $a^b = \widehat{a}^b < 1$, $a^g = 1$ *and investment* $I_D < \Pi + \pi$ *by improving type-I accuracy* ($\varepsilon^{I} < \bar{\varepsilon}$, $\varepsilon^{II} = \bar{\varepsilon}$) *and by setting* $\widehat{a}^b < 1, \underline{f} = 0$, *and* $\bar{f} = \frac{\pi(1-\widehat{a}^b)}{(1-\bar{\varepsilon})}$.

The optimal enforcement policy under the discriminating rule is shaped by the interaction of *ex post* (marginal) deterrence, focused 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 $\widehat{a}^b \in (0, 1)$ in the bad state. While *ex post* deterrence always requires to lower $\widehat{a}^b$, *ex ante* deterrence prescribes a high $\widehat{a}^b$ to increase expected profits and investment whenever the expected welfare increases with the practice.[13]

When social harm is unlikely ($\beta < \beta_0$), *ex ante* deterrence prevails and calls for a lax enforcement, implementing $\widehat{a}^b = 1$, an outcome equivalent to a per-se legality rule.[14, 15] Above this threshold, the enforcer implements $\widehat{a}^b < 1$ by properly setting the fine schedule and errors according to the incentive compatibility constraints. By lowering $\widehat{a}^b$, the enforcer reduces the negative impact of the practice on welfare, counterbalancing the higher probability of social harm, and at the same time 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 $\bar{f}$ since the incentive compatibility constraint requires a sufficiently high fine to induce the firm to choose $\widehat{a}^b < 1$ instead of 1 in the bad state. Then, ε^{I}, that acts as a substitute to the fine in affecting the investment, is reduced.

Finally, the threshold β_0 that delimits the per-se legality regime shifts to the right, implying a laxer antitrust enforcement, when the profits Π appropriated by the innovator are lower, due to a less effective patent protection.

We are further interested in characterizing the level of action $\widehat{a}^b < 1$ implemented when the likelihood of social harm β increases. Sufficient conditions for a monotonically decreasing action $\widehat{a}^b$ require some more structure, as established in the following lemma.

Lemma 3. *If* $w^g - w^b - \pi > 0$, *and* γ *sufficiently large,* $\widehat{a}^b$ *is decreasing in* β *with* $\widehat{a}^b \to 1$ *for* $\beta \to \max\{\beta_0, 0\}$ *and* $\widehat{a}^b \to \max\left\{\frac{W - \Pi - w^b\frac{\Pi}{\pi}}{2w^b + \pi}, 0\right\}$ *for* $\beta \to 1$.

[12] This is what Kaplow (2011a) defines as the chilling effect of fines on desirable actions.

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

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

[15] 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 β the enforcer would still implement the action at the highest level, $\widehat{a}^b = 1$ 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.

The enforcement policy under a discriminating rule is based on the noisy signal, that leads to type-I and type-II errors, and induces the firm to select a different level of the practice depending on the state of the world, $\widehat{a}^b < \widehat{a}^g = 1$. One may wonder, then, why the enforcer does not infer the true state of the world from the level of the practice, taking into account that the firm has a better (indeed, perfect) information on the effects. In other words, if the enforcer observes the action taken and the firm takes different actions in different states, why does not the enforcer infer, in equilibrium, the actual state of nature from that firm's choice?

A first argument simply says that a discriminating legal standard formally prevents the enforcer from basing the decision on the practice rather than on the effects, since the firm can be convicted only upon evidence of the latter. Notice that the very nature of the discriminating rule does not allow to convict based on the level of the practice disregarding, or contradicting, the evidence on the effects (the signal). Specifically, if $\sigma > x$ the relevant evidence at disposal proves the practice to be welfare detrimental even though the action associated to the good state ($a = 1$) were observed. And, conversely, if $\sigma \leqslant x$ the enforcer has collected evidence that the practice is welfare enhancing, and cannot overturn this evidence based on the fact that the firm is adopting $a < 1$.

There is however a deeper argument. In a discriminating regime, the firm optimally chooses different levels of the practice in the different states of the world. However, this is true exactly because a discriminating rule ensures that the firm will not be convicted according to the level of the practice, but only based on the effects.

Suppose the following modified discriminating rule (*MD*) applies: the enforcer convicts the firm when the practice adopted is consistent with the bad state, and she disregards the (imperfect) signal when it is in contradiction with the practice adopted. As a result this regime prevents any conviction when the practice is consistent with the good state, no matter what the signal is. Then, since profits are increasing in the practice a and the firm is never fined when it chooses the maximal action $a = 1$, this latter is the dominant strategy in both states of the world. We conclude that the *MD* regime is unable to implement a (separating) equilibrium where the firm adopts different levels of the practice in the different states.

This discussion, therefore, highlights a further feature of the discriminating rule: by committing not to base the conviction on the *ex post* observed action, it gives the proper incentives to implement a (separating) equilibrium, characterized by a lower level of the practice when it is welfare detrimental.

3.4. Optimal legal standards

We are now equipped to find the optimal regime, by selecting the legal standard, evaluated at the corresponding optimal enforcement policies, that gives the highest expected welfare.

Proposition 1 (Optimal legal standards). *If $\beta_0 > 0$, the optimal legal standard is a per-se legality rule for $\beta \leqslant \beta_0$ and the discriminating rule for $\beta \in (\beta_0, 1]$.*

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 β 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), since it prevents any intervention *ex post* on the practice when socially harmful, boosting the research investment at most.[16,17] 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, limiting the possibility 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.[18]

Our result suggests that a legal standard, specifying under which conditions a firm can be fined, restricts the degrees of freedom of the enforcement policy. Indeed, the fine schedule – a key instrument of the enforcement policy – can be applied only in the circumstances established by the legal standard in place. If, for instance, under a discriminating rule, the enforcer can levy fines according to the adopted schedule only if she is able to provide evidence that the practice is socially harmful. Hence, in our setting legal standards impose two types of constraints: the first is a standard commitment, since legal standard and enforcement policies are chosen before the firm undertakes any decision (see Section 6 for the case of a policy that is chosen after the investment is sunk); the second, as clarified above, is a restriction in the enforcement policy delimiting the cases when fines can be levied.

The results on the optimal antitrust policy and legal standards help highlighting also the relationship between competition policy, that applies to the practice adopted, and intellectual property rights protection. As already discussed, we can consider different degrees of patent (or intellectual property rights) protection by considering how the optimal antitrust intervention varies when, for given social effect W of the innovation, the ability of the innovator to privately appropriate these benefits (Π) varies.

When Π decreases, due to ineffective patent protection, the thresholds β_0 and β_3 shift to the right. Hence, the region where per-se legality (for low β) is selected expands (or starts applying). At the same time, in the region where the discriminating rule is adopted, the action $\widehat{a}^b$ implemented in the bad state increases, and may remain positive even when $\beta = 1$. Both adjustments weaken the antitrust enforcement and increase the expected profits that the innovator can obtain from the practice, boosting the investment. Conversely, an increase in Π, for given W, corresponding to a more effective patent protection, moves the thresholds β_0 and β_3 to the left, reducing or excluding the case for per-se legality and reducing the implemented action $\widehat{a}^b$ under the discriminating rule. Indeed, for β sufficiently high, in the bad state the practice is completely deterred, implementing the first best course of actions ($\widehat{a}^g = 1, \widehat{a}^b = 0$): the incentives provided by the patent protection (Π) are sufficiently high, with no need to relax the antitrust intervention. Hence, when innovative investment plays an important

[16] 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 $\underline{f} > 0$. In this case, the discriminating rule charges f when implementing the action $\widehat{a}^b$ and does not succeed to replicate the per-se legality regime.

[17] This result is reminiscent of Aghion and Tirole (1997): they show that by committing not to intervene ex-post in the selection of a project a principal can give the agent more incentives to take initiatives in finding an efficient project. When the principal's and agent's objectives are sufficiently congruent, delegation is the optimal solution. See on this issue also Cremer (1995).

[18] The role of commitment and flexibility of a legal system in affecting growth has been recently studied by Anderlini et al. (2013).

role in an industry, patent policy and competition policy act as substitutes: a weaker protection of intellectual property rights suggests to adopt a laxer competition policy towards the practices that the innovator can undertake in the market to exploit its technological excellence.

The previous discussion implicitly assumed that different institutions rule antitrust and patent issues, acting each one without coordinating with the other. In this sense, we commented on how antitrust enforcement should adapt to a more or less effective intellectual property rights protection guaranteed by the patent office. However, we can also imagine situations in which the same institution, e.g. a judge, adjudicates on the degree of patent protection and of antitrust infringment. In this case, since the expected welfare is increasing in the profits of the innovator, i.e., $\partial EW_D/\partial \Pi = W + \Delta\widetilde{W}_D - I_D > 0$, the judge would grant maximum patent protection ($\Pi = W$), reducing the lenient bias that otherwise antitrust enforcement would adopt. In a more structured analysis, finally, we may think that total welfare W decreases in the profits of the innovator Π: this can be the case when there is a deadweight loss associated to larger private profits of the firm, or in case the enforcer assigns a larger weight to consumers' surplus than to private profits. In this case we may expect that the judge would select an internal solution on the degree of patent protection. Our conclusion, however, confirms that IPR protection and antitrust act as substitutes in our environment, whereas the specific solution depends on the institutional setting that governs patent protection and antitrust intervention.

This result is reported in the following proposition:

Proposition 2 (Antitrust versus patent policy). *If IPR protection and antitrust enforcement are ruled by different institutions that act with no coordination, a more effective patent policy (a higher Π) reduces the region where per-se legality is applied, and implements a stricter discriminating rule, and vice-versa. If, instead, the same institution, e.g. a judge, adjudicates both antitrust infringements and IPR protection, a high level of patent protection (high Π) will be granted, reducing the leniency bias of antitrust intervention. In both environments, patent and competition policies act as substitutes.*

4. 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 Proposition 1 and Lemma 3, the optimal enforcement for $\beta > \beta_0$ is a discriminating rule that progressively reduces the socially harmful practice $\widehat{a}^b$ and increases the fine $\bar{f} = (\pi(1-\widehat{a}^b))/((1-\bar{\varepsilon}))$ as β increases. At the same time, type-I accuracy is 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, $\widehat{a}^b$ 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 β 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 $\beta_4 > \beta_0$ such that $\bar{f} = \pi$ and the limited liability constraint starts binding. For $\beta > \beta_4$, $\widehat{a}^b$ becomes a function of the type-II error ε^{II}, as can be seen setting $\underline{f} = 0$ in the lower bound of (8) to get[19]

$$\widehat{a}^b = \varepsilon^{II}. \tag{11}$$

By reducing ε^{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 $\widehat{a}^b$, improving marginal deterrence. The following lemma states the optimal policy under the discriminating rule and limited liability.

Lemma 4 (Optimal enforcement policy under discriminating rule and limited liability). *Under a discriminating rule, there exists a $\beta_4 > \beta_0$ such that the limited liability constraint $\bar{f} \leqslant \pi$ does not bind for $\beta \in [0, \beta_4]$ when $\bar{f}$ is optimally set. In this interval the optimal policy is the one described in Lemma 2. Instead, for $\beta \in (\beta_4, 1]$, $w^g - w^b - \pi > 0$ and γ sufficiently large the optimal policy entails for increasing β, a reduction in type I accuracy and an improvement in type II accuracy. The actions implemented are $\widehat{a}^b = \varepsilon^{II}$ and $a^g = 1$. For $\beta \in (\beta_4, 1]$, the expected welfare $EW_{\bar{D}}(\beta)$ it is decreasing in β.*

It is interesting to observe that when the limited liability constraint binds, the enforcer reduces type-I accuracy while improving type-II accuracy. From Lemma 2 we observed that type-I accuracy is improved under a discriminating regime to reduce over-enforcement and sustain the investment. When fines are capped, instead, type-I accuracy is progressively weakened as β increases. Since the implemented action in the bad state, $\widehat{a}^b$, is larger than in case no cap on fines is set, this distortion itself makes the investment more profitable with no need to costly reduce type-I errors. On the other hand, to limit this distortion, since $\widehat{a}^b = \varepsilon^{II}$, type-II accuracy is improved.

In the following proposition we summarize the optimal legal standards.

Proposition 3 (Optimal legal standards under limited liability). *When fines are capped by limited liability, the optimal legal standard for increasing values of β is:*

(i) *for $\beta \in [0, \beta_0)$ per-se legality;*
(ii) *for $\beta \in [\beta_0, \beta_4)$ the discriminating rule with type-I accuracy;*
(iii) *for $\beta \in [\beta_4, \beta_5)$ the discriminating rule with the limited liability constraint binding, with $\widehat{a}^b$ higher than in the case when no cap on fines applies (Lemma 2), $\varepsilon^I(\beta)$ increasing in β with $\varepsilon^I(1) = \bar{\varepsilon}$ and $\varepsilon^{II}(\beta)$ decreasing in β;*
(iv) *for $\beta \in [\beta_5, 1]$ per-se illegality with $\widehat{a} = 0$.*

Hence, when maximum fines are capped we find a second reason why a more rigid per-se rule may dominate the more flexible discriminating legal standard, based 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 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

[19] 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 $\widetilde{a}^b = 1 - (1-\varepsilon^{II})\frac{F}{\pi}$.

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. This argument was first proposed by Easterbrook (1984), that identified a series of filters (empirical tests) to scrutinize the predictions. 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.[20] 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.[21]

5. 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 marginal costs are increasing in its size.

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 $\Pi(a, I) = I(\Pi + \pi a) - (I^2/2)$.[22]

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: $W^b(a, I) = I(W - w^b a) \leqslant 0$ when the practice reduces welfare and $W^g(a, I) = I(W + w^g a) \geqslant 0$ 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 $I(\Pi + \pi a)$, contrary to the case of research investment, where the *ex ante* (gross profits) are $I(\Pi + \pi a)$ but the *ex post* profits in case of successful innovation are given by $\Pi + \pi a$.

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

$$I_{PS} = \Pi + \pi\widetilde{a}. \tag{12}$$

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

$$\bar{f} \in \left[\underline{f} + \frac{I\pi(1 - \widetilde{a}^b)}{1 - \varepsilon^{II}}, \underline{f} + \frac{I\pi(1 - \widetilde{a}^b)}{\varepsilon^I}\right]. \tag{13}$$

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

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

and the optimal investment in physical assets is therefore

$$I_D = \Pi + (1 - \beta)\pi + \beta\pi\widetilde{a}^b. \tag{14}$$

Notice that, both for per-se and discriminating rules, when the 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 $\widetilde{a}^b$, 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 increases 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.

[20] See Easterbrook (1984) for an early discussion of the issue and Polo (2010) for an application to selective price cuts.

[21] Our analysis of the optimal enforcement policy has focused 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 σ). 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.

[22] 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.

Proposition 4 (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 $\beta \leqslant \beta_0$ and a discriminating rule for higher β.*

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.

Finally, if property rights on the physical capital allows appropriating most of the social surplus from the investment (high Π for given W), a stricter antitrust enforcement is applied, and vice-versa.

6. Sunk investment

Since the impact of antitrust intervention (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 undertaken; secondly, which is the optimal antitrust intervention in industries where new 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 level of investment. Drawing from our previous discussion, it is evident that in this alternative case *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 $\widehat{a}^b$, can implement the first best course of actions $a^g = 1$ and $a^b = 0$. In other words, there is no need to implement an action $a^b > 0$ in the bad state to boost the investment, since the investment is already sunk. Hence, if the enforcer selects the legal standard for a given investment, a discriminating rule dominates for any value of the probability β.

Proposition 5 (Sunk investment). *If the legal standard and enforcement policy are chosen once the investment has been sunk, the enforcer applies the discriminating rule and implement the first best course of actions for any $\beta \in [0, 1]$.*

This result[23] has important implications for two relevant economic environments. In industries where the investment matters, the result underlines the importance of committing to a certain policy. Choosing the policy once the research effort is sunk leads to a too interventionist approach, abandoning per-se legality when it would be selected in the benchmark model, an instance of regulatory hold-up. Then, in environments where research investment matters, the enforcer faces a time inconsistency problem that can be solved by committing to an enforcement policy and a legal standard before the investment is chosen, for instance by adopting regulations or guidelines, or through precedents.

Instead, in industries where there is no relevant investment issue, the discriminating rule emerges as the optimal legal standard for any prior on the effects of the practice. This may be the case of mature industries where technological progress is not a major element of the competitive game. Another interesting application refers to public utilities in the early stage of liberalization, where the network infrastructure inherited from the previous monopoly phase was already in place, and antitrust policy intervened against the incumbents to prevent foreclosure on a wide range of practices, a more interventionist policy for given priors, as Proposition 5 suggests. The present debate in energy and telecommunication markets recognizes that investment in new infrastructures is becoming a primary goal of public policies, requiring some balancing between promoting competition and providing incentives to investment. This evolution in the European policy debated seems to suggest a shift from the more interventionist approach described in Proposition 5 to the more lenient one of Proposition 1.

7. Conclusions

We have studied optimal antitrust policy in a setting where firms make two choices: an *ex ante* investment in research that produces a socially valuable innovation when the investment succeeds and a business practice (tying, bundling, marketing, etc.), which increases the firm's profit the more aggressively it is pursued, but may or may not be socially valuable depending on the state of the world. Thus, a general feature of our optimal enforcement policy, refers to balancing the need to deter the practice when unlawful, a concern that is relevant *ex post*, and the attempt to sustain the investment by adopting a more lenient policy when, *ex ante*, the practice is expected to improve welfare. For instance, a per-se legality rule may be preferred to a discriminating rule when the probability of the bad state is sufficiently low. The intuition is that, a per-se legality rule prevents the enforcer from harmful over-regulation.

More generally, we have shown that optimal legal standards and enforcement policies in antitrust intervention depend on the priors of the enforcer regarding the economic effects of the practices, i.e. on the parameters (W, Π, w^g, w^b, π, β). Under this respect, our results recall the ongoing debate between different approaches in antitrust. Economic approaches that have stressed the efficiency enhancing effects of many business practices (low β), 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 (higher β), usually associated with the post-Chicago scholars, has represented the background for the effect-based approach to unilateral practices.

The debate following the judgement of the Court of First Instance on the EC vs. Microsoft decision offers some interesting elements that in our view fit the analysis of this paper. Ahlborn and Evans (2009) in their critical review of the judgement first identify the approach adopted by the Court with the ordoliberal antitrust tradition, that relies on "structural presumptions and a form-based analysis rather than an assessment of the effects of the conduct on consumer welfare". Hence, according to the authors, the Court of First Instance adopted a per-se illegality rule without evaluating the effects of the two main practices investigated: tying of the media player with the operating system, and limited interoperability with competitors' server operating systems. Secondly, they point out a different approach of the Community Courts in

[23] 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.

abuse of dominance cases compared with mergers and coordinated practices: "while in these other areas of competition law the European Court of Justice and the Court of First Instance have embraced economic reasoning and have set a high bar for the Commission, in terms of logic and evidence, there has been no sign of development under Article 82 for the last 40 years: the Court's policy continues to follow a form-based approach, based on ideas and concepts derived in the pre-Chicago world". This quotation suggests that discriminating and per-se rules may apply to different areas of competition policy. Moreover, the role of the enforcers's priors in guiding towards different legal standards is also suggested by Ahlborn and Evans (2009) when they write that "part of the answer to this question may be found in the Court's greater willingness to presume that mergers are more benign than the activities of dominant firms". And this presumption might have been reinforced by the super-dominant position that Microsoft has in the operating system market, a fact that the defendant did not contest.

To conclude, although we agree with Judge Easterbrook (1984)[24] that antitrust is an imperfect tool for the regulation of competition, we argue that our results help addressing the issue of antitrust policy design in complex environments where practices and investment decisions interact.

Acknowledgments

We are grateful to Scott Baker (the Editor) and one anonymous referee for insightful suggestions. We are also 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) Universitat Pompeu Fabra (Barcelona) and TILEC Tilburg University. All usual disclaimers apply. This paper supersedes an old version circulated under the title: "Judicial errors and innovative activity."

Appendix A.

Proof of Proposition 1. From the previous discussion, the optimal investment is $I_{PS} = \Pi + \pi\widetilde{a} - \underline{f}$ and the expected welfare is $EW_{PS}(\beta) = I_{PS}(W + Ew(\beta)\widetilde{a}) - \frac{I_{PS}^2}{2}$. Then, the maximization program $\max \mathcal{L}_{PS} = EW_{PS}(\beta) + \lambda\left(\widetilde{a} - 1 + \frac{\bar{f} - \underline{f}}{\pi}\right)$ is solved by the following first-order conditions

$$\begin{aligned} \frac{\partial \mathcal{L}_{PS}}{\partial \widetilde{a}} &= [W + Ew(\beta)\widetilde{a} - I_{PS}]\pi + Ew(\beta)I_{PS} + \lambda \geqslant 0, \\ \frac{\partial \mathcal{L}_{PS}}{\partial \underline{f}} &= -[W + Ew(\beta)\widetilde{a} - I_{PS}] - \frac{\lambda}{\pi} \leqslant 0, \\ \frac{\partial \mathcal{L}_{PS}}{\partial \bar{f}} &= \frac{\lambda}{\pi} \geqslant 0, \end{aligned} \tag{15}$$

Finally, the complementary slackness condition is

$$\lambda\left(\widetilde{a} - 1 + \frac{\bar{f} - \underline{f}}{\pi}\right) = 0. \tag{16}$$

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 λ 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 three or four possible subcases, depending on whether β_3 is larger or lower than 1. It is convenient to introduce the following expression:

$$G(\beta, \widetilde{a}) = W - \Pi + (Ew(\beta) - \pi)\widetilde{a}.$$

We discuss first case (a) where $W - \Pi - \pi < 0$.

(i) Consider first the interval $\beta \in [0, \beta_1]$ where β_1 is such that $G(\beta_1, 1) = 0$. Then, $G(0, \widetilde{a}) > 0$ for any $\widetilde{a} \in [0, 1]$ and $G(\beta, \widetilde{a})$ is increasing in $\widetilde{a}$ for $\beta < \frac{w^g - \pi}{w^g + w^b} < \beta_1$, whereas $G(\beta, \widetilde{a})$ is decreasing in $\widetilde{a}$ for $\frac{w^g - \pi}{w^g + w^b} < \beta \leqslant \beta_1$, reaching its lowest value $G(\beta_1, 1) = 0$. Hence, $G(\beta, \widetilde{a}) > 0$ for $\beta < \beta_1$ and for any $\widetilde{a} \in [0, 1]$. Then, if we set $\underline{f} = 0$, we get $\frac{\partial \mathcal{L}_{PS}}{\partial \underline{f}} = -G(\beta, \widetilde{a}) < 0$ and indeed $\underline{f} = 0$ is optimal. Moreover, $\frac{\partial \mathcal{L}_{PS}}{\partial \widetilde{a}} = G(\beta, \widetilde{a})\pi + Ew(\beta)(\Pi + \pi\widetilde{a}) > 0$ since $Ew(\beta) > 0$ for $\beta < \frac{w^g}{w^g + w^b} = \beta_2 (> \beta_1$ in case (a)). Hence, it is optimal to set $\widehat{a} = 1$.

(ii) For $\beta \in (\beta_1, \beta_2)$, $Ew(\beta) > 0 > W + Ew(\beta) - \Pi - \pi$. We solve $\frac{\partial \mathcal{L}_{PS}}{\partial \underline{f}} = -G(\beta, 1) - \underline{f} = 0$ by setting $\underline{f} = -G(\beta, 1)$. Moreover, $\frac{\partial^2 \mathcal{L}_{PS}}{\partial \underline{f}^2} = -1$, satisfying the second order conditions. Then, $\frac{\partial \mathcal{L}_{PS}}{\partial \widetilde{a}} = Ew(\beta)I_{PS} > 0$ and $\widehat{a} = 1$. Substituting $\underline{f}$ in the expression of the optimal investment we obtain $I_{PS} = W + Ew(\beta) > 0$ that is decreasing in β and equal to W when $\beta = \beta_2$.

(iii) For $\beta \in [\beta_2, \beta_3)$, $(W - \Pi)\pi + Ew(\beta)\Pi > 0 \geqslant Ew(\beta)$. Then, if we set $\frac{\partial \mathcal{L}_{PS}}{\partial \widetilde{a}} = [W + Ew(\beta)\widetilde{a} - \Pi - \pi\widetilde{a}]\pi + Ew(\beta)(\Pi + \pi\widetilde{a}) = 0$, we obtain $\frac{\partial \mathcal{L}_{PS}}{\partial \underline{f}} < 0$, implying that $\underline{f} = 0$. Solving $\frac{\partial \mathcal{L}_{PS}}{\partial \widetilde{a}} = 0$ we get $\widehat{a} = \frac{(W - \Pi)\pi + Ew(\beta)\Pi}{(\pi - 2Ew(\beta))\pi}$ which is strictly larger than zero for $\beta < \beta_3$. Substituting $\widehat{a}$ in the investment we obtain $I_{PS} = \frac{\pi W - Ew(\beta)\Pi}{\pi - 2Ew(\beta)}$ and the incentive compatibility constraint is satisfied for any $\bar{f} \geqslant \pi\left(1 - \widehat{a}\right)$. Notice that β_3 may be larger or lower than 1 according to our assumptions. In the former case, this region extends to $\beta = 1$. Conversely, if $\beta_3 < 1$, we have a further case.

(iv) For $\beta \in [\beta_3, 1]$, $0 \geqslant (W - \Pi)\pi + Ew(\beta)\Pi > Ew(\beta)$ implying $\partial \mathcal{L}_{PS}/\partial \widetilde{a} < 0$ and $\partial \mathcal{L}_{PS}/\partial \underline{f} < 0$. So that $\widehat{a} = 0$ and $\underline{f} = 0$. Substituting $\widehat{a}$ 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 $\bar{f} \geqslant \pi$. It is immediate to see that in all cases the undertake constraint is satisfied.

We can now briefly review case (b) where $W - \Pi - \pi > 0$. In this case $\beta_2 < \beta_1' < \beta_1$.

(i) Consider first $\beta \in [0, \beta_1']$: since in case (b), $\beta_1' < \beta_1$, we have $G(\beta_1', 1) > 0$, $\frac{\partial \mathcal{L}_{PS}}{\partial \underline{f}} < 0$ and $\underline{f} = 0$. Then from the definition of β_1', $\frac{\partial \mathcal{L}_{PS}}{\partial \widetilde{a}} = G(\beta_1', 1)\pi + Ew(\beta_1')(\Pi + \pi) \geqslant 0$ for $\beta \leqslant \beta_1'$ and $\widehat{a} = 1$.

(iii) works as under case (a) in the interval $\beta \in [\beta_1', \beta_3]$.

(iv), if it exists, is as under case (a). □

Proof of Lemma 2. We solve our problem by omitting the incentive compatibility constraints (8) and the undertake constraints and verifying them *ex post*. Recall from the text the expressions for the innovative investment $I_D = \Pi + (1 - \beta)[\pi - \varepsilon^I\bar{f}] + \beta[\pi\widetilde{a}^b - (1 - \varepsilon^{II})\underline{f}]$ and for the expected welfare $EW_D =$

[24] "Condemnation per-se rests on a conclusion that all or almost all examples of some category of practices are inefficient, yet we cannot reach such a judgment for any practice other than naked horizontal restraints. The traditional Rule of Reason falls prey to all of the limits of antitrust. It assumes that judges can tap a fount of economic knowledge that does not exist, and it disregards the costs of judicial decision making (including the costs of damning efficient conduct by mistake or design)." Easterbrook (1984).

$I_D\left[W+\Delta W_D-\frac{I_D}{2}\right]-\frac{\gamma}{2}(\bar{\varepsilon}-\varepsilon^I)^2-\frac{\gamma}{2}(\bar{\varepsilon}-\varepsilon^{II})^2$, where $\Delta W_D=(1-\beta)w^g-\beta w^b\widehat{a}^b$. The first order conditions are the following:

$$\frac{\partial EW_D}{\partial \widehat{a}^b}=[W+\Delta W_D-I_D]\beta\pi-\beta w^b I_D \gtrless 0$$
$$\frac{\partial EW_D}{\partial \underline{f}}=-[W+\Delta W_D-I_D]\beta(1-\varepsilon^{II})<0$$
$$\frac{\partial EW_D}{\partial \bar{f}}=-[W+\Delta W_D-I_D](1-\beta)\varepsilon^{I}<0$$
$$\frac{\partial EW_D}{\partial \varepsilon^I}=-[W+\Delta W_D-I_D](1-\beta)\bar{f}+\gamma(\bar{\varepsilon}-\varepsilon^I)\gtrless 0$$
$$\frac{\partial EW_D}{\partial \varepsilon^{II}}=[W+\Delta W_D-I_D]\beta\underline{f}+\gamma(\bar{\varepsilon}-\varepsilon^{II})\gtrless 0.$$

Let us consider the following candidate solution and check in which interval of β it holds: $\underline{f}=\bar{f}=0$ and $\widehat{a}^b=1$. Substituting in $I_D=\Pi+\pi$ and ΔW_D we have, for $\beta<\beta_0$, $\frac{\partial EW_D}{\partial \widehat{a}^b}=\beta\pi\left[W+w^g-\Pi-\pi-w^b(\frac{\Pi+\pi}{\pi})-\beta(w^g+w^b)\right]>0$. Hence, for $\beta<\beta_0$, $\frac{\partial EW_D}{\partial \widehat{a}^b}>0$, $\frac{\partial EW_D}{\partial \underline{f}}<0$ and $\frac{\partial EW_D}{\partial \bar{f}}<0$ at $\varepsilon^I=\varepsilon^{II}=\bar{\varepsilon}$. Finally, the incentive compatibility constraints (8) and the undertake constraints are clearly satisfied.

Consider next the case $\beta>\beta_0$. We set $\widehat{a}^b<1$ to obtain $\frac{\partial EW_D}{\partial \widehat{a}^b}=0$, implying that $[W+\Delta W_D-I_D]>0$. 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 (8), that is, $\bar{f}=\frac{\pi(1-\widehat{a}^b)}{(1-\bar{\varepsilon})}$. At this equilibrium, the undertake constraints are satisfied since $\Pi+\pi\widehat{a}^b-0\geqslant 0$ and $\Pi+\pi-\varepsilon^I\bar{f}=\Pi+\pi-\varepsilon^I\frac{\pi(1-\widehat{a}^b)}{(1-\bar{\varepsilon})}>\Pi+\pi\left[\frac{1-\bar{\varepsilon}-\varepsilon^I}{1-\bar{\varepsilon}}\right]>0$.

For $\beta>\beta_0$, to check the second order condition, notice that only $\widehat{a}^b$ and ε^I are set at an internal solution. Hence,

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

for γ sufficiently large. □

Proof of Lemma 3. For $\beta>\beta_0$, rearranging the first order conditions in the proof of Lemma, we get the following expression for the implemented action as a function of the optimal type-I error

$$\widehat{a}^b=\frac{(1-\bar{\varepsilon})(W-\Pi-w^b\frac{\Pi}{\pi})+(1-\beta)\left[(1-\bar{\varepsilon})w^g-(1-\bar{\varepsilon}-\varepsilon^I)(w^b+\pi)\right]}{(1-\beta)\varepsilon^I(w^b+\pi)+\beta(1-\bar{\varepsilon})(2w^b+\pi)} \quad (17)$$

where the second term at the numerator is positive for $\beta<1$ and tends to 0 for $\beta\to 1$. Instead, the sign of the first term coincides with the sign of $W-\Pi-w^b\frac{\Pi}{\pi}$ which can be shown to be positive if $\beta_3>1$. In this case, $\widehat{a}^b(\beta=1)=\frac{W-\Pi-w_b\frac{\Pi}{\pi}}{2w^b+\pi}$. If, instead, $\beta_3<1$, $W-\Pi-w^b\frac{\Pi}{\pi}$ is negative and $\widehat{a}^b(\beta\to 1)=0$, possibly with a corner solution. Finally, notice also that β_0 may be larger or lower than 0. In the latter case β is always larger than β_0 and $\widehat{a}^b<1$ for any β.

From the first order conditions we also get the following expression for the investment (always as a function of the optimal type-I error):

$$I_D=\Pi+(1-\beta)\left[\pi-\varepsilon^I\frac{\pi(1-\widehat{a}^b)}{(1-\bar{\varepsilon})}\right]+\beta\pi\widehat{a}^b. \quad (18)$$

Notice that (17) and (18) are not the equilibrium value, since they both depend on the equilibrium level of type-I error ε^I, and they can be evaluated only at the extremes of the interval. To further study the effect of β on the equilibrium value of $\widehat{a}^b$ we differentiate the first order conditions with respect to $\widehat{a}^b$, ε^I, β and we find that

$$sign\frac{d\widehat{a}^b}{d\beta}=sign\left(-\frac{\partial^2 EW_D}{\partial\beta\partial\widehat{a}^b}\frac{\partial^2 EW_D}{\partial\varepsilon^{I2}}+\frac{\partial^2 EW_D}{\partial\varepsilon^I\partial\widehat{a}^b}\frac{\partial^2 EW_D}{\partial\varepsilon^I\partial\beta}\right),$$

where

$$\frac{\partial^2 EW_D}{\partial\varepsilon^I\partial\widehat{a}^b}=\frac{\partial I_D}{\partial\varepsilon^I}\left[-\beta w^b-\frac{1}{2}\frac{\partial I_D}{\partial\widehat{a}^b}\right]+\frac{\partial^2 I_D}{\partial\varepsilon^I\partial\widehat{a}^b}[W+\Delta W_D-I_D]>0,$$

since $\frac{\partial^2 I_D}{\partial\varepsilon^I\partial\widehat{a}^b}=\frac{(1-\beta)\pi}{(1-\bar{\varepsilon})}>0$, $\frac{\partial I_D}{\partial\widehat{a}^b}>0$ and $\frac{\partial I_D}{\partial\varepsilon^I}<0$. Moreover,

$$\frac{\partial^2 EW_D}{\partial\varepsilon^I\partial\beta}=\frac{\pi(1-\widehat{a}^b)}{(1-\bar{\varepsilon})}[W+\Delta W_D-I_D]-\frac{(1-\beta)\pi(1-\widehat{a}^b)}{(1-\bar{\varepsilon})}\left[-w^g-w^b\widehat{a}^b-\frac{1}{2}\frac{\partial I_D}{\partial\beta}\right]>0$$

and

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

Multiplying the previous expression by β we notice that

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

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

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

or equivalently as

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

since $(\frac{1-\bar{\varepsilon}-\varepsilon^I}{1-\bar{\varepsilon}})$ is smaller than one and $w^g>w^b+\pi$. Then, $\frac{\partial^2 EW_D}{\partial\widehat{a}^b\partial\beta}<0$ and $\frac{d\widehat{a}^b}{d\beta}<0$ when γ (that is in the expression for $\frac{\partial^2 EW_D}{\partial\varepsilon^{I2}}$) is sufficiently large. Hence, $\widehat{a}^b$ decreases from 1 to 0 as β varies from β_0 to 1. □

Proof of Proposition 1. The per-se rule imposes to treat the actions in both states in the same way, either considering them legal or unlawful. Conversely, under the discriminating rule the practice is illegal only when a negative signal is received. To evaluate which regime is optimal, it is convenient to analyze the optimal policies under a *modified regime*, where the enforcer can implement through fines both a^g and a^b, still basing its intervention on the noisy signal σ. Notice that this case is different from the per-se rule, where the action must be treated *in the same way* irrespective of the signal, and from the discriminating rule, where the action in the good state *cannot be sanctioned*. However, this modified regime encompasses the per-se rules and the discriminating rule: if the optimal policy under this modified regime prescribes to implement the same action in both states, then the errors are irrelevant and the per-se rule is the optimal legal standard. Conversely, when the optimal policy under the modified regime implements $a^g = 1 > a^b$, it does not exploit the possibility of sanctioning the practice in the good state, an option that is exogenously prevented under the discriminating rule. This latter constraint, then would not be binding, and the optimal policy under the modified regime would be equivalent to the discriminating rule. Hence, if either of the two policy configurations is optimal in the modified regime for certain values of β, it follows that the corresponding legal standard is the optimal one.

In the modified regime, the innovative investment is $I_D = \Pi + (1-\beta)[\pi\widetilde{a}^g - \varepsilon^I\bar{f}] + \beta[\pi\widetilde{a}^b - (1-\varepsilon^{II})\underline{f}]$, while the expected welfare is $EW_D = I_D[W + \Delta\widetilde{W}_D - \frac{I_D}{2}] - \frac{\gamma}{2}(\bar{\varepsilon} - \varepsilon^I)^2 - \frac{\gamma}{2}(\bar{\varepsilon} - \varepsilon^{II})^2$, where $\Delta\widetilde{W}_D = (1-\beta)w^g\widetilde{a}^g - \beta w^b\widetilde{a}^b$. The first order conditions are:

$$\frac{\partial EW_D}{\partial \widetilde{a}^b} = [W + \Delta\widetilde{W}_D - I_D]\beta\pi - \beta w^b I_D \geqslant 0$$
$$\frac{\partial EW_D}{\partial \widetilde{a}^g} = [W + \Delta\widetilde{W}_D - I_D](1-\beta)\pi + (1-\beta)w^g I_D \geqslant 0$$
$$\frac{\partial EW_D}{\partial \underline{f}} = -[W + \Delta\widetilde{W}_D - I_D]\beta(1-\varepsilon^{II}) < 0$$
$$\frac{\partial EW_D}{\partial \bar{f}} = -[W + \Delta\widetilde{W}_D - I_D](1-\beta)\varepsilon^I < 0$$
$$\frac{\partial EW_D}{\partial \varepsilon^I} = -[W + \Delta\widetilde{W}_D - I_D](1-\beta)\bar{f} + \gamma(\bar{\varepsilon} - \varepsilon^I) \geqslant 0$$
$$\frac{\partial EW_D}{\partial \varepsilon^{II}} = [W + \Delta\widetilde{W}_D - I_D]\beta\underline{f} + \gamma(\bar{\varepsilon} - \varepsilon^{II}) \geqslant 0.$$

Following the same arguments used in the proof of Lemma 2 we show that: for $\beta \leqslant \beta_0$, $\frac{\partial EW_D}{\partial \widetilde{a}^g} > 0$ and $\frac{\partial EW_D}{\partial \widetilde{a}^b} > 0$, implying $\widehat{a}^b = \widehat{a}^g = 1$, $\underline{f} = 0$ and $\varepsilon^I = \varepsilon^{II} = \bar{\varepsilon}$; for $\beta > \beta_0$, $\frac{\partial EW_D}{\partial \widetilde{a}^b} = 0$, $W + \Delta\widetilde{W}_D - I_D > 0$ and therefore $\frac{\partial EW_D}{\partial \widetilde{a}^g} > 0$, implying $\widehat{a}^b < \widehat{a}^g = 1$. Moreover, $\underline{f} = 0$ and $\varepsilon^I < \varepsilon^{II} = \bar{\varepsilon}$.

Hence, when the enforcer can implement the actions in either state and receives a noisy signal, it implements the policy equivalent to the per-se legality rule for $\beta \leqslant \beta_0$ and the policy corresponding to the discriminating rule for higher values of β. Put another way, the constraint that prevents the enforcer from sanctioning the action in the good state, a^g, is not binding, because even when allowed to do so, the enforcer would never implement an action $\widehat{a}^g < 1$. The first order conditions for the optimal policy, then, are identical to those obtained, under the condition $a^g = 1$, in the proof of Lemma 2.

This result holds when, under a per-se rule, $W - \Pi - \pi$ is negative (case a) or positive (case b). In the former, $\beta_0 < \beta_1$ and the modified regime implements a policy equivalent to the discriminating rule for $\beta > \beta_0$, although the policy configurations prescribed by the per-se rules would still be implementable. In the latter case, direct inspection shows that, if $(W - \Pi - \pi)\pi \geqslant w^b(\Pi + 2\pi)$, then $1 < \beta_1' \leqslant \beta_0$, and the per-se legality rule would be adopted for any β, while for $(W - \Pi - \pi)\pi < w^b(\Pi + 2\pi)$, $\beta_0 < \beta_1'$. In this latter case, above β_0, the modified regime implements the discriminating rule as in case (a). We conclude that the optimal legal standard for given β does not change no matter if case (a) or (b) apply. □

Proof of Lemma 4. We denote by subscript $_{\bar{D}}$ the discriminating regime with capped fines. Combining the incentive compatibility and limited liability constraints by setting $\bar{f} = \pi$ and $\underline{f} = 0$ in (8) we obtain $\widehat{a}^b = \varepsilon^{II}$. Notice that the implemented action under limited liability constraint is higher than the one implemented without such constraint in Lemma 2, where $\frac{\partial EW_{\bar{D}}}{\partial \widetilde{a}^b} = 0$. Hence, at $\widehat{a}^b = \varepsilon^{II}$ we have $[W + \Delta W_D - I_{\bar{D}}]\beta\pi - \beta w^b I_{\bar{D}} < 0$.

Substituting the implementable actions in the expression of the investment we get

$I_{\bar{D}} = \Pi + \pi[1 - \varepsilon^I - \beta(1 - \varepsilon^I - \varepsilon^{II})]$.

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

$$\frac{\partial EW_{\bar{D}}}{\partial \varepsilon^I} = -[W + \Delta W_D - I_{\bar{D}}]\pi(1-\beta) + \gamma(\bar{\varepsilon} - \varepsilon^I) \geqslant 0$$
$$\frac{\partial EW_{\bar{D}}}{\partial \varepsilon^{II}} = [W + \Delta W_D - I_{\bar{D}}]\pi\beta - \beta w^b I_{\bar{D}} + \gamma(\bar{\varepsilon} - \varepsilon^{II}) \geqslant 0.$$

The second expression clearly holds as an equality with $\varepsilon^{II} < \bar{\varepsilon}$. Moreover, as β increases, the difference between the unconstrained and the constrained action $\widehat{a}^b$ becomes larger and the term $[W + \Delta W_D - I_{\bar{D}}]\beta\pi - \beta w^b I_{\bar{D}}$ more and more negative, implying a lower and lower ε^{II}. Turning to the first order conditions of ε^I, since $W + \Delta W_D - I_{\bar{D}} > 0$ when the limited liability constraint does not bind, this term is positive even when the constraint binds, and the first expression is solved as an equality with $\varepsilon^I < \bar{\varepsilon}$. As β increases, this term becomes smaller and at some point negative, implying an increasing ε^I, up to the point where no type-I accuracy is exerted.

Notice that for $\bar{f} = \pi$, $\underline{f} = 0$, $\widehat{a}^b = \varepsilon^{II}$ and $a^g = 1$ the undertake constraints are also satisfied.

Finally, the second order conditions hold, since

$$\frac{\partial^2 EW_{\bar{D}}}{\partial \varepsilon^{I2}} = -\left(\frac{\partial I_{\bar{D}}}{\partial \varepsilon^I}\right)^2 - \gamma < 0$$
$$\frac{\partial^2 EW_{\bar{D}}}{\partial \varepsilon^{II2}} = -\left(\frac{\partial I_{\bar{D}}}{\partial \varepsilon^{II}}\right)^2 - \gamma < 0$$
$$H_{\varepsilon^I\varepsilon^{II}} = \gamma\left[\left(\frac{\partial I_{\bar{D}}}{\partial \varepsilon^I}\right)^2 + \left(\frac{\partial I_{\bar{D}}}{\partial \varepsilon^{II}}\right)^2\right] + \gamma^2 > 0.$$

Differentiating with respect to β the expected welfare we get

$$\frac{dEW_{\bar{D}}}{d\beta} = \frac{\partial EW_{\bar{D}}}{\partial \beta} + \frac{\partial EW_{\bar{D}}}{\partial \varepsilon^I}\frac{\partial \varepsilon^I}{\partial \beta} + \frac{\partial EW_{\bar{D}}}{\partial \varepsilon^{II}}\frac{\partial \varepsilon^{II}}{\partial \beta},$$

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

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

is negative because $\frac{\partial I_{\bar{D}}}{\partial \beta} = -\pi(1 - \varepsilon^I - \varepsilon^{II})$ is negative and the same is true for the term in the second square bracket. □

Proof of Proposition 3. The argument in Proposition 1 still apply when the limited liability constraint does not bind. Due to Lemma 3, the implemented action decreases in β. For $\beta > \beta_4$, the maximum fine needed to implement the action in the bad state start binding and the action $\widehat{a}_b$ is progressively distorted upwards, while type II accuracy is improved and type I accuracy is reduced, as established in Lemma 4. The expected welfare under discriminating rule is continuous at $\beta = \beta_4$ and lower than in the case with no cap on fines for higher β, that is $EW_{\bar{D}}(\beta) < EW_D(\beta)$ for $\beta > \beta_4$. Finally, the problem of optimal discriminating rule with no caps on fines and optimal per-se rule coincide for $\beta = 1$, since only the bad state matters. Then, $EW_{\bar{D}}(1) < EW_D(1) = EW_{PS}(1)$. Then, there exists a $\beta_5 < 1$ such that the per-se illegality regime dominates for $\beta \in (\beta_5, 1]$. □

Proof of Proposition 4. The proof parallels the analysis of the benchmark case.

We first derive the optimal policies under per-se rules and discriminating rules, and then select the optimal legal standards.

Per-se rules: Recall from the text the expressions for the optimal investment $I_{PS} = \Pi + \pi\widetilde{a}$ and for the expected welfare $EW_{PS}(\beta) = I_{PS}(W + Ew(\beta)\widetilde{a}) - \frac{I_{PS}^2}{2}$. Then, the maximization program is solved by the following first-order conditions:

$$\frac{\partial EW_{PS}}{\partial \widetilde{a}} = [W + Ew(\beta)\widetilde{a} - I_{PS}]\pi + Ew(\beta)I_{PS} + \lambda \geqslant 0, \tag{19}$$

$$\frac{\partial EW_{PS}}{\partial \underline{f}} = -\frac{\lambda}{I_{PS}\pi} \leqslant 0, \tag{20}$$

$$\frac{\partial EW_{PS}}{\partial \bar{f}} = \frac{\lambda}{I_{PS}\pi} \geqslant 0, \tag{21}$$

while the complementary slackness conditions is

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

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 $(I(\Pi + \pi\widetilde{a}) - \underline{f} \geqslant 0)$, i.e. $\underline{f} \leqslant (\Pi + \pi\widetilde{a})^2$. This condition holds for sure when $\underline{f} = 0$. The analysis is almost identical to the one in Lemma 1, leading to the following regions:

(i) For $\beta \in [0, \beta_1]$ we have $W + Ew(\beta) - \Pi - \pi \geqslant 0$ and, because of $W < \Pi + \pi$, $\beta_1 < \beta_2$ which is the threshold such that $Ew(\beta_2) = 0$. Then $Ew(\beta) \geqslant 0$ for $\beta \in [0, \beta_1]$. If we set $\widehat{a} = 1$, in the first order conditions, we get $\frac{\partial EW_{PS}}{\partial \widetilde{a}} = [W + Ew(\beta) - \Pi - \pi]\pi + Ew(\beta)(\Pi + \pi) > 0$ then setting $\widehat{a} = 1$ is optimal.

(ii) For $\beta \in [\beta_1, \beta_3)$, $(W - \Pi)\pi + Ew(\beta)\Pi > 0 \geqslant Ew(\beta)$ where $\beta_3 = \frac{w^g + (W - \Pi)\pi}{w^g + w^b}$ and $\beta_3 > \beta_2$ (using $W > \Pi$). Then, if we set $\frac{\partial EW_{PS}}{\partial \widetilde{a}} = [W + Ew(\beta)\widetilde{a} - \Pi - \pi\widetilde{a}]\pi + Ew(\beta)(\Pi + \pi\widetilde{a}) = 0$, $\widehat{a}$ is interior and equal to $\frac{(W - \Pi)\pi + Ew(\beta)\Pi}{(\pi - 2Ew(\beta))\pi}$, $I_{PS} = \frac{\pi W - Ew(\beta)\Pi}{\pi - 2Ew(\beta)}$ and the incentive compatibility constraint is satisfied for any $\bar{f} \geqslant \pi(1 - \widehat{a})$.

(iii) For $\beta \in [\beta_3, 1]$, $0 \geqslant (W - \Pi)\pi + Ew(\beta)\Pi > Ew(\beta)$ implying $\partial EW_{PS}/\partial\widetilde{a} < 0$, so that $\widehat{a} = 0$. Substituting $\widehat{a}$ in the expression for the optimal investment and for the expected welfare we obtain $I_{PS} = \Pi$. Moreover, the incentive compatibility constraint is satisfied for any $\bar{f} \geqslant \pi$. It is immediate to see that in all cases the undertake constraint is satisfied.

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 ε^I and ε^{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 $\widetilde{a}^b$, that is $\frac{\partial EW_D}{\partial \widetilde{a}^b} = [W + \Delta W_D - I_D]\beta\pi - \beta w^b I_D \geqslant 0$, for the same argument developed in the case of research investment, $\widehat{a}^b = 1$ and $I_D = \Pi + \pi$. When $\beta > \beta_0$ the enforcer chooses an internal solution $\widehat{a}^b < 1$. The fines are set to meet the incentive compatibility constraint. For instance, the pair $\underline{f} = 0$ and $\bar{f} = \frac{I\pi(1 - \widehat{a}^b)}{(1 - \bar{\varepsilon})}$ 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 easily solve explicitly for the equilibrium action $\widehat{a}^b$. Solving as an equality $\frac{\partial EW_D}{\partial \widehat{a}^b} = 0$ we get the equilibrium action

$$\widehat{a}^b = \frac{(W - \Pi - w^b\frac{\Pi}{\pi}) + (1 - \beta)(w^g - w^b - \pi)}{\beta(2w^b + \pi)} \tag{22}$$

that is lower than 1 for $\beta > \beta_0$ and decreasing in β. Notice that, in case of research activity, the action $\widehat{a}^b$ in (17) evaluated at $\bar{\varepsilon} = \varepsilon^I = 0$ gives the expression above. Moreover, since (17) is increasing in $\bar{\varepsilon}$ and ε^I, it follows that the enforcer implements a higher action in case of research investment than in case of physical capital.

Optimal legal standard: the analysis to determine the optimal legal standard follows the one in Proposition 1 except that we do not have the area between β_1 and β_2. □

References

Aghion, P., Tirole, J., 1997. Formal and real authority in organizations. J. Polit. Econ. 105, 1–29.

Ahlborn, C., Evans, D., 2009. The Microsoft judgement and its implications for competition policy towards dominant firms in Europe. Antitrust Law J. 75, 887–932.

Anderlini, L., Felli, L., Immordino, G., Riboni, A., 2013. Legal institutions, innovation and growth. Int. Econ. Rev. 54, 937–956.

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.

Chang, H.F., 1995. Patent scope, antitrust policy, and cumulative innovation. Rand J. Econ. 26, 34–57.

Cremer, J., 1995. Arm's length relationships. Q. J. Econ. 110, 275–295.

Demougin, D., Fluet, C., 2008. Rules of proof, courts and incentives. Rand J. Econ. 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 Rev. 63, 9–14.

Erkal, N., 2004. The Decision to Patent, Cumulative Innovation, and Optimal Policy. Intellectual Property Research Institute of Australia, wp. 11/03.

European Commission, 2013. Commission recommendation on consistent non-discrimination obligations and costing methodologies to promote, Sept. 9, 2013. competition and enhance the broadband investment environment.

Evans, D., Padilla, J., 2005. Designing antitrust rules for assessing unilateral practices: a Neo.Chicago approach. Univ. Chicago Law Rev. 72, 73–98.

Green, J., Scotchmer, S., 1995. On the division of profit in sequential innovation. Rand J. Econ. 26, 20–33.

Gual, J., Hellwig, M., Perrot, A., Polo, M., Rey, P., Schmidt, K., Stenbacka, R., 2005. An economic approach to Article 82," Report for the DG Competition, European Commission. Compet. Policy Int. 2, 111–156.

Immordino, G., Pagano, M., Polo, M., 2011. Incentives to innovate and social harm: Laissez-faire, authorization or penalties? J. Public Econ. 95, 864–876.

Kaplow, L., 1994. The value of accuracy in adjudication: an economic analysis. J. Legal Stud. 15, 371–385.

Kaplow, L., 2011a. Optimal proof burdens, deterrence and the chilling of desirable behavior. Am. Econ. Rev. Pap. Proc. 101, 277–280.

Kaplow, L., 2011b. On the optimal burden of proof. J. Polit. Econ. 119, 1104–1140.

Kaplow, L., 2012. Burden of proof. Yale Law J. 121, 738–859.

Kaplow, L., Shavell, S., 1994. Accuracy in the determination of liability. J. Law Econ. 37, 1–15.

Kaplow, L., Shavell, S., 1996. Accuracy in the assessment of damages. J. Law Econ. 39, 191–210.

Katsoulakos, Y., Ulph, D., 2009. On optimal legal standards for competition policy: a general welfare-based analysis. J. Ind. Econ. LVII:, 410–437.

Mookherjee, D., Png, I., 1994. Marginal deterrence in the enforcement of law. J. Polit. Econ. 102, 1039–1066.

Polo, M., 2010. Anticompetitive vs competitive explanations of unilateral pratices: the identification problem. J. Compet. Econ. 6, 457–476.

Png, I.P.L., 1986. Optimal subsidies and damages in the presence of judicial error. Int. Rev. Law Econ. 6, 101–105.

Polinsky, M.A., Shavell, S., 2000. The economic theory of public enforcement of law. J. Econ. Lit. 38, 45–76.

Schwartzstein, J., Shleifer, A., 2012. An activity-generating theory of regulation. J. Law Econ. 56, 1–38.

Segal, I., Whinston, M., 2007. Antitrust in innovative industries. Am. Econ. Rev. 97, 1703–1730.

Stigler, G.J., 1970. The optimum enforcement of laws. J. Polit. Econ. 78, 526–536.