INNOCENZO
GASPARINI
INSTITUTE
FOR
ECONOMIC
RESEARCH

Institutional Members: CEPR, NBER and Università Bocconi

WORKING PAPER SERIES

Antitrust in Innovative Industries: the Optimal Legal Standards

Giovanni Immordino and Michele Polo

Working Paper n. 434

This Version: April 17, 2012

IGIER – Università Bocconi, Via Guglielmo Röntgen 1, 20136 Milano –Italy
http://www.igier.unibocconi.it

Antitrust in Innovative Industries: the Optimal Legal Standards

Giovanni Immordino
Università di Salerno and CSEF

Michele Polo
Università Bocconi and IGIER

April 17, 2012

Abstract: We study the interaction between a firm that invests in research and, if successful, undertakes a practice to exploit the innovation, and an enforcer that sets legal standards, fines and accuracy. In innovative industries deterrence on actions interacts with deterrence on research. A per-se legality rule prevails when the practice increases expected welfare, moving to a discriminating rule combined with type-I accuracy for higher probabilities of social harm. Moreover, discriminating rules should be adopted more frequently in traditional industries than in innovative environments; patent and antitrust policies are substitutes; additional room for per-se (illegality) rules emerges when fines are bounded.

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

JEL classification: D73, K21, K42, L51.

Acknowledgments: Giovanni Immordino Università di Salerno and CSEF, 84084 Fisciano (SA), Italy, giimmo@tin.it. Michele Polo, Università Bocconi, Via Sarfatti 25, 20136 Milan, Italy, michele.polo@unibocconi.it. We are indebted to Emanuela Carbonara, Jacques Cremer, Nuno Garoupa, Louis Kaplow, Yannis Katsoulakos, Dilip Mookherjee, Massimo Motta, Marco Ottaviani, Marco Pagano, Patrick Rey, Matteo Rizzolli, 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) and SIDE-ISLE 2011 (Torino). All usual disclaimers apply. This paper supersedes an old version circulated under the title: "Judicial errors and innovative activity."

1 Introduction

In recent years, the debate on competition policy has focussed on the role played by economics in improving the analysis of anticompetitive conducts. The discussion has raised issues concerning the substantial arguments as well as the legal standards that should be adopted in cases involving anticompetitive practices. Moreover, unilateral practices undertaken by technological market leaders in innovative industries have been scrutinized in a number of important cases in Europe and the US in the last decade. This paper analyzes optimal legal standards and antitrust enforcement policies towards anticompetitive practices in innovative industries.

Following the important reforms on cartel cases (article 101) and merger control, in 2009 the DG Competition of the European Commission has reshaped the enforcement of article 82 (now 102), pursuing an approach that rests on a deeper and more intelligent use of the new findings of economic analysis in the enforcement against unilateral practices.[1] A common view has emerged, labelled "effect-based" (or "more economic") as opposed to the traditional form-based approach. The novelty of these proposals refers to identifying anticompetitive practices through a careful analysis of the foreclosure effects of the conducts, beyond their formal description.

The debate on monopolization practices, as unilateral conducts are defined in the US, has developed also on the other side of the Atlantic. In 2008 the Antitrust Division of the Department of Justice issued a report (Department of Justice, 2008) on enforcement policies under Section 2 of the Sherman Act with the aim of setting clear standards. In an unusual contrast with the Antitrust Division, the Federal Trade Commission (FTC) opposed the guidelines, defined as a "blueprint for radically weakened enforcement" of Section 2. The new head of the Antitrust Division of the Department of Justice under the Obama administration, Christine Varney, decided in May 2009 to withdraw the draft paper, announcing a more aggressive approach to the enforcement of monopolization issues.

This brief summary highlights some major aspects of the recent debate. First, the academic discussion on several business practices, that are often taken as examples of unilateral anticompetitive conducts in antitrust enforcement, is quite open. While some economists argue that dominant firms adopt socially harmful practices to maintain their market power, others consider this possibility skeptically, stressing instead

[1] See Gual et al. (2005) and DG Competition (2005) and (2008).

the pursuit of superior efficiency as the driving force explaining the emergence of market leaders. In a brilliant summary of the evolution of economic thinking in antitrust, Evans and Padilla (2005) describe the pre-Chicago view as based on the recognition that dominant firms have the *ability* to adopt unilateral anticompetitive practices rather than on the investigation of their *incentives* to undertake such conducts. The Chicago revolution then took over the debate imposing "impossibility theorems" that denied any incentive to anticompetitive conducts, and strongly argued in favor of efficiency reasons behind many business practices. The authors conclude their review suggesting that the post-Chicago literature delivered a set of "possibility theorems" that lay down both efficiency and anticompetitive arguments as potential candidates to interpret the behavior of incumbents. Under this latter view, the economic features prevailing in a given market environment, concerning the type of competitors, the entry conditions, the market demand, etc. determine whether a given practice allows to foreclose the market or rather is part of the oligopolistic environment that does not cause harm to consumers.[2]

Secondly, the debate between different schools has extended from the economic arguments to be adopted in antitrust cases to the legal standards that the investigations should follow.[3] A wide range of proposals emerged, that can be roughly grouped into two sets: per-se rules that define legality or unlawfulness with reference to the conduct undertaken, and discriminating or effect-based rules that instead base the legal treatment of a certain practice on its anti-competitive or efficiency-enhancing effects.[4]

Academics, law scholars, practitioners and enforcement agencies have offered very

---

[2] These different views are partially rooted in different methodologies and analytical techniques, the Chicago approach being closer to the traditional price theory and the post-Chicago guys to the subtleties of game theoretic models. Moreover, recent developments in empirical Industrial Organization have improved our understanding of the factual relevance of many conducts, offering arguments to both sides. However, the experience suggests that a different weight on anticompetitive vs. efficiency explanations of market practices by incumbent firms may be sometimes rooted also into different priors, as the recent US discussion suggests.

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

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

different approaches on the expected motivations and effects of several business practices, a mixture of a-priori view and analytical reasoning, as well as on the legal standards that better fit the antitrust intervention in these matters.

Moving from the general debate to enforcement activity, unilateral practices undertaken by technological market leaders have been at the core of several landmark cases in the last decade in Europe and the US. In the American and European cases Microsoft was alleged of foreclosure on a number of practices such as bundling of the operating system and the browser or media player applications, loyalty rebates granted to pc producers and limited access, a mild form of refusal to deal, through a reduction in interoperability of the servers' and clients' operating systems. The record fine to Intel in the case before the European Commission was motivated, among other conducts, by foreclosure through loyalty rebates. In the last years the focus of antitrust enforcement seems to be moving towards new technological leaders as Google and Apple. The debate in competition policy has then raised new issues on the impact of antitrust enforcement on the innovative activity that characterizes these industries. For instance the commitments imposed in the EC v. Microsoft decisions to disclose the API codes of the server operating system to competitors, have been commented not only in their ability to restore competition, but also in their indirect adverse effects on the incentives to innovate. Hence, in the last decade the debate on the desirable legal standards in antitrust has often crossed the parallel issues raised by the activity of the enforcement agencies in innovative industries, identifying some additional effects that in these environments may arise.

In this paper we take into account the main ingredients of this debate studying the optimal legal standards and enforcement that an antitrust authority should adopt when regulating a certain business practice undertaken by a dominant firm. We focus our analysis on an innovative environment, where a firm breaks an initial competitive situation by introducing an innovation and this way becomes dominant, the kind of winner-takes-all competition that we often observe in high-tech industries. Once gained market power, the fresh incumbent is subject to antitrust scrutiny when undertaking commercial practices. The expected profits, then, reflect the stricter or laxer enforcement by the competition agency on the practices adopted by the innovator. In this setting antitrust enforcement faces the usual ex-post deterrence issue to affect the adoption of the practice by the incumbent. However, it has also to consider the ex-ante deterrence effect on the incentives to innovate that work through the expected profits from the practice once realized the innovation. We consider per-se and discriminating rules, deriving the optimal enforcement policies under each regime and then identifying the optimal legal standard for given expectations of the enforcer on the effects of the practice.[5]

We show how the optimal legal standard and enforcement policy vary when the enforcer's presumptions on the effects become more and more pessimistic. Specifically, a per-se legality rule prevails on the discriminating legal standard for low probability of social harm, since this regime is more effective in boosting the innovative investment. When the practice becomes more likely to reduce welfare, the enforcer adopts the discriminating rule and improves type-I accuracy to sustain investment.

We then consider three extensions of the benchmark model: first, we compare the results of the high-tech environment with the corresponding solution for an industry which has a well established and settled technology and does not face any relevant opportunity of innovation, what we call traditional industries. In these environments the discriminating rule is always dominant for any prior on social losses. Hence, interestingly, effect based rules should be adopted more frequently in traditional industries, where the only concern is deterrence of the practice when harmful, than in innovative environments. Secondly, we extend the model to include a positive effect of the new technology on profits and welfare even when the practice is not adopted, adding an additional motive to invest. In this framework, the baseline profits may be thought as guaranteed by patent protection, while the additional profits that can be obtained through the practice are affected by the antitrust policy. This way, we can consider in a simple setting the interaction of patent and competition policies. We show that, when the degree of patent protection is reduced, the region where antitrust policy opts for per-se legality, an extreme form of innovation-friendly antitrust intervention, becomes larger. In other words, patent and antitrust policies act as substitutes in our setting. Third, some additional room for per-se rules emerges, as a cost saving solution to enforcement, when fines are capped at some upper bound: per-se legality is adopted for low probability of social damages, then replaced by a discriminating rule with more and more symmetric accuracy, with per-se illegality as the optimal legal standard when the new technology is very likely to be socially harmful.

Our model contributes to the literature on antitrust and regulatory intervention

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

in industries. Immordino, Pagano and Polo (2011) propose an analytical framework similar to this paper to analyze the choice between different policy regime, namely *ex-post* law enforcement and *ex-ante* authorization, identifying when each policy is optimal. 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 two previous papers offer interesting results on law enforcement when innovative activity is a crucial component, they do not consider the choice among different legal standards that represents the focus of this paper. In Katsoulakos and Ulph (2009) a welfare analysis of legal standard is developed, which compares per-se rules and discriminating (effect based) rules. The authors identify some key elements that can help deciding the more appropriate legal standard and the cases in which type-I or type-II accuracy are more desirable. In their work the general setting is consistent with what we call traditional industries, while the case of innovative industries and the impact of enforcement on innovation, that is key in our paper, is not addressed.

Moreover, our results, although motivated with reference to competition policy and framed in terms of antitrust intervention, give useful insights in the more general debate on legal standards and accuracy in the law and economics literature[6].

The remainder of this paper is organized as follows. Section 2 presents the model. Section 3 focus on antitrust intervention in innovative industries. In Section 4 we consider the three extensions of the benchmark model. Section 5 offers some concluding comments. All proofs not following immediately from the main text are relegated to the Appendix.

---

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

2 The model

In this section we describe in detail how we model the interaction of antitrust intervention and innovative activity: we first analyze the private choices of the firm regarding the investment in research and the practice undertaken on the new technology once the innovation is realized; then we introduce the decisions of an antitrust authority on the legal standard adopted to evaluate the practice and the enforcement tools used to influence the firm's choices.

Private choices: practice intensity and research investment. 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.[7] The innovating firm, therefore, becomes dominant and subject to antitrust scrutiny. The amount of resources I that the firm invests in research determines its chances of success in the research process: for simplicity, the firm's probability of innovating $p(I)$ is assumed to be linear in I, i.e. $p(I) = I$ and $I \in [0, 1]$. The cost of learning is increasing and convex in the firm's investment and is assumed to be $c(I) = \frac{I^2}{2}$.

The (fresh) incumbent can exploit the new technology by adopting particular business strategies that yield extra-profits. Following the legal framework of antitrust intervention, business conducts are classified in a set of practices, as, for instance, interoperability, rebates or exclusive dealing contracts. A practice can be undertaken at different intensity through the choice of an action a, making the design of business strategies a matter of degree more than a yes/no decision. The set of actions is $A = [0, 1]$, where the lower bound $a = 0$ can be interpreted as not undertaking the practice at all. For instance, if the practice refers to the choice of technical compatibility with the competitors' products, the actions will coincide with different levels of interoperability. If, instead, the practice corresponds to adopting quantity rebates, the action will be the level of discount or any other relevant parameter of the scheme. In the same vein, the practice of exclusive dealing requires to specify the subset of committed customers and the proposed compensation for exclusivity, which

[7] 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.

in this latter example will be the relevant actions.

When the dominant firm undertakes the practice, this latter affects profits and welfare according to the intensity measured by the action undertaken. More precisely, the practice and associated actions yield private profits $\Pi(a) = \pi a \geqslant 0$ to the dominant firm, with the profits when the practice is not adopted ($a = 0$) normalized to zero and corresponding to the returns from "business as usual".

While the private effects of the practice are positive, the social impact is *ex-ante* uncertain. Depending on the state of nature s, the practice may increase (efficient) or decrease (anticompetitive) welfare, with an impact that depends on the action undertaken. With *ex-ante* probability β, that is common knowledge, social welfare is reduced compared to the benchmark level. We denote this case as the bad state $s = b$ and the associated welfare as $W^b(a) = -w^b a \leqslant 0$, with $w^b > 0$. In the bad state, private incentives conflict with social welfare, that is, when the incumbent increases the intensity of the practice, social welfare falls. Under certain market conditions and for given competitors' set of products, for instance, limiting interoperability restricts the rivals' ability to compete, with a stronger effect the less compatible are the products. Equivalently, when adopting rebates, more and more efficient competitors are forced to exit the larger the discount or the target quantity.

With probability $1 - \beta$, instead, a good state $s = g$ materializes: when the dominant firm undertakes the practice through actions $a \in A$, welfare increases according to the function $W^g(a) = w^g a \geqslant 0$ with $w^g > 0$. In this case, there is no conflict between private and social incentives since the practice increases both the profits of the firm and social welfare. Examples are when in the market there are alternative bundling opportunities for the competitors and limited compatibility does not reduce competition, while allowing a better match and functioning of the firm's products, or when rebates and exclusive dealing do not limit the ability of strong rivals to compete but create incentives to relationship specific investments with the clients.[8]

We can explain our modeling of uncertainty with the following argument. The way a practice and its implementation affect private profits and social welfare depends on

[8] In the benchmark model we assume that the new technology produces private and social effects only if combined with the practice, while profits and welfare do not change with respect to the competitive scenario if the practice is not adopted ($a = 0$). We choose this modeling strategy to focus on the impact of antitrust intervention (that affects the adoption of the practice), on the incentives to invest in research. In Section 4.2 we extend the model by considering a positive fixed effect of the new technology on profits and welfare, that adds to the effect of the practice described in the benchmark model.

the occurrence of a set of circumstances (market structure, conditions of entry, products offered by the competitors, etc.). This set of factual elements makes foreclosure the equilibrium of the market game or alternatively an unfeasible outcome. These features, indeed, are not yet entirely realized and cannot be observed *ex-ante*, at the time the policy is designed and the investment is sunk, but can only be described in probability terms. The *ex-ante* probability β captures the enforcer's priors that a practice leads to foreclosure. The economic model implicitly adopted by the enforcer when considering a certain practice and its implementation through the actions, what we can consider as her presumptions, is summarized in the terms $\{w^g, w^b, \pi, \beta\}$. In the remaining part of the paper we show that the optimal legal standards and enforcement policies for a certain practice depend, given the feasible policy instruments, on these parameters of the enforcer's economic model.

We impose the following restrictions on the parameters:

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

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

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

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

Public policies: legal standards, fines and accuracy. The enforcer has to design the public policies to contain the potential hazards posed by a certain practice and needs to collect information, according to the legal standards in place, to properly implement the enforcement policy. Each legal standard adopts a specific definition of what (if any) is unlawful, and requires to specify a minimum amount of evidence to convict the firm. A richer definition of unlawfulness in general requires a more complex set of information, that is more costly to collect and may lead more frequently to errors.

Following this approach, we assume that the enforcer perfectly recognizes the action chosen by the firm, i.e. any $a \in A$. Yet, the information regarding the effects of the practice 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. 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.[9] 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 A as socially damaging, committing a type-II error. Hence,

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

We assume that the signals are informative, i.e. $\varepsilon^i \leqslant \overline{\varepsilon} < \frac{1}{2}$, $i = I, II$. The enforcer can affect the level of type-i error by committing resources to refine the assessment of the effects, what is usually called accuracy[10]. In other words, the enforcer can collect additional evidence that better allows estimating, directly or indirectly, whether the practice increases or reduces welfare. We assume that the cost of reducing a type-i error, is increasing and convex, and that if no resources are devoted to this goal the error committed is equal to $\overline{\varepsilon}$.[11] More precisely, the cost of implementing an error probability ε^i is

$$g(\varepsilon^i) = \frac{\gamma}{2}(\overline{\varepsilon} - \varepsilon^i)^2.$$

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

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

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

It should be stressed that per-se legality and per-se illegality differ in the power of the enforcer to fine the firm when the practice is undertaken, and not in the fact that the practice is adopted or not in equilibrium. Indeed, we shall see that even under

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

[10] By collecting additional evidence, the enforcer is able to reduce the variance of the conditional distribution of the signal.

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

per-se illegality it may be optimal to have the firm undertaking the practice at some degree (and pay a positive fine).

Alternatively, the enforcer can adopt a discriminating legal standard (or effect-based rule) that does not consider only the actions but also their social consequences:

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

From the assumptions above, the choice of accuracy is an issue only under a discriminating rule, since 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.

Besides the level of type-I and type-II errors, the enforcer controls a third policy variable: a non decreasing fine schedule $f(a) \in [0, F]$. Since the profit function $\Pi(a)$ is increasing and linear in a, we can use with no loss of generality, within the set of non-decreasing fine schedules, the stepwise function

$$f(a) = \begin{cases} 0 & \text{if } a = 0 \\ \underline{f} \geqslant 0 & \text{if } 0 < a \leqslant \widetilde{a} \\ \overline{f} \leqslant F & \text{if } a > \widetilde{a}. \end{cases} \tag{3}$$

Notice that, under any rule, the enforcer cannot fine a firm when it does not undertake the practice ($a = 0$). In the benchmark model we do not set any upper limit on the fine that can be levied.[12]

Timing. The timing of the model is as follows. At time 0 nature chooses the state of the world $s = \{g, b\}$. At time 1, the enforcer commits to a certain legal standard $i \in \{L, IL, D\}$ and sets the fine schedule $f(a)$ and the level of the errors 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

[12]In Section 4.3 we extend the analysis to the case of caps on fines, considering the limited liability constraint ($F = \pi$).

welfare; the enforcer receives a signal σ 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.

3 Optimal legal standards and enforcement policies

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

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

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

The research investment I_{FB} is increasing in the likelihood of the good state $(1-\beta)$ and in the welfare gain w^g. Since the practice is undertaken only when it is welfare improving, discovering the new technology has a positive expected impact on welfare, and the level of investment is always positive, although decreasing in the probability of social harm due to the cost of research. The expected welfare, evaluated at the first-best policies, is

$$EW_{FB}(\beta) = \frac{[(1-\beta)w^g]^2}{2}.$$

In what follows, the policy maker is assumed not to control firm's choices directly, but to influence them via penalties: firms are free to undertake their preferred action and investment, but they are aware that public intervention may occur *ex post* in the form of fines, whenever they can be levied according to the legal standard in place. We start with per-se rules, identifying the optimal enforcement in this setting, and then move to discriminating rules and the associated optimal enforcement policy. Finally, we compare the two legal standards, evaluated at the corresponding optimal enforcement policy, and select, for different values of the prior on social harm β, the overall optimal solution.

3.1 Per-se rules

The very nature of per-se rules is to treat the practice and any associated action $a \in A$ as legal (L-rule) or unlawful (IL-rule) irrespective of the signal (effects) σ received. We analyze the optimal enforcement starting from stage 3, when the firm chooses the action, that is the level of intensity of the practice. If the research has been successful, the firm acquires a dominant position and is subject to antitrust scrutiny when choosing among the actions in A. Since the practice is equally profitable in both states of the world and per-se rules do not link the fine to the effects of the practice itself, the incumbent undertakes the same profit maximizing action, no matter if it is welfare enhancing or socially harmful. Which specific action, however, depends on the fine schedule $f(a)$ designed by the enforcer. By appropriately choosing, according to (3), the threshold level $\widetilde{a}$ and the level of fines for actions above ($\overline{f}$) and below the threshold ($\underline{f}$), the enforcer can lead the firm to choose $\widetilde{a}$. The following incentive compatibility constraint (ICC) ensures that $\widetilde{a}$ is the most profitable way (action) to implement the practice

$$\pi\widetilde{a} - \underline{f} \geqslant \pi - \overline{f}. \tag{5}$$

The undertake constraint (UC), instead, ensure that the firm (weakly) prefers to adopt the practice ($a > 0$) rather than keeping on with "business as usual" ($a = 0$), and it is relevant as long as $\widetilde{a} > 0$

$$\pi\widetilde{a} - \underline{f} \geqslant 0. \tag{6}$$

It follows that the design of the optimal fine schedule is equivalent to (indirectly) implementing a (profit-maximizing) action that maximizes welfare – which we denote $\widehat{a}$ – that is an action that the firm is willing to choose according to the incentive compatibility and undertake constraints, and that is socially optimal.

If research instead fails to produce a result, the market environment is competitive, the firm lacks any market power and is not subject to antitrust intervention, getting the competitive profits equal to zero. Then, the expected profits under per-se rules (subscript PS) are $E\Pi_{PS} = I(\pi\widetilde{a} - \underline{f}) - \frac{I^2}{2}$ and the profit maximizing investment is

$$I_{PS} = \pi\widetilde{a} - \underline{f}. \tag{7}$$

The expected welfare under per-se rules is therefore

$$EW_{PS}(\beta) = I_{PS}\left[(1-\beta)w^g - \beta w^b\right]\widetilde{a} - \frac{I_{PS}^2}{2} = I_{PS}Ew(\beta)\widetilde{a} - \frac{I_{PS}^2}{2}, \tag{8}$$

where $Ew(\beta)$ is the expected marginal welfare of an increase in the intensity of the practice. The enforcer maximizes the expected welfare setting $\widetilde{a}$, $\underline{f}$ and $\overline{f}$, subject to (5), (6) and (7).

Notice that in our setting deterrence takes two forms: *ex-post* deterrence, when the action is chosen after an innovation (marginal deterrence);[13] and *ex-ante* deterrence, since enforcement, by affecting *ex-post* the profits from the practice, influences the *ex-ante* incentives to invest in research.

In the following lemma we derive whether it is optimal to apply a per-se legality or a per-se illegality legal standard. If it is optimal not to fine the practice at any degree $a \in A$, the corresponding legal standard is per-se legality, while if a positive fine is levied at least on some actions $a \in A$, then the enforcer is applying a per-se illegality rule.

Lemma 1 (Optimal enforcement policy under per-se rules) *The optimal legal standard and enforcement policy under per-se rules are:*

(i) for $\beta \in [0, \beta_1]$, *where*

$$\beta_1 = \frac{w^g - \pi}{w^g + w^b}, \tag{9}$$

the optimal legal standard is per-se legality and the optimal enforcement implements $a^g = a^b = 1$ *and* $I = \pi$, *by setting* $\widehat{a} = 1$, $\underline{f} = 0$. *The expected welfare is* $EW_{PS}(\beta) = \pi\left[Ew(\beta) - \frac{\pi}{2}\right]$ *and is decreasing and linear in* β.

(ii) for $\beta \in (\beta_1, \beta_2)$, *where*

$$\beta_2 = \frac{w^g}{w^g + w^b}, \tag{10}$$

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

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

The optimal legal standard and enforcement policy vary with the likelihood of social harm. When the expected welfare of the practice is positive ($Ew(\beta) > 0$)

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

and social harm is unlikely ($\beta \leqslant \beta_1$), it is optimal to adopt a per-se legality regime by not fining the practice at whatever degree it is undertaken. When the likelihood of social harm becomes higher ($\beta > \beta_1$), the optimal legal standard turns to per-se illegality, allowing the enforcer to fine the practice. In a first range of values, that is for $\beta \in (\beta_1, \beta_2)$, the expected welfare of the practice is still positive and the enforcer uses the fine to reduce the level of (costly) investment but preserving the realization of the practice at the highest degree $\widehat{a} = 1$. When the social harm becomes sufficiently likely ($\beta > \beta_2$), the practice becomes socially harmful in expected terms and it is completely deterred, implementing $\widehat{a} = 0$ and eliminating any incentive to invest in research. Finally, the expected welfare in the three regions is continuous and decreasing in β.

3.2 Discriminating rules

A discriminating rule is based both on the observed actions and on the signal. It considers an action $a \in A$ as illegal if the enforcer receives a signal $\sigma > x$. Although the signal may be incorrect, we have assumed it to be informative. The enforcer, then, can indeed implement – in contrast with per-se rules – different actions in different states of the world. Since the discriminating legal standard does not allow the enforcer to levy any fine if the signal is $\sigma \leqslant x$, 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 constraint (ICC^b) is

$$\pi \widetilde{a}^b - (1 - \varepsilon^{II})\underline{f} \geqslant \pi - (1 - \varepsilon^{II})\overline{f}, \tag{11}$$

while the undertake constraint (UC^b) is

$$\pi \widetilde{a}^b - (1 - \varepsilon^{II})\underline{f} \geqslant 0. \tag{12}$$

Although the incentive compatibility constraint ICC^b to implement $\widetilde{a}^b$ puts only a lower bound on the maximum fine $\overline{f}$, when we turn to the good state, type-I errors are committed, and an excessively high $\overline{f}$ might induce the firm to undertake $a^g = \widetilde{a}^b$ rather than $a^g = 1$. [14] Hence, we have to further impose the following incentive

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

compatibility (ICC^g) and undertake (UC^g) constraints for the good state

$$\pi\widetilde{a}^b - \varepsilon^I \underline{f} \leqslant \pi - \varepsilon^I \overline{f} \tag{13}$$

and

$$\pi - \varepsilon^I \overline{f} \geqslant 0. \tag{14}$$

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

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

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

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

The innovative investment in the discriminating regime is

$$I_D = (1-\beta)\left[\pi - \varepsilon^I \overline{f}\right] + \beta\left[\pi\widetilde{a}^b - (1-\varepsilon^{II})\underline{f}\right] \geqslant 0. \tag{16}$$

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

The expected welfare under the discriminating rule is

$$EW_D = I\left[\Delta W_D - \frac{I_D}{2}\right] - \frac{\gamma}{2}(\overline{\varepsilon} - \varepsilon^I)^2 - \frac{\gamma}{2}(\overline{\varepsilon} - \varepsilon^{II})^2, \tag{17}$$

where $\Delta W_D = (1-\beta)w^g - \beta w^b \widetilde{a}^b$. The optimal policy requires therefore to set the fine schedule ($\underline{f}$, $\overline{f}$, $\widetilde{a}^b$) and the errors ε^I and ε^{II} to maximize the expected welfare under the above constraints. As before, we denote as $\widehat{a}^b$ the action that solves this program (in the bad state). In the following lemma we identify the optimal enforcement policy.

Lemma 2 (**Optimal enforcement policy under discriminating rules**) *The optimal enforcement policy under the discriminating regime is:*
(i) for $\beta \in [0, \beta_0]$, *where*

$$\beta_0 = \frac{w^g - w^b - \pi}{w^g + w^b},$$

the optimal policy implements $a^g = a^b = 1$ *and* $I = \pi$ *by setting* $\widehat{a}^b = 1$, $\underline{f} = 0$ *and the minimum level of accuracy* ($\varepsilon^I = \varepsilon^{II} = \overline{\varepsilon}$). *The optimal policy makes the*

discriminating regime equivalent to a per-se legality rule. The expected welfare is $EW_D(\beta) = \pi\left[Ew(\beta) - \frac{\pi}{2}\right]$ *and is decreasing and linear in* β*.*
ii) for $\beta \in (\beta_0, 1]$ *if* γ *is sufficiently large the optimal policy implements the actions* $a^b < 1$, $a^g = 1$ *and investment* $I < \pi$ *by improving type-I accuracy* ($\varepsilon^I < \overline{\varepsilon}$, $\varepsilon^{II} = \overline{\varepsilon}$) *and by setting* $\widehat{a}^b < 1$, $\underline{f} = 0$, *and* $\overline{f} = \frac{\pi(1-\widehat{a}^b)}{(1-\overline{\varepsilon})}$.

Moreover, $\widehat{a}^b$ *is decreasing in* β *with* $\widehat{a}^b \to 1$ *for* $\beta \to \beta_0$ *and* $\widehat{a}^b \to 0$ *for* $\beta \to 1$*. Finally, the expected welfare* $EW_D(\beta)$ *is decreasing in* β and *tends to* 0 *when* $\beta \to 1$*.*

When social harm is unlikely, the discriminating rule implements an outcome equivalent to a per-se legality rule. Notice that this occurs in an interval $[0, \beta_0]$ in which also the per-se rule opted for generalized acquittal, since $\beta_0 < \beta_1$. Above this interval, the discriminating rule allows the enforcer to implement different actions in the two states, the welfare maximizing action $a = 1$ in the good state and an action $\widehat{a}^b \in (0, 1)$ in the bad state. Even if very high fines are feasible, the enforcer implements a positive practice $\widehat{a}^b$ (which implies some social damage *ex-post*) to sustain profits, softening this way *ex-ante* deterrence on innovative activity, and turns to $\widehat{a}^b = 0$ only when β tends to 1. The level of investment under the discriminating rule is below the first best, and is affected by $\overline{f}$ and ε^I as shown in (16). Moreover, the optimal policy commands a reduction in type-I errors, softening over-deterrence and boosting the innovative investment. Indeed, this goal cannot be pursued only through a reduction in the fine $\overline{f}$ 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.

3.3 Optimal legal standards

We are now equipped to find the optimal regime, by comparing the expected welfare associated with the optimal enforcement of per-se and discriminating rules. The following proposition establishes the result.

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

The choice of the legal standard depends on the ability of the regime to ensure both *ex-post* deterrence, that is the ability to make the firm undertake the practice at the welfare maximizing level, and *ex-ante* deterrence, involving the desired level of investment in research. These effects may conflict: *ex-post* deterrence requires to discourage the practice whenever it is socially harmful, and a discriminating rule is more flexible and effective under this concern, as it allows calibrating the fine to the *ex-post* effects. *Ex-ante* deterrence, instead, requires to discourage the investment only if it is expected to reduce welfare. In other words, *ex-ante* a practice may be socially beneficial even if *ex-post* it may reduce welfare in certain circumstances. In this case, a rigid rule (per-se legality) may dominate a flexible one (discriminating) for its ability to commit not to intervene *ex-post* on the practice when socially harmful, boosting the research investment at most. 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.[15]

4 Extensions

In this section we consider three extensions of the benchmark model. First, we compare the results for innovative industries with the corresponding solution for traditional ones where the incumbent is already established and the technology is stable. Comparing these two environments, we study whether the use of richer, effect-based legal standards is more or less frequent when the innovative activity is central.

Secondly, we extend the model to include a positive constant effect of the new technology on profits and welfare (even if the practice is not adopted). In this framework, a positive benchmark level of profits may be thought as the result of patent protection, while additional profits can be obtained through the practice and are affected by the antitrust policy. This setting is therefore ideal to study if patent policy and antitrust intervention play a complementary role or act as substitutes in the policy design.

Finally, we ask how the choice of the optimal legal regime is affected by a cap on fines, in the form of a limited liability constraint.

[15] The role of commitment and flexibility of a legal systems in affecting growth has been recently studied by Anderlini et.al. (2011).

4.1 Legal standards in traditional industries

In traditional industries the incumbent derives its market power from a previous evolution of the market and adopts a well known and stable technology. We change the baseline model assuming that the firm can undertake the practice at any degree $a \in A$ with no need to innovate. Moreover, to ease the comparison we assume that private ($\Pi(a)$) and social ($W^g(a)$ and $W^b(a)$) effects of the practice are the same as in the benchmark. Legal standards and enforcement tools are unchanged as well as the timing, with the exception of stage 2, since no research investment is needed in traditional industries.

The efficient course of actions is still $a^b = 0$ and $a^g = 1$ with the first best expected welfare equal to $EW_{FB}(\beta) = (1-\beta)w^g > 0$.

Turning to the per-se rules, among all the implementable actions $\widetilde{a}$ the optimal policy selects $\widehat{a}$, the one that maximizes, subject to (5) and (6), the expected welfare

$$EW_{PS}(\beta) = \left[(1-\beta)w^g - \beta w^b\right]\widetilde{a} = Ew(\beta)\widetilde{a}, \tag{18}$$

where the optimal policy depends only on the sign of $Ew(\beta)$. The result below easily follows:

Lemma 3 (Optimal policy under per-se rules) *When $Ew(\beta) \geqslant 0$, i.e. for $\beta \leqslant \frac{w^g}{w^g+w^b} = \beta_2$, the optimal legal standard is per-se legality and the optimal enforcement policy implements $\widehat{a} = 1$, by choosing $\underline{f} < \pi$* and *$\overline{f} \geqslant \underline{f}$. When $Ew(\beta) < 0$, or* for *$\beta \in (\beta_2, 1]$, the optimal legal standard is per-se illegality and the optimal enforcement policy implements $\widehat{a} = 0$, by setting $\overline{f} \geqslant \pi$ and $\underline{f} = 0$.*

As for the case of innovative industries, when the social harm is unlikely per-se rules dictate a per-se legality regime. In traditional industries, however, the threshold for per-se legality is higher compared to innovative environments ($\beta_1 < \beta_2$), as can be appreciated comparing Lemma 1 and Lemma 3.

Turning to the discriminating rule the choice of the action in the traditional environment parallels the innovative industry case. Once again the enforcer can implement $a^b = \widetilde{a}^b$ and $a^g = 1$ satisfying incentive compatibility and undertake constraints.

The expected welfare can be written as

$$EW_D(\beta) = \left[(1-\beta)w^g - \beta w^b \widetilde{a}^b\right] - \frac{\gamma}{2}(\overline{\varepsilon} - \varepsilon^I)^2 - \frac{\gamma}{2}(\overline{\varepsilon} - \varepsilon^{II})^2. \tag{19}$$

Among the implementable actions identified by the incentive compatibility constraints for the bad state, $\widetilde{a}^b$, the enforcer chooses the one that maximizes welfare, that is $\widehat{a}^b$. Notice that the first best course of actions, $a^b = 0$ and $a^g = 1$, can be implemented under the discriminating regime by appropriately setting the fines $\overline{f}$ and $\underline{f}$ for given level of errors ε^I and ε^{II}. Moreover, since reducing errors is costly while fines are pure transfers, the first best course of actions can be implemented optimally by adopting the minimum level of accuracy, that is by setting $\varepsilon^I = \varepsilon^{II} = \overline{\varepsilon}$ and choosing the fines according to the incentive compatibility and undertake constraints.

The optimal legal standards in traditional industries are then summarized by the following proposition:

Proposition 2 (Optimal legal standards in traditional industries) *In traditional industries the discriminating rule always dominates the per-se rules and allows to replicate the first best allocation* $a^b = 0$ *and* $a^g = 1$ *by choosing the minimum level of accuracy* ($\varepsilon^I = \varepsilon^{II} = \overline{\varepsilon}$), $\underline{f} = 0$ *and any* $\overline{f} \in \left[\frac{\pi}{(1-\overline{\varepsilon})}, \frac{\pi}{\overline{\varepsilon}}\right]$.

In traditional industries, the only concern of law enforcement is *ex-post* deterrence, and treating different practices according to their effects is crucial. The discriminating rule, then, better fits this task, replicating the first best. This result contrasts with the optimal legal standard in innovative environments, where it is optimal, when the practice is *ex-ante* socially beneficial, to opt for a simpler per-se legality rule. Hence, interestingly, effect based rules should be adopted more frequently in traditional industries than in innovative environments, where per-se rules play an important role.

Indeed, the discriminating rule may have an undesirable feature in innovative industries, as the enforcer cannot be lenient when a bad signal is received. This 'rigidity' reduces the appeal of effect-based regimes when *ex-ante* deterrence must be limited at most to boost research investment. This feature, however, does not bite in traditional industries, where *ex-ante* deterrence is not an issue, explaining why discriminating rules always dominate per-se rules.

4.2 Antitrust policy v. patent policy

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

We impose the following restrictions on the parameters:[16]

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

Apart from the fixed effects, the model remains the same as in the benchmark case. Hence, we briefly sketch the differences in the analysis.[17]

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

$$I_{PS} = \Pi + \pi\widetilde{a} - \underline{f}, \tag{20}$$

and

$$EW_{PS}(\beta) = I_{PS}(W + Ew(\beta)\widetilde{a}) - \frac{I_{PS}^2}{2}. \tag{21}$$

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

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

while the expected welfare becomes

$$EW_D = I\left[W + \Delta W_D - \frac{I_D}{2}\right] - \frac{\gamma}{2}(\overline{\varepsilon} - \varepsilon^I)^2 - \frac{\gamma}{2}(\overline{\varepsilon} - \varepsilon^{II})^2. \tag{22}$$

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

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

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

Proposition 3 (Optimal legal standards with fixed effects of the innovation) *The optimal legal standard is a per-se legality rule for* $\beta \leqslant \beta_0^{'}$ *and a discriminating rule for higher* β, *where*

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

Proposition 3 shows that, qualitatively, the results on optimal legal standards are as in the benchmark model. Per-se legality initially dominates, and is then replaced, for higher β, by an effect-based rule.[18]

When the degree of patent protection is reduced, i.e. $W - \Pi$ is increased, the threshold $\beta_0^{'}$ shifts to the right and we observe an expansion of the region where antitrust policy opts for per-se legality, an extreme form of innovation-friendly antitrust intervention. In other words, patent and antitrust policies act as substitutes in our setting. This result is reported in the following corollary:

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

4.3 Limited fines and the cost of flexible rules

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

According to Lemma 2 and Proposition 1 the optimal enforcement for $\beta > \beta_0$ is a discriminating rule that progressively reduces the socially harmful practice $\widehat{a}^b$ and

[18] It is immediate to see that for $W = \Pi = 0$, the threshold and all the equilibrium expressions in Proposition 3 converge to the ones in Proposition 1.

increases the fine $\overline{f} = \frac{\pi(1-\widehat{a}^b)}{(1-\overline{\varepsilon})}$ as β increases. At the same time, type-I accuracy is improved to balance the negative effect of the increasing fine on the investment. 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 $\overline{f}$ 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, $\overline{f}$ cannot be set at at the level required to implement the action in the unconstrained solution. More precisely, there will exist a $\beta_3 > \beta_0$ such that $\overline{f} = \pi$ and the limited liability constraint starts binding. For $\beta > \beta_3$, $\widehat{a}^b$ becomes a function of the type-II error ε^{II}, as can be seen setting $\underline{f} = 0$ in the lower bound of (15) to get[19]

$$\widehat{a}^b = \varepsilon^{II}. \tag{23}$$

By reducing ε^{II} (collecting evidence on the variables that help better estimating 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 discriminating rule and limited liability.

Lemma 4 (Optimal enforcement policy under discriminating rule and limited liability) *Under a discriminating rule, there exists a $\beta_3 > \beta_0$ such that the limited liability constraint $\overline{f} \leqslant \pi$ does not bind for $\beta \in [0, \beta_3]$ when $\overline{f}$ is optimally set. In this interval the optimal policy is the one described in Lemma 2. Instead, for $\beta \in (\beta_3, 1]$ and γ sufficiently large the optimal policy entails more symmetric accuracies ($\varepsilon^I < \overline{\varepsilon}$ and $\varepsilon^{II} < \overline{\varepsilon}$). The actions implemented are $\widehat{a}^b = \varepsilon^{II}$ and $a^g = 1$. The expected welfare $EW_D(\beta)$ is decreasing in β and negative for $\beta \to 1$.*

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

In the following proposition we summarize the optimal legal standards.

Proposition 5 (Optimal legal standards under limited liability) *When fines are capped by limited liability, the optimal legal standard for increasing values of*

[19] The same qualitative argument applies for any $F \in \left(\pi, \frac{\pi}{(1-\overline{\varepsilon})}\right)$. When F is capped in the interval above, the implementable action in the bad state is $\widetilde{a}^b = 1 - (1 - \varepsilon^{II})\frac{F}{\pi}$.

β is: per-se legality ($\beta \in [0, \beta_0)$); the discriminating rule (with the limited liability constraint not binding) and with type-I accuracy ($\beta \in [\beta_0, \beta_3)$); then the discriminating rule with the limited liability constraint binding and more balanced accuracy on both errors ($\beta \in [\beta_3, \beta_4)$); and finally per-se illegality ($\beta \in [\beta_4, 1]$).

Up to the threshold β_3 the limited liability constraint does not bind, and the results correspond to the case in Proposition 1. When the likelihood of social harm increases above β_3, the limited liability constraint starts binding and the enforcer is less effective in affecting the firm's choices through fines. In this region, the dominant legal standard is still initially the discriminating rule realized, as already said, combining the maximum fine admitted with a reduction in both errors. When the social loss is very likely ($\beta > \beta_4$), the expected welfare becomes negative under a discriminating rule due to the high accuracy costs, and the more rigid per-se illegality rule replaces the discriminating rule, saving on accuracy cost although discouraging the practice in the (unlikely) good state.

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

The second reason rests on a cost saving argument: a discriminating rule better adapts to *ex-post* effects, but it requires more information and is therefore more prone to errors than a simpler, per-se rule. When fines are unlimited, this potential weakness plays a minor role, since fines act as substitutes to accuracy. When, however, fines are capped, the mix of policy instruments under a discriminating rule requires to further refine accuracy, making this regime more costly. When the practice is very likely to be harmful, then, a per-se illegality regime that completely deters it, destroying also the *ex-ante* incentives to invest, dominates a discriminating rule.[20]

[20]It is easy to show that the cost saving reason to opt for a per-se (illegality) rule applies also in traditional industries when β is sufficiently high and fines are bounded.

5 Conclusions

The optimal legal standards and enforcement policies in antitrust depend on the parameters that summarize the economic model, or the presumptions, of the enforcer. In this sense, legal standard, level of accuracy and fine schedule all depend on the priors of the enforcer regarding the economic effects of the practices. Under this respect, our results recall the debate briefly summarized in the introduction. Economic approaches that have stressed the efficiency enhancing effects of many business practices (a low 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 (a higher β), usually associated to the post-Chicago scholars, has represented the background for the effect-based approach to unilateral practices.

An interesting feature of our results refers to accuracy. We have seen that type-II accuracy can improve deterrence on actions, while the reduction of type-I error may sustain innovative investments. The possibility of refining type-I or type-II accuracy rests on the following argument. A practice may be welfare enhancing (good state) or detrimental (bad state). Each of the two possibilities can be analyzed within an appropriate model, and their empirical predictions suggest a set of observables. As long as the two sets of predictions are, at least in part, distinct, we can obtain identifying restrictions that allow to validate either of the two explanations. Then, the enforcer can collect a minimum of information – facing the default probabilities of errors ($\overline{\varepsilon}$) – 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.

To further illustrate with an example, let us consider Tirole (2005), who discusses at length the economic analysis of tying and its implications for antitrust, suggesting three possible explanations. Tying may be adopted for efficiency enhancing reasons, such as avoiding the costs of assembling complementary goods, ensuring their full compatibility, guaranteeing the quality of the components when quality is not observable, protecting the intellectual property of a main product by offering complementary goods that would require the disclosure of private and sensitive information to be produced independently. Alternatively, tying may be a tool to price discriminate, with ambiguous welfare effects. Or, finally, tying may be a foreclosure strategy by a dominant firm to monopolize a competitive market or to protect a monopolistic one. The enforcer, handling a case, has therefore a full set of factual elements to assess in order to evaluate whether the efficiency-enhancing story fits the data, making type-I error less likely, or the anticompetitive story is validated by the evidence, ensuring type-II accuracy (see Polo 2010 for an example of this identification strategy referred to selective price cuts).

Our analysis of the optimal enforcement policy has focussed on the choice of type-I and type-II accuracy, that can be chosen independently by the enforcer, while maintaining fixed the burden of proof (the threshold x of the signal σ). Kaplow (2011b), instead, investigates the optimal setting of the burden of proof. The main difference in the two approaches can be described as follows: if, for given accuracies, we change the threshold x, that is the minimum strength of evidence required to sanction a firm, we face a trade-off between a higher (lower) probability of type-I error and a lower (higher) probability of type-II errors. If, instead, type-I and/or type-II accuracies are improved, maintaining the threshold x constant, the enforcer can choose independently the probabilities of either error. In this latter case, then, the enforcer has more degrees of freedom in setting the optimal policy parameters. We leave for future research to investigate the choice of the burden of proof in innovative industries.

Appendix

Proof of Lemma 1. We solve our problem by omitting the undertake constraint (6) and verifying it *ex-post*. The maximization program is solved by the following first-order conditions

$$\begin{aligned}
\frac{\partial EW_{PS}}{\partial \widetilde{a}} &= [Ew(\beta)\widetilde{a} - I_{PS}]\,\pi + Ew(\beta)I_{PS} + \lambda \geqslant 0, \\
\frac{\partial EW_{PS}}{\partial \underline{f}} &= -\,[Ew(\beta)\widetilde{a} - I_{PS}] - \frac{\lambda}{\pi} \leqslant 0, \\
\frac{\partial EW_{PS}}{\partial \overline{f}} &= \frac{\lambda}{\pi} \geqslant 0,
\end{aligned} \tag{24}$$

Finally, the complementary slackness condition is

$$\lambda\left(\widetilde{a}-1+\frac{\overline{f}-\underline{f}}{\pi}\right)=0. \tag{25}$$

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

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

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

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

It is immediate to see that in all three cases the undertake constraint (6) is satisfied.

■

Proof of Lemma 2. The proof is organized as follows. First, we identify the equilibrium value of the policy variables; then we analyze the comparative statics of $\widehat{a}^b$ and EW_D with respect to β.

We solve our problem by omitting the incentive compatibility constraints (15) and the undertake constraints (12 and 14) and verifying them *ex-post*. The first order conditions are the following

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

where $\Delta W_D = (1-\beta)w^g - \beta w^b \widetilde{a}^b$ and I_D is given by (16).

Let us consider the following candidate solution and check in which interval of β it holds: $\underline{f} = \overline{f} = 0$ and $\widehat{a}^b = 1$. Substituting we have $I_D = \pi$ and $\Delta W_D - I_D = \left[w^g - \pi - \beta(w^g + w^b)\right] > 0$ for $\beta < \frac{w^g - \pi}{w^g + w^b} = \beta_1$. Moreover, for $\beta < \frac{w^g - w_b - \pi}{w^g + w^b} = \beta_0 < \frac{w^g - \pi}{w^g + w^b}$, $\frac{\partial EW_D}{\partial \widetilde{a}^b} = \beta\pi\left[w^g - w^b - \pi - \beta(w^g + w^b)\right] > 0$. Hence, for $\beta < \beta_0$, $\frac{\partial EW_D}{\partial \widetilde{a}^b} > 0$, $\frac{\partial EW_D}{\partial \underline{f}} < 0$ and $\frac{\partial EW_D}{\partial \overline{f}} < 0$ at $\varepsilon^I = \varepsilon^{II} = \overline{\varepsilon}$. Finally, the incentive compatibility constraints (15) and the undertake constraints (12 and 14) are clearly satisfied. The expected welfare is therefore $\pi\left[Ew(\beta) - \frac{\pi}{2}\right]$. Notice that this outcome is equivalent to the one under per-se legality.

Consider next the case $\beta > \beta_0$. We set $\widehat{a}^b < 1$ to obtain $\frac{\partial EW_D}{\partial \widetilde{a}^b} = 0$, implying that $\Delta W_D - I_D > 0$. Then $\frac{\partial EW_D}{\partial \underline{f}} < 0$ and we get $\underline{f} = 0$. Since $\underline{f} = 0$ we have $\frac{\partial EW_D}{\partial \varepsilon^{II}} = \gamma(\overline{\varepsilon} - \varepsilon^{II}) = 0$ at $\varepsilon^{II} = \overline{\varepsilon}$. Moreover, $\frac{\partial EW_D}{\partial \varepsilon^I} = 0$ for $\varepsilon^I < \overline{\varepsilon}$. Finally, $\frac{\partial EW_D}{\partial \overline{f}} < 0$ implies that $\overline{f}$ is determined by the lower bound of the constraint (15), that is, $\overline{f} = \frac{\pi(1-\widehat{a}^b)}{(1-\overline{\varepsilon})}$. Finally, notice that the undertake constraints are satisfied since $\pi\widehat{a}^b - 0 \geqslant 0$ (12) and $\pi - \varepsilon^I\overline{f} = \pi - \varepsilon^I\frac{\pi(1-\widehat{a}^b)}{(1-\overline{\varepsilon})} > \pi\left[\frac{1-\overline{\varepsilon}-\varepsilon^I}{1-\overline{\varepsilon}}\right] > 0$ (14).
To check the second order condition, notice that only $\widetilde{a}^b$ and ε^I are set at an internal solution. Hence,

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

for γ sufficiently large.

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

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

and

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

with $\widehat{a}^b \to 1$ and $I_D \to \pi$ for $\beta \to \beta_0$ and $\widehat{a}^b \to 0$ and $I_D \to 0$ for $\beta \to 1$. Therefore, the expected welfare tends to 0 when $\beta \to 1$. Notice that the expressions above 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 analyze the effect of β on the equilibrium value of $\widehat{a}^b$ we can differentiate the first order conditions with respect to $\widehat{a}^b$, ε^I and β. Then, we have that $sign\frac{d\widehat{a}^b}{d\beta} = sign(-\frac{\partial^2 EW_D}{\partial\beta\partial\widehat{a}^b}\frac{\partial^2 EW_D}{\partial\varepsilon^{I2}} + \frac{\partial^2 EW_D}{\partial\varepsilon^I\partial\widehat{a}^b}\frac{\partial^2 EW_D}{\partial\varepsilon^I\partial\beta})$ where $\frac{\partial^2 EW_D}{\partial\varepsilon^I\partial\widehat{a}^b} = \frac{\partial I_D}{\partial\varepsilon^I}\left[-\beta w^b - \frac{1}{2}\frac{\partial I_D}{\partial\widehat{a}^b}\right] + \frac{\partial^2 I_D}{\partial\varepsilon^I\partial\widehat{a}^b}\left[\Delta W_D - I\right] > 0$, since $\frac{\partial^2 I_D}{\partial\varepsilon^I\partial\widehat{a}^b} = \frac{(1-\beta)\pi}{(1-\overline{\varepsilon})} > 0$, $\frac{\partial I_D}{\partial\widehat{a}^b} > 0$ and $\frac{\partial I_D}{\partial\varepsilon^I} < 0$. $\frac{\partial^2 EW_D}{\partial\varepsilon^I\partial\beta} = \frac{\pi(1-\widehat{a}^b)}{(1-\overline{\varepsilon})}\left[\Delta W_D - I_D\right] - \frac{(1-\beta)\pi(1-\widehat{a}^b)}{(1-\overline{\varepsilon})}\left[-w^g - w^b\widehat{a}^b - \frac{1}{2}\frac{\partial I}{\partial\beta}\right] > 0$.

Finally,

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

Multiplying the previous expression by β we notice that

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

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

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

or equivalently as

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

since $\left(\frac{1-\overline{\varepsilon}-\varepsilon^I}{1-\overline{\varepsilon}}\right)$ 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.

Finally, differentiating with respect to β the expected welfare we get

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

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

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

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

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

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

Instead, the discriminating rule (Lemma 2) gives

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

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

For $\beta \in (\beta_1, \beta_2)$, per-se rule implements $a^g = a^b = 1$ and $I = Ew(\beta)$ by setting $\widehat{a} = 1$ and, $\underline{f} = [\pi - Ew(\beta)]$. The same allocation can be implemented also under a discriminating rule by setting $\varepsilon^I = \varepsilon^{II} = \bar{\varepsilon}$, $\widehat{a}^b = 1$ and $\underline{f} = [\pi - Ew(\beta)] \,/\, [(1-\beta)\bar{\varepsilon} + \beta(1-\bar{\varepsilon})]$, adjusting the fine with respect to the PS regime to take into account the errors. Although implementable, this allocation is not optimal under a discriminating rule, and therefore $EW_D(\beta) > EW_{PS}(\beta)$ in this interval. Finally, for $\beta \in [\beta_2, 1]$, $EW_{PS}(\beta) = 0$ while $EW_D(\beta)$ is decreasing and equal to zero only at $\beta = 1$. ■

Proof of Lemma 3. It is immediate to check that the $\widehat{a}=1$, any $\underline{f}<\pi$ and any $\overline{f}\geqslant\underline{f}$ satisfy both the incentive compatibility constraint (5) and the undertaking constraint (6). The same is true for $\widehat{a}=0$, $\overline{f}\geqslant\pi$ and $\underline{f}=0$. ■

Proof of Proposition 2. It is immediate to check that the first best allocation $a^b=0$ and $a^g=1$ and the optimal policy $\varepsilon^I=\varepsilon^{II}=\overline{\varepsilon}$, $\underline{f}=0$ and any $\overline{f}\in\left[\frac{\pi}{(1-\overline{\varepsilon})},\frac{\pi}{\overline{\varepsilon}}\right]$ satisfy both the incentive compatibility (15) the undertaking constraints (12) and (14). ■

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

Step 1. First of all, we find the **optimal policy under per-se rules**. While the logic of the proof is as the one of Lemma 1 we find one more region and the thresholds are different. Therefore, we find it useful to go through it. Recall from the text the expressions for the optimal investment $I_{PS}=\Pi+\pi\widetilde{a}-\underline{f}$ 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

$$\begin{aligned}\frac{\partial EW_{PS}}{\partial\widetilde{a}} &= [W+Ew(\beta)\widetilde{a}-I_{PS}]\,\pi+Ew(\beta)I_{PS}+\lambda\geqslant 0,\\ \frac{\partial EW_{PS}}{\partial\underline{f}} &= -[W+Ew(\beta)\widetilde{a}-I_{PS}]-\frac{\lambda}{\pi}\leqslant 0,\\ \frac{\partial EW_{PS}}{\partial\overline{f}} &= \frac{\lambda}{\pi}\geqslant 0,\end{aligned}\tag{26}$$

Finally, the complementary slackness condition is

$$\lambda\left(\widetilde{a}-1+\frac{\overline{f}-\underline{f}}{\pi}\right)=0.\tag{27}$$

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

(i) For $\beta\in\left[0,\beta_1^{'}\right]$ we have $W+Ew(\beta)-\Pi-\pi\geqslant 0$ where $\beta_1^{'}=\frac{w^g-\pi+W-\Pi}{w^g+w^b}$. Moreover, $W<\Pi+\pi$ implies that $\beta_1^{'}<\beta_2$ so that $Ew(\beta)\geqslant 0$ for $\beta\in\left[0,\beta_1^{'}\right]$.

Then, if we set $\underline{f} = 0$ and $\widehat{a} = 1$, the investment is $I_{PS} = \Pi + \pi$ and, substituting in the first order conditions, we get $\frac{\partial EW_{PS}}{\partial \underline{f}} = -[W + Ew(\beta) - \Pi - \pi] < 0$ and $\frac{\partial EW_{PS}}{\partial \widetilde{a}} = [W + Ew(\beta) - \Pi - \pi]\,\pi + Ew(\beta)\Pi > 0$ then setting $\underline{f} = 0$ and $\widehat{a} = 1$ is optimal. Hence, $\overline{f}$ is not needed to define the fine schedule. Finally, $EW_{PS} = (\Pi + \pi)\left[W + Ew(\beta) - \frac{\Pi+\pi}{2}\right]$.

(ii) For $\beta \in \left(\beta_1^{'}, \beta_2\right)$, $Ew(\beta) > 0 > W + Ew(\beta) - \Pi - \pi$ and the first order condition $\partial EW_{PS}/\partial \underline{f} = 0$ holds for $W + Ew(\beta)\widetilde{a} - I_{PS} = 0$. Then $\frac{\partial EW_{PS}}{\partial \widetilde{a}} = Ew(\beta) I_{PS} > 0$ and $\widehat{a} = 1$. Substituting in $\partial EW_{PS}/\partial \underline{f} = 0$ and solving we get $\underline{f} = -W - Ew(\beta) + \Pi + \pi > 0$. Substituting $\underline{f}$ in the expression of the optimal investment we obtain $I_{PS} = W + Ew(\beta) > 0$ that is decreasing in β and equal to W when $\beta = \beta_2$. Finally, $EW_{PS} = \frac{(W+Ew(\beta))^2}{2}$.

(iii) For $\beta \in [\beta_2, \beta_2^{'})$, $(W - \Pi)\pi + Ew(\beta)\Pi > 0 \geqslant Ew(\beta)$ where $\beta_2^{'} = \frac{w^g + (W-\Pi)\pi}{w^g + w^b}$ and $\beta_2^{'} > \beta_1^{'}$ (using both $W < \Pi + \pi$ and $W > \Pi$). Then, if we set $\frac{\partial EW_{PS}}{\partial \widetilde{a}} = [W + Ew(\beta)\widetilde{a} - \Pi - \pi\widetilde{a}]\,\pi + Ew(\beta)(\Pi + \pi\widetilde{a}) = 0$ we have that $\underline{f} = 0$ since $\frac{\partial EW_{PS}}{\partial \underline{f}} < 0$, $\widehat{a}$ is interior and equal to $\frac{W-\Pi}{\pi - 2Ew(\beta)} + \frac{Ew(\beta)\Pi}{(\pi - 2Ew(\beta))\pi}$ and $I_{PS} = \frac{\pi W - Ew(\beta)\Pi}{\pi - 2Ew(\beta)}$. Finally, $EW_{PS} = I_{PS}\left[W + Ew(\beta)\widehat{a} - \frac{I_{PS}}{2}\right]$. Substituting $\widehat{a}$ and I_{PS} and rearranging we get $EW_{PS} = \frac{[rW - Ew(\beta)]^2}{2\pi[\pi - 2Ew(\beta)]}$. Differentiating w.r.t. β we get:

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

Hence, in this region EW_{PS} is decreasing and convex in β.

(iv) For $\beta \in \left[\beta_2^{'}, 1\right]$, $0 \geqslant (W - \Pi)\pi + Ew(\beta)\Pi > Ew(\beta)$ implying $\partial EW_{PS}/\partial \widetilde{a} < 0$ and $\partial EW_{PS}/\partial \underline{f} < 0$. So that $\widehat{a} = 0$, $\underline{f} = 0$ 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 $\overline{f} \geqslant \pi$. It is immediate to see that in all cases the undertake constraint (6) is satisfied.

Step 2. Second, we find the **optimal policy under discriminating rules**. Recall from the text the expressions for the innovative investment $I_D = \Pi + (1 - \beta)\left[\pi - \varepsilon^I \overline{f}\right] + \beta\left[\pi\widetilde{a}^b - (1 - \varepsilon^{II})\underline{f}\right]$ and for the expected welfare $EW_D = I\left[W + \Delta W_D - \frac{I_D}{2}\right] - \frac{\gamma}{2}(\overline{\varepsilon} - \varepsilon^I)^2 - \frac{\gamma}{2}(\overline{\varepsilon} - \varepsilon^{II})^2$. The proof follows closely the one in Lemma 2. We only underline three differences: first, going through the same steps in Lemma 2 from $\frac{\partial EW_D}{\partial \widetilde{a}^b} = 0$ we find a new threshold $\beta_0^{'} = \frac{w^g - w^b - \pi - w^b \frac{\Pi}{\pi} + (W - \Pi)}{w^g + w^b}$ (instead of β_0) such that for $\beta \leqslant \beta_0^{'} : \widehat{a}^b = 1$, $\overline{f} = \underline{f} = 0$, $I_D = \Pi + \pi$ and the expected welfare is $(\Pi + \pi)\left[W + Ew(\beta) - \frac{\Pi+\pi}{2}\right]$. Notice that once again this outcome is equivalent to the one under per-se legality.

Second, differently from Lemma 2, showing that EW_D is decreasing in β is not enough to completely characterize the optimal legal standard. Indeed, we also need to show that for $\beta > \beta_0^{'}$ the expected welfare is concave in β.

Therefore, differentiating two times with respect to β the expected welfare we get

$$\frac{d^2 EW_D}{d\beta^2} = \frac{\partial^2 EW_D}{\partial\beta^2} + \frac{\partial^2 EW_D}{\partial\varepsilon^{I2}}\left(\frac{\partial\varepsilon^I}{\partial\beta}\right)^2 + \frac{\partial^2 EW_D}{\partial\widehat{a}^{b2}}\left(\frac{\partial\widehat{a}^b}{\partial\beta}\right)^2 + 2\frac{\partial^2 EW_D}{\partial\varepsilon^I\partial\widehat{a}^b}\left(\frac{\partial\varepsilon^I}{\partial\beta}\frac{\partial\widehat{a}^b}{\partial\beta}\right),$$

where $\frac{\partial^2 EW_D}{\partial\beta^2} = 2\frac{\partial I_D}{\partial\beta}\left[-w^g - w^b\widetilde{a}_D - \frac{1}{2}\frac{\partial I_D}{\partial\beta}\right] > 0$, $\frac{\partial^2 EW_D}{\partial\varepsilon^{I2}} < 0$, $\frac{\partial^2 EW_D}{\partial\widehat{a}^{b2}} < 0$ while $\frac{\partial^2 EW_D}{\partial\varepsilon^I\partial\widehat{a}^b}\left(\frac{\partial\varepsilon^I}{\partial\beta}\frac{\partial\widehat{a}^b}{\partial\beta}\right)$ is ambiguous in sign. The expected welfare is then always decreasing in β and concave when γ (that is in the expression for $\frac{\partial^2 EW_D}{\partial\varepsilon^{I2}}$) is sufficiently large.

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

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

Proof of Lemma 4. Combining the incentive compatibility and limited liability constraints by setting $\overline{f} = \pi$ and $\underline{f} = 0$ in (15) we obtain

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

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

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

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

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

that hold as equalities with internal solutions $\varepsilon^I < \overline{\varepsilon}$ and $\varepsilon^{II} < \overline{\varepsilon}$. Notice that for $\overline{f} = \pi$, $\underline{f} = 0$, $\widehat{a}^b = \varepsilon^{II}$ and $a^g = 1$ the undertake constraints (12 and 14) are also satisfied.

Finally, the second order conditions hold, since

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

Differentiating with respect to β the expected welfare we get

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

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

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

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

Bibliography

Anderlini, L., Felli L., Immordino, G. and A. Riboni, (2011), “Legal institutions, innovation and growth,” CEPR Discussion Papers 8433.

Arrow, K.J., (1962), “Economic Welfare and the Allocation of Resources for invention,” in: Nelson, R.R. (Ed.), The Rate and Direction of Inventive Activity: Economic and Social Factors, Princeton University Press, Princeton, pp. 609-626.

Dasgupta, P., and Stiglitz, J., (1980), “Uncertainty, Industrial Structure, and the Speed of R&D,” *The Bell Journal of Economics* **11**, 1-28.

Demougin D. and C. Fluet, (2008), “Rules of Proof, Courts and Incentives,” *Rand Journal of Economics*, **39**, 20-40.

Department of Justice (2008), “Competition and Monopoly: Single-Firm Conduct under Section 2 of the Sherman Act.”

DG Competition (2005), “Discussion Paper on the Application of Article 82 of the Treaty to Exclusionary Abuses.”

DG Competition (2008), “Draft Guidance on the Commission’s Enforcement Priorities in Applying Article 82 EC Treaty to Abusive Exclusionary Conducts by Dominant Undertakings.”

Easterbrook F., (1984), “The Limits of Antitrust,” *Texas Law Review*, **63**, 9-14.

David E. and J. Padilla, (2005), “Designing Antitrust Rules for Assessing Unilateral Practices: a Neo.Chicago Approach,” *The University of Chicago Law Review*, **72**, 73-98.

Gual, J., Hellwig, M., Perrot, A., Polo, M., Rey, P., Schmidt, K. and R. Stenbacka, (2005), “An Economic Approach to Article 82,” Report for the DG Competition, European Commission, published on *Competition Policy International*, **2**, 111-156.

Immordino, G., Pagano, M. and M. Polo, (2011), “Incentives to Innovate and Social Harm: Laissez-faire, Authorization or Penalties?,” *Journal of Public Economics,* **95**, 864-876.

Immordino, G. and M. Polo, (2008), “Judicial Errors and Innovative Activity,” CSEF working paper n. 196.

Kaplow, L., (1994), "The Value of Accuracy in Adjudication: an Economic Analysis," *Journal of Legal Studies*, **15**, 371-385.

Kaplow, L., (1995), "A Model of the Optimal Complexity of Legal Rules," *The Journal of Law, Economics and Organization*, **11**, 150-163.

Kaplow, L., (2011a), "Optimal Proof Burdens, Deterrence and the Chilling of Desirable Behavior," *American Economic Review: Papers & Proceedings,* **101**, 277–80.

Kaplow, L., (2011b), "On the Optimal Burden of Proof," *Journal of Political Economy*, **119**, 1104-1140.

Kaplow, L., (2012), "Burden of Proof," *Yale Law Journal,* **121**, 738–859.

Kaplow, L. and S. Shavell, (1994), "Accuracy in the Determination of Liability," *Journal of Law and Economics*, **37**, 1-15.

Kaplow, L. and S. Shavell, (1996), "Accuracy in the Assessment of Damages," *Journal of Law and Economics*, **39**, 191-210.

Yannis, K. and D. Ulph, (2009), "On Optimal Legal Standards for Competition Policy: a General Welfare-Based Analysis," *Journal of Industrial Economics*, LVII: 410-437.

Lando, H. (2006), "Does wrongful conviction lower deterrence?" *Journal of Legal Studies,* **35**, 327-337.

Lear (2006), "The Cost of Inappropriate Intervention and Non Intervention Under Article 82," Report for the Office of Fair Trade, London, UK.

Lee, T., and Wilde, L.L., (1980), "Market Structure and Innovation: A Reformulation," *Quarterly Journal of Economics* **94**, 429-36.

Loury, G., (1979), "Market Structure and Innovation," *Quarterly Journal of Economics* **93**, 395-410.

Mookherjee, D. and I. Png, (1994), "Marginal Deterrence in the Enforcement of Law," *Journal of Political Economy* **102**, 1039-66.

Png I.P.L., (1986), "Optimal Subsidies and Damages in the Presence of Judicial Error," *International Review of Law and Economics*, **6**, 101-105.

Polinsky, M.A. and S. Shavell, (2000), "The Economic Theory of Public Enforcement of Law," *Journal of Economic Literature*, **38**, 45-76.

Polo M., (2010), "Anticompetitive vs Competitive Explanations of Unilateral Pratices: the Identification Problem," *Journal of Competition Law and Economics*, **6**, 457-476.

Posner, R.A., (1992), "*The Economic Analysis of Law*," 4th Edition. Boston: Little Brown.

Segal I., and M. Whinston, (2007), "Antitrust in Innovative Industries," *American Economic Review*, **97**, 1703-1730.

Shavell, S., (1980), "Strict Liability versus Negligence," *The Journal of Legal Studies*, **9**, 1-25.

Shavell, S., (1992), "Liability and the Incentive to Obtain Information about Risk," *Journal of Legal Studies,* **21**, 259-270.

Shavell, S., (2007), "Do Excessive Legal Standards Discourage Desirable Activity?" *Economics Letters*, **95**, 394-397.

Stigler, G.J., (1970), "The Optimum Enforcement of Laws," *Journal of Political Economy*, **78**, 526-536.

Tirole J., (2005), "The Analysis of Tying Cases: a Primer," *Competition Policy International*, **1**, 1-25.