Institutional Members: CEPR, NBER and Università Bocconi

WORKING PAPER SERIES

Optimal Legal Standards in Antitrust: Traditional v. Innovative Industries

Giovanni Immordino, Michele Polo

Working Paper n. 420

This Version: September 9, 2011

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

Optimal Legal Standards in Antitrust: Traditional v. Innovative Industries

Giovanni Immordino
Università di Salerno and CSEF

Michele Polo
Università Bocconi and IGIER

September 9, 2011

Abstract: A dominant firm undertakes a given business practice that is regulated by an antitrust enforcer by the choice of a legal standard, fines and accuracy. In traditional industries the incumbent and technology are already established, while in innovative industries the successful innovator becomes dominant. In the former case, marginal deterrence is key to enforcement, and discriminating rules are always dominant when fines are unbounded, or they are replaced with per-se illegality when fines are capped and the practice is likely to be socially harmful. In innovative industries marginal deterrence interacts with average deterrence (the impact of enforcement on innovation effort). Then, per-se legality is preferred when the practice is likely to be welfare beneficial, moving to a discriminating rule when social harm becomes more likely. When fines are capped, per se-legality, discriminating rule and per-se illegality are alternatively chosen when the practice is more and more likely to be socially harmful.

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 Nuno Garoupa, Louis Kaplow, Yannis Katsoulakos, Dilip Mookherjee, Massimo Motta, 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 Mannheim 2011, Cresse 2011 and Otranto 2011 conferences. All usual disclaimers apply.

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 and concentration cases. The discussion has raised issues concerning the substantial arguments to use in cases as well as the legal standards that should be adopted. This paper analyzes optimal legal standards and enforcement policies in antitrust, comparing traditional and innovative market environments.

In the European Union, during the first years of the decade the Commission has adopted important reforms on cartel cases (article 101) and merger control. In 2009 the DG Competition 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") and 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. The previous case suggests a link between the legal standards and enforcement policies on one side and the a-priori view of the enforcers with respect to the relevance of anticompetitive and efficiency effects of certain practices on the other.

This brief summary highlights some major aspects of the recent debate. First, while some economists argue that dominant firms adopt socially harmful practices to maintain their market power, others consider this possibility skeptically, stressing instead the pursuit of superior efficiency as the driving force explaining the emergence of market leaders. 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

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

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 behaviour 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, ect. 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.

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 v. efficiency explanations of market practices by incumbent firms may be sometimes rooted also into different priors or pre-analitical presumptions, as the recent US discussion suggests.

Secondly, the debate between different schools has extended from the economic arguments that should be adopted in antitrust cases to the legal standards that the investigations should follow.[2] 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. These latter 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.

Third, legal standards vary in their flexibility as well as in the amount, complexity and costs of the information required to identify an unlawful conduct and in the likelihood of committing errors. Hence, the choice of the appropriate level of fines and accuracy may vary across legal standards. Per-se rules (per-se legality or per-se illegality) are based on a simpler definition of unlawfulness, referred only to the practice undertaken. Hence, although per-se rules cannot distinguish desirable from socially harmful actions, they require to collect

[2] 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".

a simpler set of information, they are less costly and hardly lead to errors. If errors are unlikely, also accuracy plays a minor role in the enforcement of per-se rules. Conversely, discriminating rules identify unlawfulness with social harm and allow the enforcer designing a policy contingent on welfare effects. However, they may lead more frequently to errors and therefore accuracy has a larger scope in the design of the optimal enforcement.

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 consider two different economic environments in which a firm can have market power. In traditional industries the incumbent is already established (derives its market power from the previous evolution of the market) and the technology is stable. In innovative industries, instead, an initial competitive situation is broken by a firm that introduces an innovation and, this way, becomes dominant. In this latter case, therefore, innovation and market power come together.

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.[3] These presumptions express the view of the enforcer on the expected effects of a certain business practice.

In this framework we analyze the optimal choice of the legal standards, the associated enforcement policies and the level of accuracy in traditional and innovative industries for unbounded and capped fines, deriving a full set of prescriptions. Indeed, we show how in different environments the optimal legal standard and enforcement policy vary when the enforcer's presumptions on the effects becomes more and more pessimistic. In the simplest setting of traditional industries and unlimited fines the discriminating rule is always dominant for any prior on social losses. When we depart from this benchmark, some room for per-se rule is re-established. When fines are capped in traditional industries, the enforcer chooses a discriminating rule when the practice is not likely to be harmful, reducing type-II errors to improve marginal deterrence, while she opts for per-se illegality when the likelihood of social harm is sufficiently high, controlling less accurately but more cheaply the actions of the incumbent. Conversely, in innovative industries a per-se legality rule replaces the discriminating legal standard for low probability of social harm, being more effective in boosting the innovative investment, while the enforcer adopts the discriminating rule and improves type-I accuracy to sustain investment when the practice is more likely to reduce welfare. Finally, when fines are capped in innovative industries per-se legality is adopted for low probability of social damages, then replaced by a discriminating rule with more and

[3] See Easterbrook (1984).

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 in industries. Immordino, Pagano and Polo (2011) propose an analytical framework similar to this paper, analyzing the choice between different policy regime, namely ex-post law enforcement and ex-ante authorization and focus the analysis on the case of innovative industries, identifying the conditions when each of the policies 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 in a simple setting, comparing per-se rules and discriminating (effect based) rules characterized by a higher probability of errors. 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 here call traditional industries, while the case of innovative industries and the impact of enforcement on innovation 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[4].

The remainder of this paper is organized as follows. Section 2 presents the model. Section 3 and 4 focus on antitrust intervention in traditional and innovative industries respectively. Section 5 offers some concluding comments. All the proofs are in the Appendix.

[4] Judicial errors and their reduction, i.e. accuracy, are a central concern in law enforcement: they have been analyzed in the standard model of law enforcement proposed by Kaplow (1994), Kaplow and Shavell (1994, 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

We model the interaction of an antitrust enforcer and a dominant firm that undertakes a certain business practice. As the recent debate has recognized, depending on the economic environment, the effects of the practice may be positive (efficiency enhancing) or negative (anticompetitive). We analyze two different situations. In traditional industries the firm is an already established incumbent that possibly undertakes the practice using a standard technology. Alternatively, we analyze innovative industries in which the market is initially competitive but market power and competitive advantages are created, and further exploited through the practice, if an innovation is realized. In this latter case, therefore, innovation and dominance come together, a feature that is often observed in high-tech industries characterized by a "winner-takes-all" type of competition. In our analysis the enforcer can adopt different legal standards and display appropriate enforcement policies to influence the behaviour of the firm in these environments.

Private choices: practice intensity and research investment. Following the legal framework of antitrust intervention, business conducts are classified in a set of practices. In our model we analyze how the firm undertakes the practice at different degrees of intensity through the choice of an action a. With no loss of generality, the set of actions is $A = [0, 1]$, with the lower bound $a = 0$ that can be interpreted as not undertaking the practice at all. For instance, if the practice is the setting of technical compatibility with the competitors' products, the action refers to different levels of interoperability. If instead the practice corresponds to adopting quantity rebates, the action is 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 compensation for exclusivity proposed, that play in this latter example the role of actions.

In our discussion we shall refer to a firm with market power as dominant or incumbent. 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 $\Pi(a) = \pi a$ to the firm. Hence, the profits when the practice is not adopted ($a = 0$) are normalized to zero and correspond to the returns from "business as usual".

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, 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. For certain market conditions and competitors set of products, for instance, the existence of switching costs and the like, 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 boundling opportunities for the competitors and limited compatibility does not reduce competition, while allowing a better match and functioning of the products of the firm, or when rebates or exclusive dealing do not limit the ability of strong rivals to compete but create incentives to relationship investment with the clients.

This assumption can be interpreted this way: the way a practice and its implementation affect private profits and social welfare depends on 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. The ex-ante probability β captures the enforcers priors that a practice leads to foreclosure. We assume β to be common knowledge. 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 $\left\{w^g, w^b, \pi, \beta\right\}$. 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 consider two different environments (industries): in the *traditional* one the incumbent has already established its market power and is developing its business strategies given the current technology (innovation is not part of the economic landscape). In this case, therefore, the choice of the incumbent when adopting the practice refers to the action $a \in A$, and the problem for the enforcer is to design the antitrust policy to influence how the incumbent undertakes the practice. The traditional environment refers to industries in which the technology is mature and stable and the established incumbent derives its market power from the past evolution of the industry. This case corresponds to the standard setting considered in the optimal law enforcement approach.

We compare this case with an *innovative* environment in which the market is initially competitive and characterized by fragmentation and symmetry among firms, none of which has market power. However, 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.[5] We maintain the same notation by assuming that the profits of the (newly established) dominant firm if the innovation is successful are, $\Pi(a) = \pi a$, while in case the research process fails the initial competitive situation keeps on, with profits normalized to zero. Analogously, in case of innovation the welfare effects of the practice undertaken by the dominant firm may be positive ($W^g(a) = w^g a$) or negative ($W^b(a) = -w^b a$), while welfare is normalized to zero if the research process fails.[6]

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 $p(I)$ of innovating 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. For simplicity we assume $c(I) = \frac{I^2}{2}$.

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, what is called the burden of proof.[7] A richer definition of unlawfulness in general requires a more complex set of information, that are more costly to collect and may lead more frequently to errors.

---

[5] We do not model competition in research and patent races, but rather adopt the approach first proposed by Arrow (1962) to study the incentives to invest in research. We further discuss this issue in the final comments.

[6] We are implicitly assuming that the new technology produces private and social effects only if combined with the practice. Alternatively, the new technology might create and extra-profit and a welfare gain per-se, and this benchmark level might be further affected by undertaking the practice. This case leads to similar prediction but increases the number of parameters of the model. Hence, in developing our analysis we prefer to opt for a simpler model, discussing these extensions in Section 5.

[7] On the burden of proof see, for instance, Kaplow (2011) and Demougin and Fluet (2008). In this paper we maintain, within each legal standard, the burden of proof fixed and instead in some cases we allow the enforcer improving the accuracy.

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 $\sigma = g, b$ on the state of the world, that is whether the incumbent's practice is welfare enhancing or decreasing. 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 case of 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,

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

We assume that the signals received 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. 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, $i = I, II$, is increasing and convex, and that if no resources are devoted to this goal the error committed is equal to $\overline{\varepsilon}$.[8] More precisely, for type-i error, $i = I, II$, 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 consider the effects of those actions. More precisely, 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. Indeed, we shall see that even under 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:

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

D *discriminating*: any action $a \in A$ is legal unless the enforcer receives a signal $\sigma = b$;

In applying a given legal standard the enforcer controls three policy variables, the level of type-I and type-II errors (accuracies) as well as the 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

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

In the following we shall consider two cases: unbounded maximum fines ($F >> \pi$) and fines bounded by limited liability ($F = \pi$).[9]

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, more rigid but less prone to errors, and discriminating rules, that combine more flexibility and a more challenging informational requirement.

In the next sections, both for the traditional and the innovative environment, we will first identify the optimal enforcement policies (accuracies and fine schedule) within a given legal standard, and then we will select the optimal legal standard comparing the outcome obtained under each of them when the enforcement policies are implemented optimally.

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 enforcement policy set by the enforcer, in the innovative environment the firm chooses the research investment I, innovates with probability $p(I) = I$ and in this case also learns the state of the world $s = b, g$. In the traditional environment, no investment is required for the already established incumbent. At time 3, the firm chooses an action, conditional on what it learnt in the previous stage. Finally, at time 4 the action undertaken determines the private profits and the social welfare; the enforcer receives a signal $\sigma = g, b$ 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.

We can now move to the analysis of the optimal legal standard and enforcement policies distinguishing traditional and innovative industries.

[9] Notice that, under any rule, the enforcer, contrary to a regulator, cannot fine a firm when it does not undertake the practice ($a = 0$).

3 Antitrust intervention in traditional industries

In traditional industries the incumbent derives its market power from the previous evolution of the market and develops its business strategies adopting a well known and stable technology. The relevant choice, from a private and public perspective, is therefore how the practice is implemented (if any), that is the choice of action $a \in A$. To evaluate the benefits of public intervention we start by identifying the first-best outcome (FB), which would obtain if the enforcer could control the firm's action choice directly.

Let us denote by a^s the action chosen in state $s = b, g$. The welfare maximizing choices are clearly $a^b = 0$ (do not undertake the practice when socially harmful) and $a^g = 1$ (implement the practice at the highest degree when welfare enhancing) with expected welfare equal to

$$EW_{FB}(\beta) = (1-\beta)w^g > 0.$$

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, 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.

3.1 Per-se rules

Turning to the second-best, let us consider, for a given fine schedule $f(a)$, the firm's choice when a per-se rule applies. 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. Hence, facing an enforcement policy that does not discriminate according to the effects, under per-se rules the incumbent undertakes the same 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:

$$\pi\widetilde{a} - \underline{f} \geqslant \pi - \overline{f}. \tag{4}$$

The undertake constraint (UC), instead, ensure that the firm (weakly) prefers to adopt the practice ($a > 0$) rather than keeping with "business as usual"[10] ($a = 0$), and is relevant as

[10] The undertake constraint corresponds to the participation constraint in standard mechanism design problems. In our problem we prefer to use this more intuitive label.

long as $\widetilde{a} > 0$:

$$\pi\widetilde{a} - \underline{f} \geqslant 0. \tag{5}$$

Hence, the design of the optimal fine schedule is equivalent to (indirectly) implementing a (profit-maximizing) action $\widetilde{a}$, that is an action that the firm is willing to choose according to the incentive compatibility and undertake constraints. Among all the implementable actions $\widetilde{a}$, the optimal policy selects the one that maximizes the expected welfare (the subscript PS refers to per-se rules)

$$EW_{PS}(\beta) = \left[(1-\beta)w^g - \beta w^b\right]\widetilde{a} = Ew(\beta)\widetilde{a}, \tag{6}$$

which we denote $\widehat{a}$[11]. Notice that any $\widetilde{a} \in A$ can be implemented setting $\underline{f} = 0$ and $\overline{f} \geqslant \pi$. Then, the optimal policy depends only on the effect of the practice on the expected welfare $Ew(\beta)$. We implicitly derive the optimal legal standard (per-se legality or per-se illegality) from the optimal enforcement policy: if no fine is given for any action $a \in A$ the corresponding legal standard is per-se legality, while if a positive fine is set at least for some actions, then the enforcer is (implicitly) following a per-se illegality rule. The result below immediately follows:

Lemma 1. *When $Ew(\beta) \geqslant 0$, that is for $\beta \in \left[0, \overline{\beta}\right]$, where*

$$\overline{\beta} = \frac{w^g}{w^g + w^b}, \tag{7}$$

the optimal legal standard is per-se legality and the optimal enforcement policy implements $\widehat{a} = 1$, by choosing any $\underline{f} < \pi$. When $Ew(\beta) < 0$, or $\beta \in \left(\overline{\beta}, 1\right]$, the optimal legal standard is per-se illegality and the optimal enforcement policy implements $\widehat{a} = 0$, by setting $\overline{f} \geqslant \pi$.

3.2 Discriminating rule

A discriminating rule is based both on the observed actions and on the signal received about the state of the world, and considers an action $a \in A$ as illegal if the enforcer receives a signal $\sigma = b$. Although the signal may be incorrect, we have assumed it to be informative. The enforcer, then, can indeed implement – differently from the 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 $\sigma = g$, the fine schedule $f(a)$ applies only when the signal of the bad state is received.

[11] This is the standard marginal deterrence problem in law enforcement. See Mookherjee and Png (1994).

In the bad state, the practice is socially harmful but the enforcer receives a signal $\sigma = b$ with probability $(1-\varepsilon^{II})$. With probability ε^{II}, instead, the signal is $\sigma = g$ and the enforcer cannot set any fine. Given the fine schedule $f(a)$, the incentive compatibility constraint in the bad state (ICC^b) is:

$$\pi\widetilde{a}^b - (1-\varepsilon^{II})\underline{f} \geqslant \pi - (1-\varepsilon^{II})\overline{f}, \tag{8}$$

while the undertake constraint (UC^b) is:

$$\pi\widetilde{a}^b - (1-\varepsilon^{II})\underline{f} \geqslant 0. \tag{9}$$

In the good state ($s = g$) the practice is socially beneficial but the enforcer receives a signal $\sigma = b$ with probability ε^I. Although the incentive compatibility constraint ICC^b to implement $\widetilde{a}^b$ puts only a lower bound on the maximum fine $\overline{f}$, when type-I errors are committed an excessively high $\overline{f}$ would induce the firm to undertake $a^g = \widetilde{a}^b$ rather than $a^g = 1$ in the good state[12]. Hence, we have to further impose the following incentive compatibility (ICC^g) and undertake (UC^g) constraints:

$$\pi\widetilde{a}^b - \varepsilon^I\underline{f} \leqslant \pi - \varepsilon^I\overline{f} \tag{10}$$

and

$$\pi - \varepsilon^I\overline{f} \geqslant 0. \tag{11}$$

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.,

$$\overline{f} \in \left[\underline{f}_D + \frac{\pi(1-\widetilde{a}^b)}{1-\varepsilon^{II}}, \underline{f} + \frac{\pi(1-\widetilde{a}^b)}{\varepsilon^I}\right]. \tag{12}$$

The expected welfare can be written as

$$EW_D(\beta) = \left[(1-\beta)w^g - \beta w^b\widetilde{a}^b\right] - \frac{\gamma}{2}(\overline{\varepsilon} - \varepsilon^I)^2 - \frac{\gamma}{2}(\overline{\varepsilon} - \varepsilon^{II})^2. \tag{13}$$

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.

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

3.3 Optimal legal standards in traditional industries

Having identified the optimal enforcement policy under per-se and discriminating rules, we can now move to the choice of the optimal legal standards. The following Proposition summarizes the results.

Proposition 1 (Optimal legal standards in traditional industries) *In traditional industries the discriminating rule always dominate 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* $\overline{f} \in \left[\frac{\pi}{(1-\overline{\varepsilon})}, \frac{\pi}{\overline{\varepsilon}}\right]$.

With unbounded fines the discriminating rule always dominates the per-se rules by implementing the welfare maximizing course of actions. This result, although expected, is not entirely obvious: when comparing per-se and discriminating rules, indeed, the latter is more flexible under a certain respect, allowing to set a fine schedule contingent on the states of the world. However, a disciminating rule is also more rigid under a different perspective, since it prevents the enforcer from setting a fine on the practice when the signal is good, an outcome that is not precluded under a per-se illegality regime. However, sanctioning a welfare enhancing action is not desirable in the present setting, and the second source of rigidity, that would reduce the appeal of a discriminating rule, does not bite in this case. We shall observe that in innovative industries this feature is not necessarely true.

Notice that, avoiding to spend resources to refine the assessment of the effects, the enforcer adopts the minimum level of accuracy through a protocol of investigation that acquires a very limited set of evidence.

The pivotal role of fines in enforcement suggests to analyze the alternative case when the maximum feasible fine F is capped at some upper bound. For instance, when the limited liability cap $F = \pi$ applies, the first best course of actions described in Proposition 1 is no more implementable, since $\overline{f}$ cannot be set at a level sufficiently high to induce the firm to take $a^b = 0$ in the bad state. In this case, a^b becomes a function of the type-II error ε^{II}. Setting $\underline{f} = 0$, from the lower bound of (12) we get[13]

$$\widehat{a}^b = \varepsilon^{II}. \tag{14}$$

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 $\widehat{a}^b$,

[13] The same qualitative argument applies for any $F \in \left(\pi, \frac{\pi}{(1-\overline{\varepsilon})}\right)$, since the incentive compatibility constraints to obtain the first best course of actions cannot be met. When F is capped in the interval above, the implementable action in the bad state is $\widetilde{a}^b = 1 - (1 - \varepsilon^{II})\frac{F}{\pi}$.

improving marginal deterrence. The following Proposition describes the optimal policies and legal standards in the limited liability case.

Proposition 2 (Optimal legal standards with limited fines). *When fines are capped by limited liability, that is $F = \pi$, the optimal legal standard is the discriminating rule for low β and per-se illegality for high β.*

More precisely, there exists a threshold $\widetilde{\beta} \in (\overline{\beta}, 1)$ such that for $\beta \in \left[0, \widetilde{\beta}\right)$ the discriminating rule is optimal and implements actions $a^g = 1$ and $\widehat{a}^b = \max\left\{\overline{\varepsilon} - \frac{\beta w^b}{\gamma}, 0\right\}$, by setting $\underline{f} = 0$, $\overline{f} = \pi$, $\varepsilon^I = \overline{\varepsilon}$ and improving type-II accuracy: $\varepsilon^{II} = \max\left\{\overline{\varepsilon} - \frac{\beta w^b}{\gamma}, 0\right\}$. For $\beta \in \left[\widetilde{\beta}, 1\right]$ the per-se illegality rule dominates implementing $a^g = a^b = 0$.

The adoption of a per-se illegality rule when the practice is likely to be socially harmful and fines are capped is driven by the lower probability (zero probability in our simplified setting) of committing errors when following a simpler legal standard, that in turn makes the control of actions more effective and less costly. When fines are capped, indeed, marginal deterrence is hampered and a discriminatory rule does not succeed to implement always the first best course of actions $a^g = 1$ and $a^b = 0$. In this case the enforcer cannot prevent the firm from taking, in the bad state, a socially damaging action $a^b > 0$. Even if the first best action 0 is not implemented in the bad state, the discriminating rule allows the enforcer inducing a more efficient course of actions compared to any per-se rule. However, to obtain this result, the enforcer has to spend to improve type-II accuracy. For very high β, per-se illegality becomes the dominant legal standard, allowing to save on accuracy costs and to obtain the efficient action 0 in the bad state at the cost of a suboptimal action (0 rather than 1) in the (unlikely) good state.

4 Antitrust intervention in innovative industries

In innovative industries, the firm gains market power if it is able to discover the new technology, an outcome that requires to invest in research. Starting from the first best benchmark, the efficient course of actions in case of successful research is the same as in the traditional environment, that is $a^b = 0$ and $a^g = 1$. The expected welfare is therefore $EW_{FB}(\beta, I) = I(1-\beta)w^g - \frac{I^2}{2}$ that yields the optimal investment level

$$I_{FB} = (1-\beta)w^g, \tag{15}$$

which is increasing in the likelihood of the good state $(1-\beta)$ and in the associated welfare gain w^g. Notice that, since the actions are efficiently chosen, the practice is undertaken (with the highest intensity) only when it is welfare improving. Hence, discovering the new technology has a positive expected impact on welfare. Therefore, the level of investment is always positive, although decreasing in the probability of social harm, since the enforcer takes into account the private cost of research. The expected welfare, evaluated at the first-best policies, is

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

4.1 Per-se rules

At stage 3, if the research has been successful, the firm acquires a dominant position and chooses among the actions in A given the fine schedule $f(a)$. If research instead fails to produce a result, the firm, which lacks any market power and is not sanctioned, gets the competitive profits equal to zero.[14] Upon discovery, therefore, the firm (which becomes dominant and is subject to antitrust scrutiny), adopts $\widetilde{a}$ according to the incentive compatibility and undertake constraints, with expected profits $E\Pi_{PS} = I(\pi\widetilde{a} - \underline{f}) - \frac{I^2}{2}$, choosing the investment

$$I_{PS} = \pi\widetilde{a} - \underline{f}. \tag{16}$$

The expected welfare under per-se rules is therefore

$$EW_{PS}(\beta) = I_{PS}Ew(\beta)\widetilde{a} - \frac{I_{PS}^2}{2}. \tag{17}$$

The enforcer maximizes the expected welfare setting the parameters of the fine schedule, $\widetilde{a}$, $\underline{f}$ and $\overline{f}$, subject to the incentive compatibility and undertake constraints. Notice that now the enforcer has to take into account not only the effect of the policy on the actions, the usual marginal deterrence problem, but also the impact on the expected profits and the incentives to invest, an additional effect that we can label as average deterrence. The following Lemma states the optimal enforcement policy under per-se rules.

Lemma 2: *The optimal legal standard and enforcement policy under per-se rules are:*
(i) for $\beta \in \left[0, \underline{\beta}\right]$, *where*

$$\underline{\beta} = \frac{w^g - \pi}{w^g + w^b}, \tag{18}$$

[14] In this setting therefore, the firm obtains zero profits when the innovation is not discovered and the market is competitive as well as in case of innovation when the practice is not undertaken ($a = 0$). Alternatively, we might assume that when the innovation is realized but the practice is not adopted the firm hets a positive profit Π. As we discuss in section 5, the main results would not change but the regions are identified by a larger set of parameters. For this reason we prefer to keep the benchmark model as simple as possible.

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 $\beta \in \left(\underline{\beta}, \overline{\beta}\right)$, *where*

$$\overline{\beta} = \frac{w^g}{w^g + w^b}, \tag{19}$$

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 $\beta \in \left[\overline{\beta}, 1\right]$, *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* $EW_{PS}(\beta) = 0$.

As in the case of traditional industries, when the social harm is unlikely ($\beta < \underline{\beta}$) per-se rules dictate a per-se legality regime. In innovative industries, however, this result applies to a narrower interval ($\underline{\beta} < \overline{\beta}$), as can be appreciated comparing Lemma 1 and Lemma 2. The enforcer, indeed, takes into account not only the actions implemented but also the cost of research (the term $-\pi$ in the threshold $\underline{\beta}$). Hence, although for $\beta \in \left(\underline{\beta}, \overline{\beta}\right)$ the practice increases the expected welfare ($Ew(\beta) > 0$) and the same course of actions as in the per-se legality rule is implemented, compared to this latter regime the enforcer reduces the level (and cost) of investement by progressively increasing the fine. The only way to set a positive fine, then, is to consider the practice as illegal, that is to adopt a per-se illegality rule. Finally, it is only when the probability of social harm becomes sufficiently high ($\beta > \overline{\beta}$) to make the expected marginal welfare of the new actions $Ew(\beta)$ negative that the optimal policy prevents the adoption of the practice ($a = 0$), eliminating therefore also any incentive to invest in research.

4.2 Discriminating rules

Turning to the discriminating rule, at stage 3 when research is successful, the choice of the action in the innovative environment parallels the traditional industry. Once again the enforcer can implement $a^b = \widetilde{a}^b$ and $a^g = 1$ satisfying the incentive compatibility and undertake constraints.

At stage 2, the firm decides the level of investment that maximizes

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

The innovative investment in the discriminating regime is then

$$I_D = (1-\beta)\left[\pi - \varepsilon^I \overline{f}\right] + \beta\left[\pi\widetilde{a}^b - (1-\varepsilon^{II})\underline{f}\right] \geqslant 0. \tag{20}$$

Notice, by comparing (20) with (15), that the level of investment in the discriminating regime is below the first best for any set of policy parameters. Moreover, 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

$$EW_D = I\left[\Delta W_D - \frac{I_D}{2}\right] - \frac{\gamma}{2}(\overline{\varepsilon}-\varepsilon^I)^2 - \frac{\gamma}{2}(\overline{\varepsilon}-\varepsilon^{II})^2, \tag{21}$$

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 incentive compatibility and undertake constraints. As before, we denote as $\widehat{a}^b$ the action that solves this program. In the following Lemma we identify the optimal policy.

Lemma 3: *The optimal enforcement policy under the discriminating regime depends on the probability of social harm β:*
(i) for $\beta \in \left[0, \underline{\underline{\beta}}\right]$, where

$$\underline{\underline{\beta}} = \frac{w^g - w^b - \pi}{w^g + w^b}.$$

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 $\beta \in \left(\underline{\underline{\beta}}, 1\right]$ 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 $\beta \to \underline{\underline{\beta}}$ 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 the same interval $\left[0, \underline{\underline{\beta}}\right]$ in which also the per-se rule opted for generalized acquittal, since $\underline{\underline{\beta}} < \underline{\beta}$. Above this interval, the discriminating rule allows the enforcer implementing 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, indeed, implements an action $\widehat{a}^b$ greater than zero (which implies some social damage ex-post) in order to ex-ante sustain profits and research, that is to soften average deterrence, turning 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 (20). 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.

When $\beta > \underline{\underline{\beta}}$, therefore, the optimal policy under a discriminating rule adopts a mixture of high fine and type-I accuracy to reduce the investment and the action $\widehat{a}^b$. In the same interval, the optimal policy under the per-se illegality rule (Lemma 2), instead, was designed first to reduce the investment but not the action, and then (for $\beta > \overline{\beta}$) reverted to implementing $\widehat{a} = 0$ completely discouraging the investment.

Before moving to the analysis of the optimal legal standard, we further analyze the enforcement policy under a discriminating rule when fines are capped by a limited liability rule, that is $F = \pi$. The policy problem is equivalent to the one analyzed in Lemma 3, with the further constraint that $\overline{f} \leqslant \pi$. Since for $\beta > \underline{\underline{\beta}}$ the optimal policy implements $\widehat{a}^b < 1$ by setting a positive fine $\overline{f} = \frac{\pi(1-\widehat{a}^b)}{(1-\overline{\varepsilon})}$ and $\widehat{a}^b$ is decreasing in the likelihood of social harm β, the optimal fine $\overline{f}$ itself is increasing in β. As long as the limited liability constraint does not bind, therefore, the optimal policy under the discriminating rule is the one described in Lemma 3. However, there will exist a $\widehat{\beta} > \underline{\underline{\beta}}$ at which $\overline{f} = \pi$ and the limited liability constraint starts binding. Indeed, this occurs when the action implemented in the bad state is $\widehat{a}^b = \overline{\varepsilon}$. Hence, we have to analyze the optimal policy for $\beta \in \left[\widehat{\beta}, 1\right]$. In this interval the limited liability constraint affects the level of the implementable action. The following Lemma states the optimal policy under discriminating rule and limited liability.

Lemma 4 (Optimal policy under limited liablity): *There exists a $\widehat{\beta} > \underline{\underline{\beta}}$ such that the limited liability constraint $\overline{f} \leqslant \pi$ does not bind for $\beta \in \left[0, \widehat{\beta}\right]$ when $\overline{f}$ is optimally set. In this interval the optimal policy is the one described in Lemma 3. Instead, for $\beta \in \left(\widehat{\beta}, 1\right]$ 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 on fines binds, the enforcer implements a balanced reduction in both errors, a lower type-I error to sustain the investment and to soften average deterrence and more type-II accuracy to improve marginal deterrence.

4.3 Optimal legal standards in innovative industries

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. We analyze both the case when the enforcer can use unlimited fines and when instead the sanctions are capped by a limited liability rule. The following propositions establish the main results.

Proposition 3 (Optimal legal standards in innovative industries) *When fines are unbounded the optimal legal standard is a per-se legality rule for* $\beta \leqslant \underline{\underline{\beta}}$ *and a discriminating rule for higher* β.

The main difference with the traditional industries case rests on the adoption of a per-se legality standard when the practice is very likely to increase welfare under the new technology. In this case under a per-se legality rule the enforcer commits not to fine the practice when it is socially damaging. The resulting effect is an increase in the incentives to invest and in the probability to discover the new technology.

When fines are capped by limited liability, the comparison of legal standard is enriched by additional effects. For $\beta \in \left[0, \underline{\underline{\beta}}\right]$, the discriminating rule is equivalent to a per-se legality regime, which in turn dominates the per-se illegality rule. When the likelihood of social harm increases, at some point $\widehat{\beta} > \underline{\underline{\beta}}$ the limited liability constraint on fines starts binding and the enforcer has to use different types of accuracy to improve marginal deterrence and to sustain the investment. In this region the expected welfare under the discriminating rule is decreasing and becomes negative for $\beta \rightarrow 1$. Hence, there is a threshold $\overline{\overline{\beta}}$ such that for higher β per-se illegality – which has a non-negative expected welfare – becomes the dominant rule. The following Proposition summarizes this result.

Proposition 4 (Optimal legal standards under limited liability). *When fines are capped by limited liability, the optimal legal standard for increasing values of* β *is: per-se legality* ($\beta \in \left[0, \underline{\underline{\beta}}\right)$)*; the discriminating rule without the limited liability constraint and with type-I accuracy* ($\beta \in \left[\underline{\underline{\beta}}, \widehat{\beta}\right)$)*; then the discriminating rule with the limited liability constraint and a more balanced accuracy on both errors* ($\beta \in \left[\widehat{\beta}, \overline{\overline{\beta}}\right)$)*; and finally per-se illegality* ($\beta \in \left[\overline{\overline{\beta}}, 1\right]$)*.*

Hence in innovative industries, when the limited liability constraint restricts the level of fines available to the enforcer, we have, for increasing values of the likelihood of social harm, the full range of legal standards. We have obtained a much richer outcome compared with the case of traditional industries with unlimited fines, where the discriminating rule was dominant for any value of β.

5 Concluding comments

Our model delivers several prescription on the optimal legal standards, accuracy and enforcement policies for environments that differ in two relevant dimensions. The first refers to the source of market power, that makes the practice a relevant concern for the antitrust intervention: in traditional industries the incumbent is already established and the technology is stable, while in innovative industries innovation and dominance come together. The second dimension is the set of feasible policy instruments, namely the level of admitted fines, that may be unbounded or capped.

In the resulting four environments we have considered per-se (legality and illegality) rules, that base the notion of illegality, and the associated enforcement policy, on the nature of the practice adopted, and a discriminating rule that defines unlawfulness as the emergence of a social harm due to the practice chosen by the firm. Due to the richer definition of unlawfulness, discriminating rules may entail judicial errors and require the enforcer to choose the level of accuracy. For each environment we have first analyzed the optimal policy for given legal standard and then selected the optimal legal standard.

Our results deserve some comments.

First, the equilibrium outcomes and the different regions depend on the parameters $\{w^g, w^b, \pi, \beta\}$ 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, as those proposed by the Chicago school, have also campained for per-se legality rules, while a more articulated reconstruction of the competitive and anticompetitive effects of those practices, usually associated to the post-Chicago scholars, has represented the background for the effect-based approach to unilateral practices.

Secondly, the legal rules considered differ in their relative flexibility under several dimensions, where full flexibility should be intended as the ability to implement any action in any state of the world by appropriately shaping the fine schedule. At a first glance, a discriminating rule is more flexible being able to set fines contingent on the effects. This feature entirely explains why in the simplest setting (traditional industries and no upper bound on fines) a discriminating legal standard dominates for any prior β of the social harm. However, per-se rules potentially give rise to another form of flexibility, since they allow (under a per-se illegality regime) to levy a fine even on a socially beneficial practice as well as being lenient (when opting for per-se legality) on socially harmful practices. We have seen that these outcomes can be desirable in innovative industries, making per-se rules more appealing in some cases. In innovative industries, indeed, the enforcer adopts a per-se legality standard when the new technology is very likely to be welfare enhancing and the incentives to innovation prevail in the policy problem, while it opts for a per-se illegality rule, sanctioning the practice even if socially beneficial, when the very likely social harm commands to discourage the investment in research.

A third interesting feature of our results refers to accuracy. We have seen that type-II accuracy can improve marginal deterrence, while the reduction of type-I error may soften average deterrence sustaining innovative investments. In the intermediate interval where the discriminating rule is applied, the enforcer initially reduces type-I errors (Lemma 3) while then, once fines are capped, it reverts to more symmetric accuracy and reduces both errors (Lemma 4).[15]

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 ($\bar{\varepsilon}$) – or enrich the set of evidence. As long as the enforcer collects information on the (empirical) predictions of the competitive model, she is able to refine the assessment of the efficiency-enhancing effects, reducing the probability of condemning an innocent firm, that is a type-I error. Conversely, additional evidence of the anti-competitive explanation implements a better type-II accuracy. 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, guarantee-

[15] In this paper the 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. Instead, Kaplow (2011) has shown that – changing the burden of proof – the enforcer faces a trade-off between a higher (lower) probability of type-I error and a lower (higher) probability of type-II errors.

ing 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. Polo (2010) offers an example of this identification strategy referred to selective price cuts.

Fourth, by comparing the level of innovative activity corresponding to the optimal legal standard with the first best level in (15), it is simple to show that if there is no upper bound on feasible fines, the optimal policy always involves under-investment since private incentives (profits) are lower than social ones (welfare). If instead, feasible fines cannot exceed the firm's profits, the optimal policy might entail over-investment in the region where the discriminating rule is optimal and the limited liability constraint is binding, whereas under-investment occurs elsewhere. This result depends in part on the way we have modelled the innovative process.

On the one hand, in the analysis of the innovative industries we have assumed that the new technology creates an extra-profit and a welfare gain or loss only when combined with the practice (when $a > 0$). One may argue, however, that the new technology may increase the private and social payoffs independently of the practice, although this latter may further magnify the results. In order to easily account for the independent effect of the new technology on profits and welfare we might add a fixed positive term to the profits and welfare that occurs even in the absence of the practice ($a = 0$), as for instance assuming that when research is successful $\Pi(a) = \Pi + \pi a$, $W^g(a) = W + w^g a$ and $W^b(a) = W - w^b a$. This way we would magnify the private and the social effects of the new technology independently of the practice adopted (and of the effects of the enforcement policy on the practice itself). In other words, these terms would create a private and social incentive to realize the innovation orthogonal to the choice of the practice. Since in this paper we want to focus on antitrust intervention on the practice and how it is affected in innovative industries, we preferred to drop this component and the associated fixed terms.

These latter, in any case, would not affect private and public choices, but would only change the thresholds delimiting the different regions. For instance, it is easy to see that the threshold $\underline{\beta} = \frac{w^g - \pi}{w^g + w^b}$, the upper bound to per-se legality when per-se rules are designed in innovative industries, would become

$$\underline{\beta}' = \frac{w^g - \pi + W - \Pi}{w^g + w^b} = \underline{\beta} + \frac{W - \Pi}{w^g + w^b}.$$

This threshold (and, in a similar way, the others), would then shift according to the impact of the new technology on private and social payoffs. Specifically, if the private effect Π is larger than the social one W, then the thresholds would shift to the left, leading to a tighter control of the actions: since in this case the private incentives to investment become stronger, the public policy is less biased to sustain the investment and more concerned with controlling the actions, opting for a stricter legal standard and policy. Conversely, when the new technology affects social welfare more than private profits ($W > \Pi$) the thresholds would shift to the right, enlarging the region where more lenient legal rules are optimal.

On the other hand, in modelling the innovative process we have followed the approach first proposed in Arrow (1962) to compare the incentive to innovate under different market structures, the distinguishing feature being the assumption that there is just one firm that has the possibility of investing in research. In this setting underinvestment depends on the misalignment of private and social payoffs. Quite often, however, several firms invest at the same time in research, engaging in patent races. How our setting can be adjusted to include this different feature of the innovative process? In the literature on patent races we find two different approaches which also produce the private-public misalignment just mentioned, but introduce two further reasons for over-investment. In Dasgupta and Stiglitz (1980, section 3) the patent race takes the form of a winner-takes-all competition, resulting in excessive individual effort. Lee and Wilde (1980), Loury (1979) and Dasgupta and Stiglitz (1980, Section 4), instead, model uncertainty in a way which creates a smoother probability of success as a function of the individual efforts, leading to excessive participation in equilibrium. These additional reasons for over-investment are not present in our model. Hence, by combining our approach with a different modelling of patent races we may have more investment in innovation than in the present setting. Qualitatively, we would expect the same impact (stricter legal standards and policies) as the one seen above when the new technology improved private more than social payoffs ($\Pi > W$).

Appendix

Proof of Proposition 2. First of all, substitute $\underline{f} = 0$ and $\overline{f} = \pi$ in (12), obtaining $\widehat{a}^b = \varepsilon^{II}$. Substituting in the expected welfare we get:

$$EW_D(\beta) = (1-\beta)w^g - \beta w^b \varepsilon^{II} - \frac{\gamma}{2}(\overline{\varepsilon} - \varepsilon^I)^2 - \frac{\gamma}{2}\left(\overline{\varepsilon} - \varepsilon^{II}\right)^2, \tag{22}$$

Improving accuracy is costly and type-I accuracy does not affect private actions; hence, we have $\varepsilon^I = \overline{\varepsilon}$. Conversely, type-II accuracy makes the incentive compatibility constraints slacker. Differentiating with respect to ε^{II} and solving we get

$$\varepsilon^{II} = \widehat{a}^b = \overline{\varepsilon} - \frac{\beta w^b}{\gamma}.$$

If $\gamma \geqslant \frac{w^b}{\overline{\varepsilon}}$, the two expressions are non negative for $\beta \in [0,1]$. Then substituting in the expression of the expected welfare we have

$$EW_D(\beta) = Ew(\beta) + \beta w^b(1-\overline{\varepsilon}) + \frac{(\beta w^b)^2}{2\gamma}.$$

Then recalling from Lemma 1 that for $\beta \in \left[0, \overline{\beta}\right]$, per-se rules at the optimal policy give an expected welfare of $EW_{PS}(\beta) = Ew(\beta)$, it immediately follows that the discriminating rule dominates the per-se illegality regime. For $\beta \in \left(\overline{\beta}, 1\right]$, the optimal per-se rule gives a welfare equal to zero if $F = \pi$. Hence, the optimal policy is the discriminating (per-se illegality) rule as long as $EW_D(\beta) > 0$ ($\leqslant 0$). Differentiating $EW_D(\beta)$ with respect to β

$$\frac{\partial EW_D(\beta)}{\partial \beta} = -(w^g + w^b\overline{\varepsilon}) + \beta\frac{(w^b)^2}{\gamma} \leqslant 0 \quad \text{for} \quad \gamma \geqslant \frac{w^b}{\overline{\varepsilon}}.$$

Moreover, $EW_D(\overline{\beta}) > 0 = EW_{PS}(\overline{\beta})$, it is convex and $EW_D(1) = w^b\left[-\overline{\varepsilon} - \frac{w^b}{\gamma}\right] \leqslant 0$ for $\gamma \geqslant \frac{w^b}{\overline{\varepsilon}}$. Hence, there must be a value $\widetilde{\beta} > \overline{\beta}$ such that $EW_D(\beta) < 0$ for $\beta > \widetilde{\beta}$ and per-se illegality dominates.

Consider now the case $\gamma \leqslant \frac{w^b}{\overline{\varepsilon}}$. The optimal policy under a modified per-se illegality rule works as in the previous case. However, for $\beta \geqslant \frac{\overline{\varepsilon}\gamma}{w^b}$ we have $\widehat{a}^b = \varepsilon^{II} = 0$ and the expected welfare becomes $EW_D(\beta) = (1-\beta)w^g - \frac{\gamma}{2}\overline{\varepsilon}^2$ that is decreasing in β and negative for $\beta = 1$. Hence, for a sufficiently high β $EW_D(\beta) < 0 = EW_{PS}(\beta)$ and per-se illegality dominates. ■

Proof of Lemma 2. The maximization program is solved by the following first-order conditions

$$\begin{aligned}
\frac{\partial EW_{PS}}{\partial \widetilde{a}} &= [Ew(\beta)\widetilde{a} - I_{PS}]\,\pi + Ew(\beta)I_{PS} + \lambda \geqslant 0, \\
\frac{\partial EW_{PS}}{\partial \underline{f}} &= -\left[Ew(\beta)\widetilde{a} - I_{PS}\right] - \frac{\lambda}{\pi} \leqslant 0, \\
\frac{\partial EW_{PS}}{\partial \overline{f}} &= \frac{\lambda}{\pi} \geqslant 0,
\end{aligned} \tag{23}$$

Finally, the complementary slackness condition is

$$\lambda\left(\widetilde{a} - 1 + \frac{\overline{f} - \underline{f}}{\pi}\right) = 0. \tag{24}$$

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 $\beta \in \beta \in [0, \underline{\beta}]$ 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 $\frac{\partial EW_{PS}}{\partial \underline{f}} = -[Ew(\beta) - \pi]\widetilde{a} < 0$ and setting $\underline{f} = 0$ is optimal. Moreover, $\frac{\partial EW_{PS}}{\partial \widetilde{a}} = [Ew(\beta) - \pi]\pi\widetilde{a} + Ew(\beta)\pi\widetilde{a} > 0$ and $\widehat{a} = 1$. Hence, $\overline{f}$ is not needed to define the fine schedule.

(ii) For $\beta \in (\underline{\beta}, \overline{\beta})$, $Ew(\beta) > 0 > Ew(\beta) - \pi$ 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 $\beta = \overline{\beta}$.

(iii) For $\beta \in [\overline{\beta}, 1]$, $0 \geqslant Ew(\beta) > Ew(\beta) - \pi$ 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 $\overline{f} \geqslant \pi$. ■

Proof of Lemma 3. 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 (12) and verifying it ex-post. The first order conditions are the following

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

where $\Delta W_D = (1-\beta)w^g - \beta w^b \widetilde{a}^b$ and I_D is given by (20).

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 $\Delta W_D - I_D = [w^g - \pi - \beta(w^g + w^b)] > 0$ for $\beta < \frac{w^g - \pi}{w^g + w^b}$. Moreover, for $\beta < \frac{w^g - w_b - \pi}{w^g + w^b} < \frac{w^g - \pi}{w^g + w^b}$, $\frac{\partial EW_D}{\partial \widetilde{a}^b} = \beta\pi\,[w^g - w^b - \pi - \beta(w^g + w^b)] > 0$. Hence, for $\beta < \frac{w^g - w^b - \pi}{w^g + w^b} = \underline{\underline{\beta}}$, $\frac{\partial EW_D}{\partial \widetilde{a}^b} > 0$, $\frac{\partial EW_D}{\partial \underline{f}} < 0$ and $\frac{\partial EW_D}{\partial \overline{f}} < 0$ at $\varepsilon^I = \varepsilon^{II} = \overline{\varepsilon}$. Finally, the incentive compatibility constraints (12) is 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 $\beta > \underline{\underline{\beta}}$. We set $\widehat{a}^b < 1$ to obtain $\frac{\partial EW_D}{\partial \widehat{a}^b} = 0$, 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 (12), that is, $\overline{f} = \frac{\pi(1-\widehat{a}^b)}{(1-\overline{\varepsilon})}$.
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 $\beta > \underline{\underline{\beta}}$, 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 $\beta \to \underline{\underline{\beta}}$ 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 $\underline{\underline{\beta}}$ 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 $\frac{\partial I_D}{\partial\beta}=-\frac{1-\varepsilon^I-\overline{\varepsilon}}{(1-\overline{\varepsilon})}\pi(1-\widehat{a}^b)$ 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 Lemma 4. Combining the incentive compatibility and limited liability constraints by setting $\overline{f}=\pi$ and $\underline{f}=0$ in (12) 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

$$\frac{\partial EW_D}{\partial\varepsilon^I} = [\Delta W_D-I_D]\frac{\partial I_D}{\partial\varepsilon^I}+\gamma(\overline{\varepsilon}-\varepsilon^I)\geqslant 0$$

$$\frac{\partial EW_D}{\partial\varepsilon^{II}} = [\Delta W_D-I_D]\frac{\partial I_D}{\partial\varepsilon^{II}}-\beta w^b\frac{\partial\widetilde{a}^b_D}{\partial\varepsilon^{II}}+\gamma(\overline{\varepsilon}-\varepsilon^{II})\geqslant 0$$

that hold as equalities with internal solutions $\varepsilon^I < \overline{\varepsilon}$ and $\varepsilon^{II} < \overline{\varepsilon}$.

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 $EW_D(1) = -\varepsilon^{II2}\pi\left(w^b + \frac{\pi}{2}\right) < 0.$ ■

Proof of Proposition 3. Given Lemma 2, the per-se rules give

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

Instead, the discriminating rule (Lemma 3) gives

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

Let us compare the expected welfare in the different regimes for increasing values of β. For $\beta \in \left[0, \underline{\underline{\beta}}\right]$, both D and PS are equivalent to the per-se legality regime. In the interval $\left(\underline{\underline{\beta}}, \underline{\beta}\right]$ 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 $\beta \in \left(\underline{\beta}, \overline{\beta}\right)$, 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 $\varepsilon^I = \varepsilon^{II} = \overline{\varepsilon}$, $\widehat{a}^b = 1$ and $\underline{f} = [\pi - Ew(\beta)] \,/\, [(1-\beta)\overline{\varepsilon} + \beta(1-\overline{\varepsilon})]$, 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 $\beta \in [\overline{\beta}, 1]$, $EW_{PS}(\beta) = 0$ while $EW_D(\beta)$ is decreasing and equal to zero only at $\beta = 1$. ■

Bibliography

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 Dominique and Claude 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 Frank (1984), The Limits of Antitrust, *Texas Law Review*, **63**, 9-14.

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

Gual, Jordi, Martin Hellwig, Anne Perrot, Michele Polo, Patrick Rey, Klaus Schmidt, and Rune Stenbacka (2005), *An Economic Approach to Article 82*, Report for the DG Competition, European Commission, published on *Competition Policy International*, 2005, 2: 111-156.

Immordino, Giovanni, Marco Pagano and Michele Polo (2011), "Incentives to Innovate and Social Harm: Laissez-faire, Authorization or Penalties?," *Journal of Public Economics,* **95**, 864-876.

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

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

Kaplow, Louis (2011), Optimal Proof Burdens, Deterrence and the Chilling of Desirable Behavior, *mimeo*

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

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

Katsoulakos Yannis and David 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, Dilip 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, Mithcell A. and Steven Shavell (2000), "The Economic Theory of Public Enforcement of Law," *Journal of Economic Literature*, **38**, 45-76.

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

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

Segal Ilya, and Michael Whinston (2007), "Antitrust in Innovative Industries," *American Economic Review*, **97**(5), 1703-1730.

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

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

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

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

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