Optimal Patent Screening with Imperfect Enforcement*

Gerard Llobet
CEMFI and CEPR

Álvaro Parra
Dalhousie

Javier Suarez
CEMFI and CEPR

July 21, 2025

Abstract

Using an industry-dynamics innovation model, we explore the interplay between patent screening and patent enforcement. When patent enforcement is imperfect, genuine sequential innovators may see their market access blocked by potentially abusive infringement claims from prior innovators. In this case, the patent's office leniency towards obvious innovators may contribute to reduce the presence of blocking incumbent monopolists, encouraging innovation and improving welfare. This result is robust across multiple extensions of the baseline framework.

JEL Codes: L26, O31, O34.
Keywords: Innovation, Patent Screening, Patent Enforcement, Industry Dynamics.

*We are thankful to the editor and two referees for their comments. This work is based on a previous paper entitled "Entrepreneurial Innovation, Patent Protection, and Industry Dynamics." Financial support from the European Commission (grant CIT5-CT-2006-028942) and the Spanish Ministry of Science and Innovation through the ECO2008-00801, the Consolider-Ingenio 2010 Project "Consolidating Economics" and the Regional Government of Madrid, through grant S2015/HUM-3491 is gratefully acknowledged. Álvaro Parra is partly supported by funding from the Social Sciences and Humanities Research Council of Canada. Address for correspondence: Gerard Llobet, CEMFI, Casado del Alisal 5, 28014 Madrid, Spain. Tel: +34-914290551. Fax: +34-914291056. Email: llobet@cemfi.es.

1 Introduction

Innovation, entry, exit, and the level of competition are complex interrelated phenomena in every industry, and especially so in those that are more dynamic and technologically intensive. Many of these industries rely on patents and other intellectual property rights (IPRs) as a source of temporary market power that allows successful innovators to obtain a return from their innovation investments. Intellectual property, however, is a double-edged sword for the technological progress of an industry since, once innovators become established incumbents, it may discourage subsequent innovation. Incumbents can use the threat of litigation to fend off entry or to obtain a favorable settlement in licensing negotiations with future innovators. Under perfect IPR enforcement this trade-off between current and future innovation is optimally resolved by protecting significant innovators from minor subsequent innovators but not from significant subsequent innovators (Hopenhayn *et al.*, 2006).

In this paper we study the optimal patent policy in situations where IPR enforcement is imperfect. We show that reducing the barriers to entry for low value innovations during the patent granting process and/or by courts, when confronting accusations of patent infringement, may increase overall innovation and enhance social welfare. This result arises from a new mechanism that operates dynamically at the industry level. If high quality innovators face the possibility of being blocked or obliged to pay licensing fees by prior innovators, a higher leniency towards low quality innovation may be socially optimal. Low quality innovators bring competition to business niches previously monopolized by high quality innovators. By doing so, they reduce the incentives of the incumbents to legally oppose entry, effectively improving the returns to subsequent high-quality innovators.

Our results are obtained within a sequential-innovation model of an industry made of many market niches. Successful developers of improved versions of the product corresponding to each niche contribute to welfare and appropriate temporary monopoly profits as in a standard quality ladder model with limit pricing (Grossman and Helpman, 1991;

Aghion and Howitt, 1992). Innovation can be of two types, genuine or obvious. Genuine innovations represent an improvement large enough over the state of the art to entitle the owner to a valid patent while also being outside the breadth of previous patents in the corresponding niche. In contrast, an obvious innovation represents a small inventive step that, under perfect screening and enforcement conditions, would not be enough to either obtain a valid patent or to avoid being considered infringers of prior patents in their market niches.

In every period, potential innovators observe market conditions and invest in R&D until the quasi-rents from innovation are dissipated. After investment, firms discover whether their product constitutes a genuine or an obvious improvement. Every innovator, genuine or obvious, applies for a patent, as is this the only way to appropriate rents. All genuine innovations obtain patents, whereas the probability of granting patents to obvious innovations depends negatively on the effort of the patent office in screening them out. Patent enforcement is also probabilistic meaning that, although in purity genuine innovation should not be regarded to infringe prior patents whereas obvious innovation should, courts can make mistakes along both dimensions (Lemley and Shapiro, 2005; Farrell and Shapiro, 2008).[1]

Innovators face uncertainty on whether their product will reach a market niche monopolized by a prior genuine innovator which will oppose to entry by alleging the infringement of its patent rights or a competitive niche where such an opposition does not occur. This assumption captures the so-called "tragedy of anticommons" (Heller and Eisenberg, 1998).[2] If the innovator lands on a monopolized niche, the incumbent goes to court. If the court rules against the incumbent, both firms compete in the same niche, driving the prior incumbent's profits to zero and turning the entrant into the new monopolist (of its

[1]Other works in which courts make probabilistic rulings include Spier (1994), Daughety and Reinganum (1995), and Landeo *et al.* (2007). By adopting this probabilistic approach, we also abstract away from the traditional patent length and breadth discussion (Scotchmer, 2004).

[2]Lemley (2008) argues that, due to the large number of overlapping rights, firms decide to innovate first and deal later with the lawsuits that ensue from existent patent holders. This strategy is also supported by the large proportion of patents brought to court that end up being invalidated (Allison and Lemley, 1998).

incremental product quality improvement). If the court rules in favor of the incumbent, the innovator can still enter by negotiating a license. Frictions may preclude licensing from taking place, resulting in situations where the innovation goes to waste. If the innovator lands on a competitive niche, entry is accommodated.[3] Industry dynamics follows from the fact that the successful entry of genuine innovators turns market niches into (or maintains them as) monopolized ones, while obvious entry turns them into (or maintains them as) competitive ones.

In this context, we show that when patent enforcement is imperfect (in the sense that incumbents can prevent firms with genuine innovations from entering the market) the leniency of the patent grating and enforcement system towards obvious innovations has an inverted U-shaped effect on innovation and social welfare. This leniency changes market structure in a way that effectively reduces the barriers to entry faced by future innovators. In a counteracting manner, it also reduces the value of the rents that successful genuine entrants obtain conditional on entry. Our results show that for low leniency levels the first effect may dominate in promoting innovation and increasing welfare, while the second effect tends to dominate thereafter.

Patent licensing plays an important role in this setup. We show that when it is possible without frictions, the leniency towards obvious innovators that maximizes innovation and social welfare coincide, as welfare grows proportionally to the rate of technological progress. This is not the case if there are frictions that prevent licensing in some instances, in the sense that an incumbent and an entering genuine innovator fail to reach an agreement with some probability, resulting in the innovation going to waste. In this context it is optimal to foster the entry of obvious innovators beyond the level that maximizes innovation, as there is an extra gain from reducing the hurdle faced by future genuine innovators: avoiding the waste of the innovations blocked because licensing fails.

Importantly, these results are not driven by our assumptions on how licensing takes

[3] As further discussed below, we also consider a variation of the model in which this assumption is relaxed, giving raise to patent trolls.

place. Our baseline formulation considers the most competition-friendly case, where a genuine entrant that loses in court pays a licensing fee to produce and compete with the incumbent, who remains in the market. We also analyze an alternative arrangement where the incumbent that wins in court buys out the innovation of the entrant (avoiding competition between the new product and the previous top-quality product). This strategy increases the total profits earned in the niche, as it allows the incumbent to monopolize a larger quality differential with respect to the highest-quality competing alternative. We show that these higher profits actually increase the incentives to innovate and, yet, our main results on the potential desirability of some leniency in patent screening (and the rationale for it) remain qualitatively unchanged.[4]

The mechanism explaining the social value of leniency towards obvious innovations in our baseline model hinges on the assumption that incumbents with obvious innovations obtain no profits and, therefore, they have no incentives to defend their patents in court. Recent literature, however, has brought attention to the fact that in some instances these patent holders, sometimes denoted as *patent trolls*, might be able to extract a licensing payment from a future entrant using (the threat of) a patent infringement lawsuit. We analyze this possibility in a variation of the model in which all incumbents, regardless of the quality of their innovation, pose the same legal obstacle to potential entrants. In this case our entry-facilitating mechanism does not operate and innovation is reduced relative to the baseline setup. However, some screening leniency towards obvious innovators remains socially desirable when court enforcement is imperfect. This occurs because the prospect of becoming a patent troll if the innovation is not genuine also provides an incentive to innovate (given the uncertainty about the quality of the innovation when the investment is undertaken).

The article is organized as follows. In the remainder of this section, we discuss the

[4]For simplicity, our baseline model abstracts from the static dead-weight loss from market power. However, we also show that our results are robust to that. In this case, however, the leniency rate that maximizes social welfare turns out to be lower, as promoting entry further increases niche turnover and reduces monopoly power.

related literature. Section 2 introduces the baseline model. Section 3 shows that, for a given time-invariant screening leniency towards obvious innovations, the model displays a unique steady-state equilibrium. We provide the comparative statics of such an equilibrium and characterize the socially optimal patent screening leniency. Section 4 discusses the effects of static deadweight losses arising from market power, of incumbents able to accumulate innovations, of patent trolling linked to successful obvious innovation, and of licensing prior to litigation. Section 5 discusses the empirical and policy implications of the analysis and concludes. All proofs are in the Appendix.

Related Literature To our knowledge, this is the first article providing a formal framework to understand the interaction between patent screening and enforcement, as well as their impact on innovation and market dynamics. Our analysis builds, however, upon several strands of the literature.

In our model, innovation is sequential (or cumulative). The literature on sequential innovation has studied various dimensions of patent policy, including patentability requirements (Scotchmer and Green, 1990; O'Donoghue, 1998), patent breadth and length (O'Donoghue *et al.*, 1998), forward protection (Denicolò, 2000; Denicolò and Zanchettin, 2002), or lack of protection (Bessen and Maskin, 2009). This literature has also covered aspects such as antitrust implications (Segal and Whinston, 2007), optimal buyouts schemes (Hopenhayn *et al.*, 2006), growth and industry dynamics (Denicolò and Zanchettin, 2014), and product-market competition (Marshall and Parra, 2019). We add to this literature by exploring the dynamic trade-offs that shape the interaction between patent screening and enforcement.

Our contribution is also related to the literature on entry efficiency and IPRs. Patent screening and enforcement limit appropriability in two ways: affecting the duration of the incumbency status (as in O'Donoghue *et al.*, 1998) and acting as an entry barrier to future innovations (Gilbert and Newbery, 1982). We show that these mechanisms are in tension, making imperfect patent screening desirable. In this respect, we complement the analysis of Parra (2019), which studies the optimal patent design when market structure is endogenously determined by (an exogenous) patent strength.

We also add to the literature that endogenizes the effective strength of patent rights. Caillaud and Duchêne (2011) study how congestion at the patent office leads to poorer screening, inducing low-quality innovators to apply for a patent. Atal and Bar (2014) explore the equilibrium consequences on R&D, patent quality, and welfare of introducing a two-tier patent system discriminating on innovation quality. Schankerman and Schuett (2021) study various policy instruments, such as screening intensity and screening fees, in a framework where patent challenges emerge endogenously, and court rulings are perfect. Picard and van Pottelsberghe de la Potterie (2013) analyze the interaction between patent screening and enforcement in a static framework. We complement previous contributions by uncovering the dynamic implications of imperfect patent enforcement on the optimal patent screening policy.

Finally, our research is also related to the debate among US scholars on whether the patent office has been "rationally ignorant" (Lemley, 2001; Lemley and Sampat, 2008) or "irrationally ignorant" (Frakes and Wasserman, 2019) when devoting limited resources to the review of patent applications, implying the grating of many invalid patents. Such a debate is based on comparing the direct costs of ex ante screening with the direct costs of ex post litigation. In our paper we abstract from those costs and identify an innovation enhancing effect associated with the presence of invalid patents that may justify the social desirability of not having a full screen-out of invalid patents by the patent office when patent enforceability is imperfect.

2 The Model

Consider a discrete-time infinite-horizon model of innovation in an industry comprised of a measure-one continuum of business niches. Time is denoted by t and agents' discount factor is $\beta < 1$. Each niche has an incumbent producer, protected by a patent, supplying the highest quality product available in that niche. In each period, potential new producers invest in R&D aiming to improve the quality of existing products and to enter a niche. As further explained below, successfully accessing the market requires obtaining a patent with which to defend the innovation from future imitators, as well as surmounting the opposition of incumbents claiming patent infringement. A patent office imperfectly screens the innovation's quality and decides whether to grant a patent, while decisions on patent infringement lawsuits are made by evidence-based courts with incomplete information about the quality of the innovation.

The Industry There is a measure-one continuum of homogeneous infinitely-lived consumers purchasing at each business niche. For each niche product, consumers have a unit-inelastic demand from which they derive utility of the form $U_t = Q_t - P_t$, where Q_t is the product's quality and P_t its price. Utility is additive across dates and niches. Within a niche, firms compete in price and through innovation in the context of a quality ladder. So, Q_t denotes the best available product quality at date t and P_t the equilibrium price charged for such a quality. Production costs are assumed to be zero.

Innovators In every period, an endogenous mass $m_t \in [0, 1]$ of potential innovators undertake an R&D investment with cost normalized to one to obtain a new product. Innovation is subject to two sources of uncertainty. First, it might be genuine or obvious with exogenous probabilities α and $1 - \alpha$, respectively. Second, it is untargeted, in the sense that there is uncertainty regarding over which of existing products it will improve upon. We capture that by assuming that each innovation may end up challenging any of the existing niches with equal probability.[5]

A genuine innovation constitutes a quality improvement $\pi > 0$ upon the incumbent product. When allowed to reach the market with the protection of a patent, the owner

[5]The assumption that innovation is untargeted implies that innovation undertaken by incumbents would be no different from innovation undertaken by outsiders, since incumbents would have a zero probability of innovating into their own niche. Our analysis abstracts, however, from incumbents' capability to undertake innovation directed to their own niche, an issue which would give rise to trade-offs and complexities which cannot be discussed within the limited length of this paper.

of a genuine innovation becomes the new incumbent, effectively monopolizing the corresponding niche. Price competition with the previous incumbent and the unit-demand assumption imply that when the innovator successfully enters, it earns a per-period profit flow of π while monopolizing the new top-quality product.[6]

An obvious innovation, in turn, provides a quality improvement of negligible size $\varepsilon < \pi$, meaning that the new product is an almost perfect substitute for the previous incumbent product. If allowed to access the market, an obvious innovator competes away the rents of the existing incumbent and, as the new incumbent, obtains per-period profits of ε.[7]

Obtaining a patent Innovators can apply for a patent at a cost assumed to be zero. We assume that holding a patent is essential for innovators to profit from their innovations because, otherwise, other producers would imitate the leading quality and compete away any rents, making market entry unprofitable.[8] With negligible patent application and entry costs, applying for a patent and attempting to enter the market if the application succeeds is the innovators' optimal decision regardless of their innovation's quality.[9]

Patents are granted by the patent office after a screening process. The patentability of an innovation requires usefulness, novelty, and non-obviousness. In our quality ladder formulation, we assume that genuine innovations satisfy the three requirements, while obvious innovations fail on their non-obviousness, since they are too similar to the previous top quality. We assume that the patent office always successfully identifies genuine inno-

[6]In equilibrium, genuine monopolists charge a price $P_t = \pi$ for their product capturing as profit the full quality improvement of their innovations, and generating no deadweight loss. Section 4.1 discusses an alternative environment where firms invest in cost-reducing innovations, and consumers' demand is not unit-inelastic. In that case, the exercise of market power implies a deadweight loss, altering the welfare implications but not our main results.

[7]In equilibrium, as with their genuine counterparts, obvious monopolists charge a price $P_t = \varepsilon$ for their product capturing as profit the full quality improvement of their innovations, and generating no deadweight loss.

[8]The model regards patents as the only means to sustain a technological lead. In practice, firms may profit from their innovations using first-mover advantages and trade secrets. The former are prevalent in markets with switching costs, network externalities, or reputation. The latter presents an unexplored trade-off in the context of sequential innovation: avoiding making the developed knowledge public versus a weaker position to enter and to defend the market from entry.

[9]Thus, it is irrelevant whether the innovator learns the quality of the innovation before or after accessing the market. If the patent application and entry costs were not negligible and the innovator learned the quality of the innovation after accessing the market, the strategy would remain optimal provided that π is large enough.

vations, granting patents to them, while wrongly granting patents to obvious innovations with probability λ. So λ measures the patent office's leniency towards obvious innovations. This *leniency rate* can be interpreted as the result of an insufficient search for prior art or an overestimation of the distance to such an art by the patent office.

Attempted entry and patent enforcement We assume that the resulting mass of new innovators granted a patent at date t, $[\alpha + (1-\alpha)\lambda]m_t$, gets uniformly distributed over the existing niches so that all have the same probability of being challenged by just one new innovator.[10] Leading incumbent firms in each niche lose their status whenever a subsequent innovator is allowed to access the market. An incumbent can respond to attempted entry by suing the potential entrant for the infringement of its patent. We assume that incumbents with genuine innovations sue their challengers, aiming to protect their profits π. In contrast, incumbents with obvious innovations accommodate entry without litigation, thus allowing the innovators to become the new incumbents.[11]

Courts make independent probabilistic evidence-based decisions on each patent infringement case. Courts rule in favor of obvious and genuine innovators with exogenous probabilities denoted by μ_0 and μ_1, respectively. Due to the sequential nature of innovation, court decisions boil down to assessing the size of the inventive step of the corresponding innovations relative to the incumbent's technology. Absent any noise in the assessment, genuine innovations should be regarded outside the breadth of the incumbent's patent, implying $\mu_1 = 1$, while obvious innovations should be regarded within that breadth, implying $\mu_0 = 0$. Reflecting noise in the court examination process, we assume instead that patent enforcement is imperfect and $0 < \mu_0 < \mu_1 < 1$.

[10]This formulation simplifies the analysis by abstracting from entry congestion—as found in the literature on random search—and patent races (e.g., Loury, 1979; Lee and Wilde, 1980).

[11]This could be justified assuming that innovators would succeed in persuading courts about the invalidity of the obvious incumbent's patent, voiding its infringement allegations against potential entrants, and making the legal defense of its incumbency unprofitable. In Section 4.3 we analyze the alternative situation in which obvious incumbents can engage in *patent trolling* to extract rents from subsequent innovators.

Licensing If the court rules that an innovation infringes the incumbent's patent, the incumbent and the infringer can engage in an ex-post licensing negotiation with probability $1 - \varphi$, where $\varphi \in [0, 1]$ captures exogenous frictions (e.g. financial constraints of the infringer) that prevent reaching a mutually beneficial agreement. We model the negotiation assuming that, at this point, both the incumbent and the infringer know the quality of the innovation, and the incumbent makes a take-it-or-leave-it offer to license its technology to the infringer who, if accepting it, becomes the new incumbent monopolist, exploiting its own innovation.[12] If the licensing offer is rejected or the negotiation does not take place, the incumbent keeps its status, and the infringer's innovation goes to waste.

Simplifying assumption and notation conventions For simplicity, we focus on the case where $\varepsilon \to 0$ thus abstracting from the value of obvious innovation throughout the analysis. Under this assumption, we can simplify terminology and refer to niches monopolized by a genuine innovator as *monopolized niches* and to the niches monopolized by an obvious innovator as *competitive niches*. We will denote the proportion of monopolized niches by $x_t \in [0, 1]$. The crucial difference between both types of niches is the hurdle that a subsequent innovator faces when attempting to enter them.

Under the assumption $\varepsilon \to 0$, incumbents in competitive niches extract zero value from their position in the market. The expression for the endogenous value of incumbency in a monopolized niche, which we denote by v_t, will be derived below.

Timing Figure 1 summarizes the timing of events in a given period from the perspective of an entering innovator (labeled E) that faces the potential legal opposition of an established incumbent (labeled I). In the industry as a whole, there is an initial stage in which an endogenous mass m_t of potential innovators invest in R&D in order to develop an innovation. For any of these innovators, the quality of their innovation is then realized, after which the patent office determines whether to grant a patent or not. Innovators that receive a patent attempt to enter their assigned niches. When the niche is

[12] In Section 4.4 we show the robustness of our results to the situation with asymmetric information in which only the infringer observes the realized quality of its innovation when this negotiation takes place.

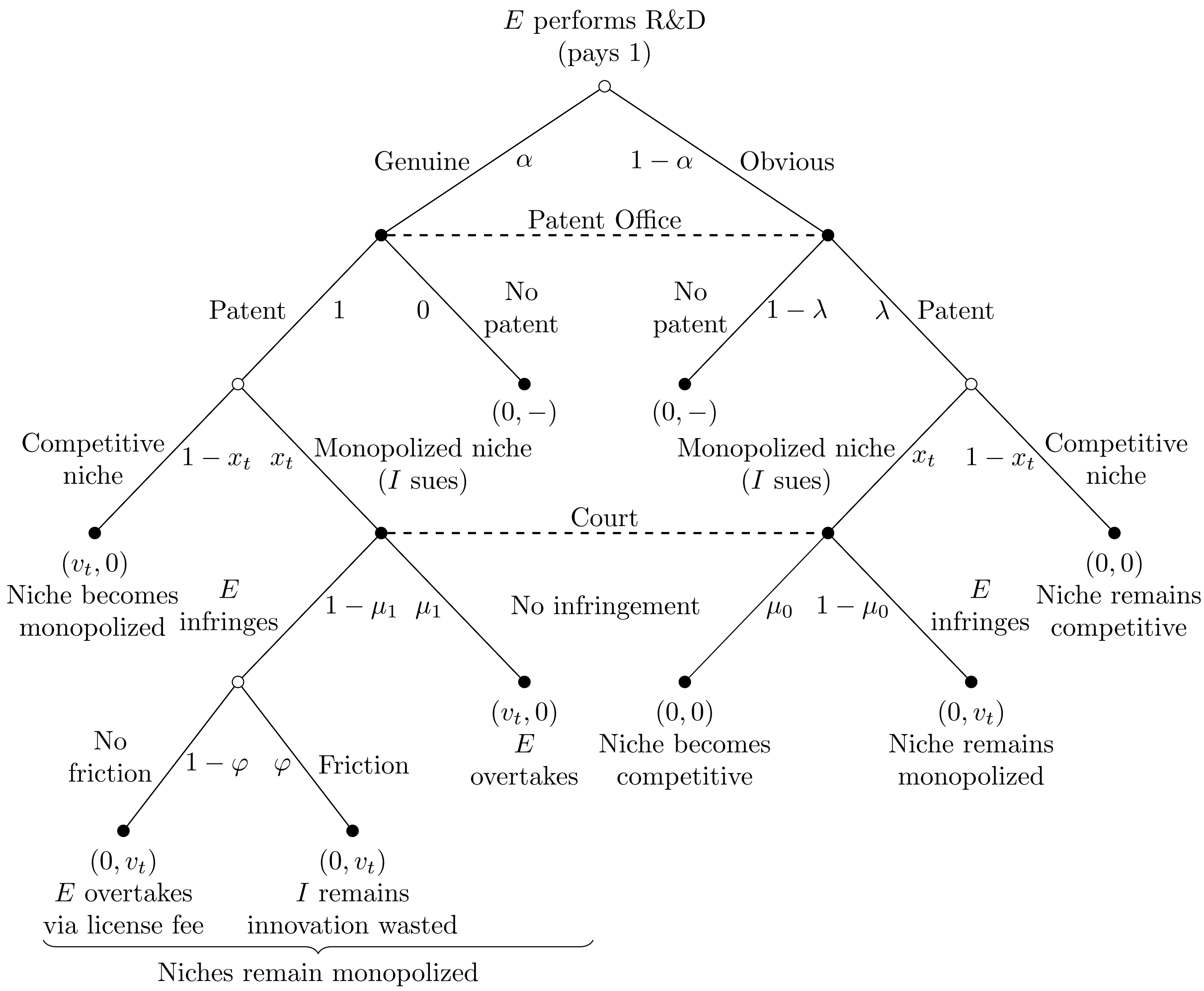


Figure 1: Game tree in the baseline model. The tree represents possible contingencies from the perspective of an entering innovator (E) in period t. Pairs at the terminal nodes describe the payoffs of the entering innovator and, if applicable, the incumbent (I) of the niche in which entry is attempted, in this order. Labels on branches describe probabilistic outcomes or decisions at each node.

monopolized by a prior genuine innovator, litigation ensues, and a court dictates whether the innovation infringes the incumbent's patent or not. Finally, if the court rules in favor of the incumbent and the innovation is genuine, ex-post negotiation on the licensing of the incumbent's patent to the entrant may ensue.[13]

As reflected in the terminal nodes of the tree in Figure 1, the sequence of events ends up with either the innovator replacing some previous incumbent or the incumbent preserving its prior position. The final outcomes across possible contingencies differ in terms of

[13]With $\varepsilon \to 0$, the negotiation when the innovation is obvious does not make sense since the infringer would not be willing to pay more than zero to enter the niche.

the turnover or not of the leading producer within each niche, whether the innovation is exploited or goes to waste, and the resulting monopolized versus competitive status of the niche. Below we analyze how these outcomes influence the innovation flow (m_t), the proportion of monopolized niches (x_t), the probability with which an innovation succeeds in accessing the market (p_t), the value of monopolist incumbency (v_t), and the welfare associated with the innovation process (W_t).

3 Equilibrium of the Industry

We can recursively characterize the value of an incumbent monopolist at date t using the following Bellman-type equation:

$$v_t = \pi + \beta\{1 - m_{t+1}[(1-\alpha)\lambda\mu_0 + \alpha\mu_1]\}v_{t+1}. \tag{1}$$

This value equals the sum of the current flow of monopoly profits π and the discounted value of having the possibility of maintaining the monopoly position one period later, βv_{t+1}. The incumbent has the opportunity to retain its monopoly position unless an innovator attempts to enter the niche and wins in court. Innovation at each niche happens with probability m_{t+1} and is of obvious or genuine quality with probabilities $1-\alpha$ and α, respectively. The terms in $\lambda\mu_0$ and μ_1 account for the compounded probabilities with which innovations of each quality obtain both a favorable assessment by the patent office and a positive court ruling.

Importantly, v_t is independent of whether licensing takes place or not after a genuine innovator is declared to infringe the incumbent's patent and, hence, independent of φ. To see this, notice that if the incumbent confronts an infringing genuine innovator at t+1, with probability φ a friction precludes licensing, so the incumbent remains in the niche, with continuation value v_{t+1}. With probability $1-\varphi$, the incumbent and the infringer negotiate for the latter to become the new incumbent. Since entry leads the incumbent to earn zero profits thereafter, it requests a payment of at least v_{t+1} to give up its incumbency. On the other hand, the infringer is willing to pay at most the value of being an incumbent in the niche. Hence, v_{t+1} is the only feasible licensing fee (and the incumbent's payoff) independently of whether licensing occurs or not. This also explains why if the infringer cannot pay v_{t+1} (e.g., because it has no liquid wealth and borrowing against v_{t+1} is limited by incentive issues), licensing will not take place and its innovation will go to waste.[14]

The transitions across the two classes of niches, contingent on each possible realization of the relevant uncertainty in each period are summarized in Figure 1. Accordingly, the law of motion governing the evolution of the proportion of monopolized niches, x_t, between t and $t+1$ can be described as follows:

$$x_{t+1} = x_t[1 - m_{t+1}(1-\alpha)\lambda\mu_0] + (1-x_t)\alpha m_{t+1}, \tag{2}$$

where the first term represents the proportion of niches already monopolized at t that remain monopolized at t+1. These are all the monopolized niches that do not experience the successful entry of obvious innovators since only that entry turns them competitive. The second term in the right hand side of (2) measures the previously competitive niches that become monopolized due to the entry of genuine innovators.

The equilibrium innovation flow m_t is determined by a free-entry condition. Developing an innovation has a cost normalized to 1. Innovation leads to a monopolistic incumbency of value v_t, only if innovation is genuine and the potential opposition of the previous incumbent is surmounted; otherwise the innovator obtains zero. Thus, an innovator obtains the value v_t with probability

$$p_t = \alpha(1 - x_t + x_t\mu_1). \tag{3}$$

That is, the innovation has to be genuine and the firm must either land in a competitive niche or in a monopolized niche and innovator wins in court, which occurs with probability

[14]In Section 4.2, we consider the case in which the incumbent monopolist can buy out the genuine innovation from the infringer, avoiding that it goes to waste. We exclude this possibility in the baseline case studied here because the accumulation of monopoly power on subsequent innovations by an incumbent might be questionable from an anti-trust perspective.

μ_1. The free entry of innovators implies that in an equilibrium involving an interior innovation flow $m_t \in (0,1)$ we must have

$$p_t v_t = 1, \tag{4}$$

so that the innovator obtains zero net present value from the investment. As in the case of v_t, licensing frictions do not affect this free-entry condition since, as already explained, irrespectively of reaching a deal or not with the incumbent, the genuine entrant is left with a zero payoff from entry if losing in court.

3.1 Steady-State Equilibrium

Our analysis focuses on the interior-innovation equilibrium of the model characterized by equations (1)–(4). They determine four key endogenous variables at each date t: x_t, p_t, m_t, and v_t. To ease notation, we denote the steady-state value of these variables simply by x, p, m, and v, respectively.

Lemma 1. *The model has a unique steady-state equilibrium with an interior innovation flow $m \in (0,1)$ if and only if $\pi \in \left(\underline{\pi}, \underline{\pi} + \beta x^{-1}\right)$, where $\underline{\pi} = (1-\beta)/p$. This equilibrium is given by*

$$x = \frac{\alpha}{\alpha + (1-\alpha)\lambda\mu_0}, \tag{5}$$

$$p = \alpha \frac{\alpha\mu_1 + (1-\alpha)\lambda\mu_0}{\alpha + (1-\alpha)\lambda\mu_0}, \tag{6}$$

$$m = \frac{\pi p - (1-\beta)}{\beta\left[\alpha\mu_1 + (1-\alpha)\lambda\mu_0\right]}, \tag{7}$$

$$v = \frac{1}{p}. \tag{8}$$

In this equilibrium, innovation flows in until the expected value of developing an innovation, pv, equals the unit research cost. Intuitively, a higher innovation flow increases the speed at which successful innovators get replaced by subsequent ones, reducing the value of incumbency. A higher value of the profit parameter π has a positive impact on the equilibrium value of m, while it does not affect any of the other equilibrium variables highlighted in Proposition 1. This means that the impact of π on the value of incumbency is fully offset by the change in the innovation flow m, without altering the steady state values of the proportion of monopolized niches x or the probability of successful entry p, whose steady-state values are fully determined by parameters describing the proportion of genuine and obvious innovations, α, the leniency of the patent office towards obvious innovations, λ, and the rates at which courts rule in favor of innovations of each type, μ_0 and μ_1.[15]

It is also worth noting that parameters λ and μ_0 always appear in combination, as $\lambda\mu_0$, which represents the rate at which obvious innovations, if landing in a monopolized niche, succeed in replacing the incumbent monopolist. Notice also that the parameter φ capturing frictions in licensing negotiations between incumbent monopolists and infringing genuine innovators does not affect any of the variables highlighted in Proposition 1. However, φ will have an impact on the welfare calculations below, since it affects the probability with which infringing genuine innovations go to waste.

Assumption 1. *We assume $\pi \in (\underline{\pi}, \underline{\pi} + \beta x^{-1})$ for every relevant value of the leniency rate λ.*

Assumption 1 restricts the profit parameter π so that the steady-state equilibrium of the model involves an interior entry flow $m \in (0, 1)$ for every relevant value of the policy parameter λ. By fully differentiating equations (1)–(4), we obtain the comparative statics of the steady-state equilibrium summarized next.

Corollary 1. *The effect of marginal changes in the parameters on the steady-state equilibrium variables x, p, m, and v have the signs shown in the following table:*

	π	β	α	λ	μ_0	μ_1	φ
x	0	0	+	−	−	0	0
p	0	0	+	+	+	+	0
m	+	+	+	?	?	+	0
v	0	0	−	−	−	−	0

Besides the already commented impact of parameters π and φ, the table shows that the effects of the discount factor β are qualitatively identical to those of the profit parameter

[15]Throughout the paper we interpret having $\mu_0 > 0$ and $\mu_1 < 1$ as saying that courts are imperfect. The imperfection might come from pure mistakes or from systematic biases (as in Jacobsmeyer, 2018).

π. We now turn to the effect of the remaining parameters on each of the equilibrium variables.

The proportion of niches operating under monopoly, x, is increasing in the probability that an innovation is genuine, α, and decreasing in the rate at which obvious innovations succeed in accessing monopolized niches, $\lambda\mu_0$. Intuitively, the higher the probability that a firm with an obvious patent arises and is allowed to produce, $(1-\alpha)\lambda\mu_0$, the more often monopolist incumbents lose their positions. In contrast, the probability with which genuine innovators succeed in court vis-a-vis an incumbent patent holder, μ_1, does not affect x since the niche will remain in any case monopolized by either the prior incumbent or the genuine innovator.

Due to the free-entry condition, the value of incumbency v is inversely related to the probability with which innovators profitably reach the status of new monopolists, p. Such probability increases in the proportion of innovations that are genuine, α, and the probability that courts rule in favor of genuine innovators when confronting a monopolist incumbent, μ_1. More surprisingly, p is also increasing in the rate at which obvious innovations reach monopolized niches, $\lambda\mu_0$, even if this does not give them the status of incumbent monopolists. This effect occurs because this entry reduces the proportion of monopolized niches x and, with it, the probability that a genuine innovator faces opposition from an incumbent monopolist. This pro-innovation effect of the success of obvious innovation explains the ambiguous sign attached to the impact of $\lambda\mu_0$ on the innovation flow, which we will further analyze below.

The effects of other parameters on the innovation flow are standard. Together with the profit parameter π and the discount factor β, an increase in the courts' probability of ruling in favor of a genuine innovator, μ_1, or in the probability of obtaining a genuine innovation, α, also fosters the investment in innovation, as it increases the probability of profitably reaching the market.

The next proposition fully characterizes the impact of the patent office's leniency rate

λ on the innovation flow.[16]

Proposition 1 (Screening leniency promotes innovation)**.** *The relationship between the innovation flow m and the patent office's leniency rate λ is single-peaked. Specifically, there exists a threshold $\widehat{\mu}_1 = 1/\sqrt{\alpha\Pi} \in (0,1)$ such that the innovation-maximizing leniency rate involves: i) full screening, $\lambda_m = 0$, if $\mu_1 \geq \widehat{\mu}_1$, and ii) screening leniency with*

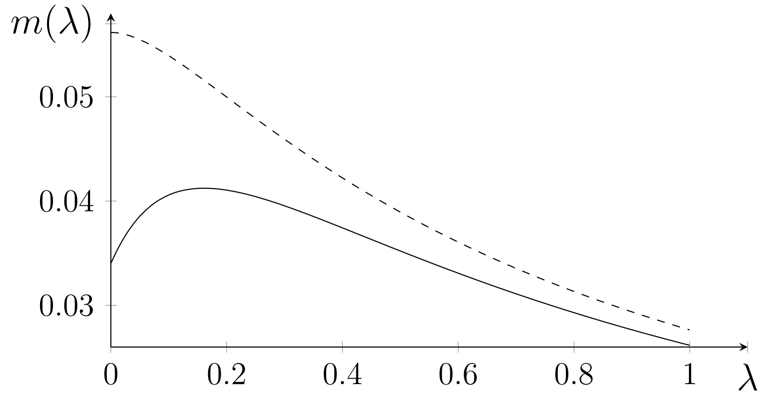
$$\lambda_m = \frac{\alpha\big((1-\mu_1)\sqrt{\alpha\Pi} - (\alpha\Pi\mu_1 - 1)\big)}{(1-\alpha)\mu_0(\alpha\Pi - 1)} > 0, \tag{9}$$

if $\mu_1 < \widehat{\mu}_1$.

This result uncovers an interesting non-monotonic relationship between innovation and the leniency rate. The main driver of this result is that a change in λ engenders two effects with opposite signs. On the one hand, a higher leniency rate reduces, as noted above, the proportion of monopolized niches x in which genuine innovators are challenged in court, reducing the innovators' hurdle to successfully enter the market. On the other hand, a higher leniency decreases the value of incumbency, v, increasing the probability that successful innovators once established as monopolist incumbents see their rents competed away by an entrant with an obvious innovation.

When the probability of success in court of an entrant with a genuine innovation, μ_1, is close to one, the first effect becomes irrelevant (since incumbents' opposition is futile against genuine innovators), and innovation monotonically decreases with λ. In contrast, when μ_1 is sufficiently low, the entry-facilitating effect associated with having a lower proportion of monopolized niches dominates for lower values of the leniency rate giving rise to an interior innovation-maximizing value of λ. This is illustrated in Figure 2. As we will show next, this result has important implications for the discussion on the socially optimal level of patent-office leniency towards obvious innovations λ in contexts where the patents of prior genuine innovators impose a significant hurdle (a low μ_1) to the successful market access of subsequent genuine innovators.

[16] As previously noted, a parallel result could be established in terms of the parameter μ_0 as it always enters the relevant expressions multiplied by and together with λ



Note: Parameter values are $\alpha = 0.1$, $\pi = 0.3$, $\beta = 0.98$ —implying $\widehat{\mu}_1 = 0.816$— and $\mu_0 = 0.25$.

Figure 2: Impact of screening leniency on innovation. The solid line depicts a case (with $\mu_1 = 0.75$) in which the innovation flow $m(\lambda)$ is maximized under strictly positive screening leniency, $\lambda > 0$. The dashed line shows the case with $\mu_1 = \widehat{\mu}_1 = 0.816$ in which full screening maximizes innovation.

Corollary 2. *When $\mu_1 < \widehat{\mu}_1$, λ_m decreases in both μ_0 and μ_1.*

At entry-maximizing leniency, both genuine and obvious entrant success rates are substitutes for leniency. Since only obvious entrants with a patent can turn a niche competitive, a higher success rate in court, μ_0, can be compensated with a smaller leniency, creating the same rate of competitive niche creation and the same entry rate. The pro-competitive effect of allowing obvious innovations is only required to compensate for the mistake of not allowing genuine entrants. As the success rate of genuine entrants increases, μ_1, less leniency is needed.

3.2 Optimal Patent Screening

The leniency rate λ affects welfare through the innovation flow m and the rate at which genuine innovations access the market, allowing society to benefit from the product quality improvement π that they originate. In our quality ladder model, those improvements become permanent welfare gains either in the form of producers' profits (while the genuine innovators operate as incumbent monopolists) or consumers' surplus (once those innovators lose their incumbency).

The present value of the social welfare created by a genuine innovation, conditional on market access, is given by $\Pi \equiv \pi/(1-\beta)$. The probability with which an innovation is genuine and obtains access the market is, in turn, given by $\alpha[1 - x(1-\mu_1)\varphi]$ since this access only fails if the innovation is considered to have infringed an incumbent's patent and the subsequent licensing negotiations fail. The total per-period net addition to welfare associated with the innovation flow in steady-state can be expressed as

$$W(\lambda) = m(\lambda)\left\{\alpha\left[1 - x(\lambda)(1-\mu_1)\varphi\right]\Pi - 1\right\}, \tag{10}$$

where we are making explicit the dependency of equilibrium variables m and x on λ.[17] Assumption 1 implies that the expression in braces in (10) is always positive.[18] Thus, welfare is positive.

The derivative of the welfare function with respect to the leniency rate λ is

$$W'(\lambda) = m'(\lambda)\left\{\alpha\left[1 - x(\lambda)(1-\mu_1)\varphi\right]\Pi - 1\right\} - \alpha(1-\mu_1)\varphi\Pi m(\lambda)x'(\lambda). \tag{11}$$

Leniency affects welfare through two channels: the innovation flow, $m(\lambda)$, and the probability of landing in a monopolized niche, $x(\lambda)$, since for $\varphi > 0$ that implies a positive probability that the genuine innovation goes to waste. When this second entry-facilitating channel is not socially relevant ($\varphi = 0$), maximizing welfare is equivalent to maximizing the innovation flow, and the results in Proposition 1 are directly applicable to the socially optimal leniency rate, which coincides with λ_m.

When the entry-facilitating channel is relevant ($\varphi > 0$), Corollary 1 implies $x'(\lambda) < 0$, making the second term in (11) overall positive, and pushing, other things equal, towards a higher socially-optimal leniency rate. Reducing the proportion of monopolized

[17] We can obtain the same expression for W as the sum of the additions of consumer surplus, ΔCS, and producer surplus, ΔPS, that occur in every period. We then have $W = \Delta CS + \Delta PS$, where

$$\Delta CS = m(\lambda)x(\lambda)\left\{\alpha[\mu_1 + (1-\mu_1)(1-\varphi)] + (1-\alpha)\lambda\mu_0\right\}\Pi = m(\lambda)\left\{\alpha\left[1 - x(\lambda)(1-\mu_1)\varphi\right]\Pi\right\},$$

since consumers receive the surplus from prior genuine innovations every time the incumbent of a monopolized niche loses its incumbency. On the other hand, in steady state, the fraction of monopolized niches remains constant and, hence, also the overall surplus appropriated by producers conditional on entry. However, innovators pay an aggregate cost $m(\lambda)$ per period. So $\Delta PS = -m(\lambda)$, which combined with the expression for ΔCS implies (10).

[18] Specifically, Assumption 1 implies $\pi > (1-\beta)p^{-1}$. Replacing the profit lower-bound into the square brackets we obtain $\alpha\left[1 - x(\lambda)(1-\mu_1)\varphi)\right]\Pi - 1 > \alpha(1-\mu_1)(1-\varphi)/[\alpha\mu_1 + (1-\alpha)\lambda\mu_0] \geq 0$.

niches beyond the level that maximizes the innovation flow, further reduces the chances of confronting an incumbent monopolist that, by succeeding in court, would prevent a genuine innovation from being implemented.

When the innovation-maximizing leniency rate is just zero, the second effect may not be strong enough to make the welfare-maximizing leniency rate strictly positive (this is clearly the case from $\mu_1 \to 1$, when the second term in (11) approaches zero). In contrast, when λ_m is strictly positive, the entry-facilitating channel makes the socially optimal leniency rate strictly higher than λ_m.

Proposition 2 (Screening leniency enhances welfare)**.** *Social welfare is maximized at a unique value of the patent office's leniency rate $\lambda_w(\varphi)$. Without licensing frictions, this rate equals the innovation-maximizing rate λ_m. Otherwise, the welfare-maximizing leniency rate $\lambda_w(\varphi)$ is weakly larger than λ_m. In particular, if λ_m is interior, then the welfare-maximizing rate entails a strictly larger screening leniency, ($\lambda_w(\varphi) > \lambda_m > 0$), for $\varphi > 0$. Moreover, the optimal leniency increases as the licensing frictions increase.*

So far we have assumed that patent screening is costless. If screening were costly and, in particular, if the screening cost function satisfied Inada-type conditions with respect to the precision rate $1 - \lambda$, the optimal leniency rate would always be positive. Proposition 2 goes further by saying that, even if screening is costless, there may be circumstances (specifically if μ_1 is sufficiently low) in which granting patents to some obvious innovations enhances welfare.

Propositions 1 and 2 point to two complementary forces behind this result. Both forces originate from how the access of obvious innovators to the market reduces the proportion of monopolized niches x and, through them, the entry hurdle of subsequent genuine innovators when $\mu_1 < 1$. The first force, by reducing such a hurdle, can boost innovation (for $\mu_1 < \widehat{\mu}_1$). The second force, which is only relevant in the presence of licensing frictions ($\varphi > 0$), by reducing x it additionally reduces the proportion of genuine innovations that go to waste (arising because a license from the incumbent monopolist is not obtained after being found to infringe its patent). Thus, a positive level of screening leniency may be socially optimal even with $\varphi = 0$ and it increases with φ.

These implications of our analysis can be seen as the result of a model à la Dixit and Stiglitz (1977) that features insufficient entry due to the innovators' inability to fully appropriate the value of their innovations. In our model, genuine innovators may fail to appropriate the value of investment either because of they are blocked by the patent rights of previous innovators or because they might be displaced by subsequent innovators. Protecting incumbents against genuine entrants would damage the appropriability of innovations' value through the first channel but increase it through the second channel. Due to discounting, however, the first channel dominates, making it socially optimal to have $\mu_1 = 1$, as in the standard sequential innovation setup (e.g., Hopenhayn *et al.* (2006)).

When $\mu_1 = 1$, no hurdle to the market access of genuine innovations exists and the entry-facilitating motive to be lenient towards obvious innovators disappears. Hence, the optimal λ becomes zero (or equivalently, the optimal μ_0 becomes zero). More generally, if the barrier-to-entry effect is sufficiently small ($\mu_1 \to 1$), the optimal screening strategy consists of thoroughly screening out low-value innovations, consistent with the classical results in Green and Scotchmer (1995) and O'Donoghue *et al.* (1998). The following corollary states these points.

Corollary 3. *Social welfare is maximized by setting no hurdle to genuine innovations,* $\mu_1 = 1$. *In that case, it is also socially optimal to fully weed out obvious innovations either with full ex-ante screening,* $\lambda = 0$, *or by blocking them in court,* $\mu_0 = 0$.

In contrast, if courts are fallible ($\mu_1 < 1$), the presence of incumbent monopolists acts as a barrier-to-entry for genuine innovators. In this case, for the entry-facilitating reasons already explained at length above, it can be socially optimal that both the patent office and courts are lenient towards obvious innovations ($\lambda > 0$ and $\mu_0 > 0$).

4 Robustness and Extensions

In this section, we develop several extensions of the baseline model and show that the main insights from our previous analysis are robust to these changes. First, we consider the case in which innovation reduces production costs (rather then improving product quality) and demand is not unit-inelastic, in which case the exercise of monopoly power entails a static inefficiency. The second extension develops the case in which incumbents can buy infringing innovations, preventing them to go to waste but also extending their monopoly power over multiple steps in the quality ladder. The third extension considers the situation in which successful obvious innovators can act as patent trolls. That is, they can use the threat of an infringement court ruling against subsequent genuine innovators as a way to extract rents. The fourth and final extension analyzes the effects of asymmetric information as well as the usage of out-of-court settlements.

4.1 Cost-saving Innovations and Static Inefficiency

The baseline model considers quality-improving innovations in an environment where demand is inelastic. In that context, genuine innovators during their incumbency as monopolists extract all the surplus from consumers, avoiding static deadweight losses and simplifying the welfare analysis. We now provide a tractable framework, inspired in Marshall and Parra (2019), that extends our analysis to a case of cost-reducing innovations where the exercise of market power involves welfare losses.

In this setup, there is also a continuum of niches of size 1. In each niche, the good produced is homogeneous. Firms compete in prices, facing a demand function $q(p) = a/p$. Each genuine innovation decreases the existing marginal cost by a proportion $1-\delta$ where $\delta \in (0,1)$. That is, if z represents the baseline marginal cost, after i genuine innovations the marginal cost becomes $z_i = \delta^i z$.

Lemma 2. *In a sequential cost-saving innovations setup, the profit flow π and the deadweight loss ℓ generated by a genuine innovation are independent of the baseline marginal*

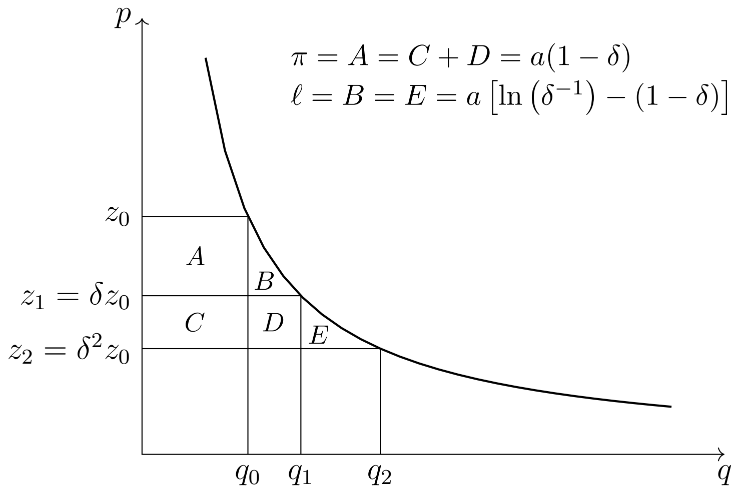


Figure 3: Monopoly profits and deadweight losses with cost-saving innovations. This figure identifies as areas under the relevant demand curve the profits (π) and deadweight losses (ℓ) associated with two subsequent successful genuine cost-saving innovations. The innovator reducing the marginal cost from z_0 to z_1, monopolizes the market by selling at price $p = z_0$ (the cost of the next most efficient producer), generating profits A and deadweight losses B. Likewise, the innovator reducing the marginal cost to z_2 sells the product at price z_1, generating profits $C + D = A$ and deadweight losses $E = B$.

cost z and the cumulative number of innovations i. They are equal to $\pi = a(1-\delta)$ and $\ell = a\left[\ln(\delta^{-1}) - (1-\delta)\right] > 0$, as illustrated in Figure 3.

Under this formulation, because profits are invariant to the cumulative number of innovations, firm behavior and the industry dynamics described in Section 3 go through without alterations, generating the same equilibrium conditions and expressions for the steady-state equilibrium variables reported in Lemma 1.

The expression for the welfare function, however, needs to be revised to account for the deadweight loss ℓ. In particular, the main difference with respect to the baseline model is that the successful access of any innovator to a monopolized niche converts the deadweight loss ℓ into consumer surplus. Thus, social welfare increases by ℓ on a permanent basis. The resulting revised expression for social welfare is provided in the following lemma.

Lemma 3. *Social welfare in the cost-saving innovations setup can be written as*

$$W = m(\lambda)\left\{\alpha[1 - x(\lambda)(1-\mu_1)\varphi](\Pi + L) - 1\right\}, \tag{12}$$

where $L \equiv \ell/(1-\beta)$.

The new term L in W captures the gain of consumer surplus associated with the removal of prior monopoly power due to the successful market entry of genuine and obvious innovators. Expression (12) coincides with (10) when $L = 0$.[19] The next proposition shows the implications of the presence of this deadweight loss for the socially optimal leniency towards obvious innovations.

Proposition 3 (Optimal leniency and static inefficiency)**.** *The unique welfare maximizing leniency rate,* $\lambda_d(\varphi, L)$*, i) is contained in the interval* $[\lambda_m, \lambda_w(\varphi)]$*; ii) equals the innovation-maximizing leniency rate,* λ_m*, when there are no licensing frictions (*$\varphi = 0$*) or genuine innovators are always allowed by courts (*$\mu_1 = 1$*); iii) strictly exceeds the innovation-maximizing rate (*$\lambda_d(\varphi, L) > \lambda_m$*) when* $L > 0$ *and the innovation-maximizing rate entails screening leniency,* $\lambda_m > 0$*; iv) converges to the welfare-maximizing leniency rate without static inefficiency when there is no deadweight loss (*$\lambda_d(\varphi, 0) = \lambda_w(\varphi)$*); and v) increases with the licensing friction* φ *and decreases in the deadweight loss* L*.*

Unsurprisingly, the socially optimal leniency rate is the same as in the benchmark setup when $L = 0$. Moreover, as in such a setup, the socially optimal leniency rate coincides with the innovation maximizing one when there are no licensing frictions ($\varphi = 0$) or legal hurdle to the entry of genuine innovation is null ($\mu_1 = 1$). More interestingly, when screening leniency contributes to maximizing innovation ($\lambda_m > 0$), the welfare maximizing leniency rate entails strictly more leniency ($\lambda_d(\varphi, L) > \lambda_m$) but strictly less than in the absence of static inefficiency ($\lambda_d(\varphi, L) < \lambda_w(\varphi)$). This result reveals that the cost of static inefficiency, L, pushes the socially optimal leniency rate closer to the innovation maximizing rate, as innovation favors the turnover of incumbent monopolists and, thus,

[19] L enters (12) in a way parallel to π, which is the present value of the profit flow associated with genuine innovation. This is so in spite of the fact that the removal of market power from prior incumbent monopolists and the subsequent recovery of consumer surplus L i) does not occur when genuine innovators access a non-monopolized niche, and ii) occurs when obvious innovators access a monopolized niche. It turns out, however, that the rate at which these two processes occur is endogenously equalized in steady state (so that the fraction of monopolized niches remains constant), explaining the result.

the recovery of deadweight losses from their exercise of market power. As a result, as the deadweight loss L increases, the optimal leniency $\lambda_d(\varphi, L)$ approaches the innovation maximizing leniency rate, λ_m, from above. Favoring innovation and turnover among incumbent monopolists (relative to the case with $L = 0$) increases welfare even if that implies an increase in the equilibrium proportion of monopolized niches, x.

4.2 Buying Out Infringing Innovations

In the benchmark model we assume that when courts rule against a genuine innovator, this firm can license the incumbent's technology to enter the niche and sell its higher-quality product. In this scenario, the innovator faces the latent competition of the old incumbent, thus obtaining a profit flow equal to π. In this section, we allow for an alternative arrangement where the incumbent buys out the innovation of the infringing genuine innovator who then exits the market. As a result, the incumbent expands the quality gap with its closest competitor, increasing its profit flow by an extra π. Over time, the incumbent piles up the improvements of genuine innovators defeated in court until an innovator eventually defeats the incumbent. In contrast with the baseline model, here we assume that the incumbent always has the resources to buy out the entering innovation, avoiding the frictions in the baseline model.[20]

We start by recursively characterizing the present value of profits of an incumbent who has accumulated n improvements at time t as

$$v_t^n = n\pi + \beta\left(\{1 - m_{t+1}\left[\alpha + (1-\alpha)\lambda\mu_0\right]\} v_{t+1}^n + m_{t+1}\alpha\left(1-\mu_1\right) v_{t+1}^{n+1}\right). \qquad (13)$$

This expression indicates that, at t, the incumbent collects per-period profits of $n\pi$. At $t+1$, three things may occur. As in the benchmark model, if no innovator challenges the niche, the incumbent preserves the existing monopoly, and its continuation value

[20] Throughout the paper we have interpreted $\varphi > 0$ as the result of financial constraints that could prevent innovators from licensing the incumbent's technology. At the same time, in the benchmark model, licensing did not result in any private gains, implying that firms had no real incentives to reach an agreement. In contrast, in this case, when an incumbent buys a genuine innovation, per-period producer surplus increases by π, encouraging firms to overcome licensing frictions and reinforcing our assumption in this case.

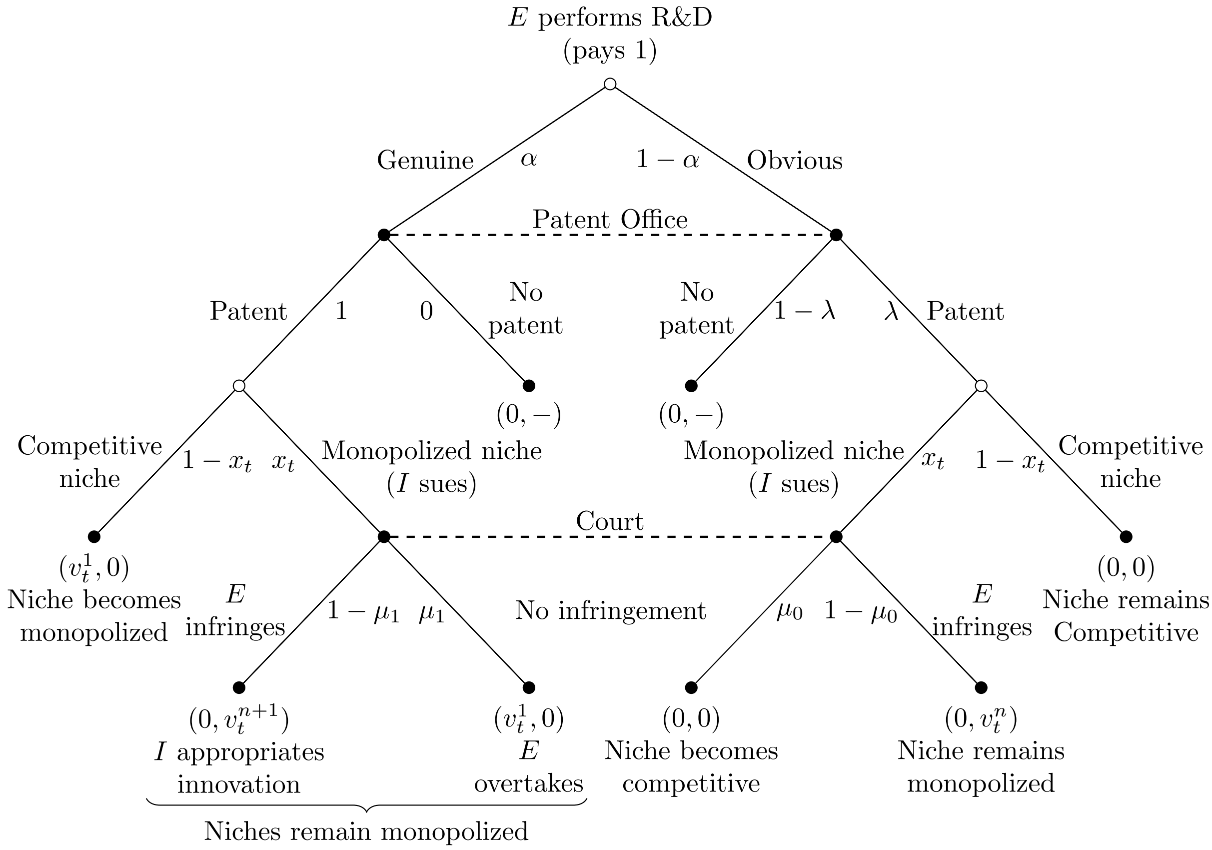


Figure 4: Game tree when the incumbent buys out infringing innovations. Tree depicted for the case in which the incumbent of the monopolized niche in which innovation may occur has already accumulated n innovations.

becomes v_{t+1}^n. If a genuine innovator is considered to infringe the incumbent's patent, which occurs with probability $m_{t+1}\alpha(1-\mu_1)$, the incumbent makes a take-it-or-leave-it offer equal to the innovator's outside option, equal to 0. This results in a future value of v_{t+1}^{n+1}. Finally, whenever the court rules in favor of an entering innovator, obvious or genuine, the incumbent loses the value of incumbency and its payoff becomes zero. Figure 4 specifies all the cases in the game tree.

In a steady state, where $v_t^n = v_{t+1}^n = v^n$ for all t and n, we can verify, after guessing that v^n is linear in n, that the following expression for v^n solves (13) for all n:

$$v^n = \frac{\pi n}{1-\beta[1-m(\alpha\mu_1+(1-\alpha)\lambda\mu_0)]} + \frac{\pi m\alpha\beta(1-\mu_1)}{\left(1-\beta\left\{1-m\left[\alpha\mu_1+(1-\alpha)\lambda\mu_0\right]\right\}\right)^2}. \quad (14)$$

This expression is linear in n. Differently from the baseline model in which (1) implies a negative impact of the innovation flow on the value of incumbency, here the effect of

m on v is ambiguous. The reason is that a higher innovation rate, m, has two effects pointing in opposite directions. It increases the probability that an innovator challenges the incumbent and wins in court, decreasing the duration of incumbency and its value. At the same time, it also increases the probability that a genuine innovator is found to infringe the incumbent's patent and the incumbent can benefit from buying out the innovation, increasing its profit flow. It is easy to check that $\mu_1 > 1/2$ is a necessary and sufficient condition for the first effect to strictly dominate for all values of λ and m.

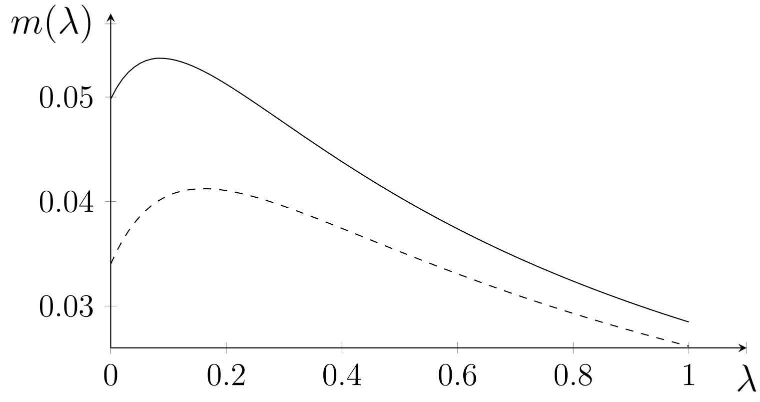
The law of motion of the proportion of monopolized niches x_t and the probability with which an innovator is genuine and successfully reaches the market p_t are still described by (2) and (3), respectively. In steady-state, the proportion of monopolized niches, x, and the probability that innovation leads to profits, p, have the same steady-state values as in the baseline model. The free entry condition determining the steady-state innovation flow, if interior, is now $pv^1 = 1$. Replacing p from (6) and v^1 from (14) in this expression, we can obtain the steady-state innovation flow as

$$m(\lambda) = \frac{(1-\beta)}{2\beta\left[\alpha\mu_1 + (1-\alpha)\lambda\mu_0\right]}\left\{(\Pi\alpha - 2) + \sqrt{\Pi\left[\Pi\alpha^2 - 4(\alpha - p)\right]}\right\}. \tag{15}$$

Assumption 1 guarantees that the expression within the square root is non-negative and that the innovation flow is positive. The next result compares this rate with its counterpart in the baseline model.

Proposition 4 (Incumbent buyouts increase innovation)**.** *Under imperfect enforcement ($\mu_1 < 1$), the innovation flow when the incumbent can buy out infringing innovations is larger than when the incumbent licenses its patent to the entrant.*

As Figure 5 illustrates, the innovation flow in this case can also reach a maximum at an interior value of the leniency rate. For all values of λ, the innovation flow is higher than in the baseline model. Intuitively, for any given level of the innovation flow, the odds that an innovator appropriates the value of being an incumbent monopolist are the same as in the baseline model, while the value of incumbency is higher due to the possibility of



Note: Parameter values are $\alpha = 0.1$, $\pi = 0.3$, $\beta = 0.98$, $\mu_1 = 0.75$, and $\mu_0 = 0.25$.

Figure 5: Innovation when the incumbent buys out infringing innovations. This figure shows a case in which the innovation flow is maximized by a positive leniency rate λ both in the baseline model (dashed line) and with incumbent buyouts (solid line).

accumulating monopoly rents over several subsequent innovations. This spurs innovation, making v^1 decline until the free entry condition holds again.

Notice that, since in this scenario there are no licensing frictions, the welfare function (10) becomes $W(\lambda) = m(\lambda)(\alpha\Pi - 1)$. Consequently, the welfare-maximizing leniency matches the innovation-maximizing leniency. As illustrated in Figure 5, there are values of the parameters of the model for which setting a strictly positive leniency rate λ remains socially optimal.

4.3 Patent Trolling

In the baseline model, niches occupied by incumbents with obvious patents are competitive and assumed to create no barrier to entry for subsequent innovators. We justified this assumption on the absence of incentives for producers in these niches to legally defend a market position that generates no rents. However, this assumption implies that they cannot use the patents on their obvious innovations as a means to extract rents from future genuine innovators.[21]

[21] This baseline scenario could be further justified by arguing that entrants to niches occupied by prior obvious innovators would succeed in making courts declare the incumbent innovations obvious and hence invalid in the first place.

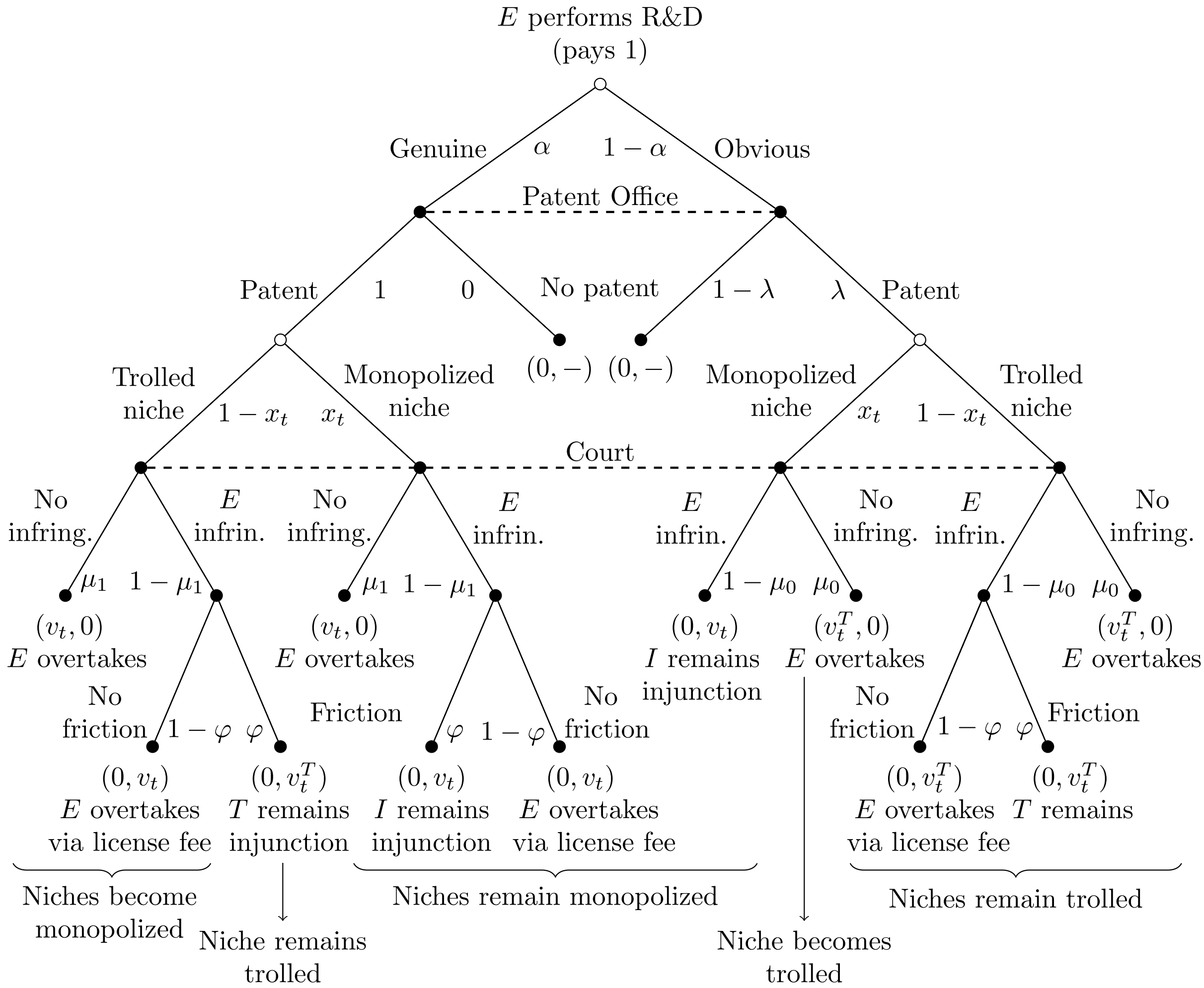


Figure 6: Game tree under patent trolling. Patent trolling in the model describes the situation in which incumbents with prior obvious innovations (in "trolled niches") are able to extract rents from subsequent genuine innovators if they win a patent infringement case against them.

In this section, we explore the opportunistic behavior that arises when these firms, denoted as patent trolls, are able to win their patent infringement case in court and, if that happens, force genuine innovators to pay a licensing fee to occupy their niche. This behavior implies a transfer from genuine innovators to trolls (the former obvious innovators). This reduces the prospect of profitably entering a niche as a genuine innovator but simultaneously increases the value of entering as an obvious innovator. Such a value was zero in the baseline model and will be denoted by v_t^T in this extension. Beyond that, the equations describing the law of motion of the fraction of monopolized niches, x_t, and the value of being an incumbent monopolist, v_t, remain unchanged.

We assume that trolls can make a take-it-or-leave-it offer to those genuine innovators who, after attempting to enter their niche, are considered to be infringers by courts, which occurs with probability $1-\mu_1$. The value of trolling will then satisfy the following recursive expression:

$$v_t^T = \beta \left\{ \alpha m_{t+1} (1-\mu_1)(1-\varphi) v_{t+1} + [1 - m_{t+1} (\alpha + (1-\alpha) \lambda \mu_0 - \alpha \varphi (1-\mu_1))] v_{t+1}^T \right\}. \tag{16}$$

Consistently with the game tree depicted in Figure 6, a troll in period $t+1$ can only extract a payoff v_{t+1} from licensing its patent when a genuine innovator enters its niche, is defeated in court, and the licensing negotiation succeeds. This explains the first term inside the curly brackets in (16). The second term accounts for the value associated with the probability of remaining as a troll, which is one minus the probability with which the troll licenses its patent or losses in court against an entering innovator of either type.

The new free-entry condition involves two modifications relative to the baseline model. First, an innovator which turns out to be genuine has a lower probability of appropriating the value of monopoly incumbency; this probability is now just μ_1 independently of whether the entered niche was monopolized by a prior genuine incumbent or occupied by a troll. Second, an innovator with an obvious innovation which successfully enters a niche (which happens with probability $\lambda\mu_0$) obtains the value of being a troll. The resulting free-entry condition is then

$$\alpha \mu_1 v_t + (1-\alpha) \lambda \mu_0 v_t^T = 1. \tag{17}$$

In the steady-state, the proportion of monopolized niches coincides with the baseline model, whereas the value of being an incumbent monopolist and a troll can be written, using (16) and (17), as the following functions of the steady-state innovation flow:

$$v(m) = \frac{\pi}{1 - \beta \left\{ 1 - m \left[(1-\alpha) \lambda \mu_0 + \alpha \mu_1 \right] \right\}}, \qquad v^T(m) = \beta^T(m) v(m). \tag{18}$$

where $\beta^T(m) < 1$ is an endogenous discount factor that reflects the expected discounted time at which the troll will successfully license its patent for a fee $v(m)$ to a genuine innovator who loses in court:

$$\beta^T(m) = \frac{\beta m \alpha (1-\mu_1)(1-\varphi)}{1-\beta+\beta m \{\alpha[1-\varphi(1-\mu_1)] + (1-\alpha)\lambda\mu_0\}} < 1. \tag{19}$$

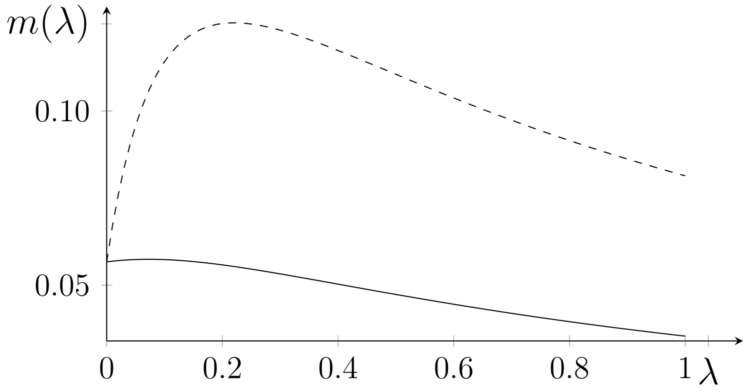
In the polar cases in which financial or other bargaining frictions prevent licensing with probability one ($\varphi = 1$) or incumbents always succeed in court ($\mu_1 = 1$), we have $\beta^T(m) = 0$ and trolling is ineffective. This occurs because licensee fees are the mechanism that allows trolls to appropriate rents from defeated entrants.

The following proposition shows that, with $\mu_1 < 1$ and $\varphi > 0$, patent trolling is necessarily detrimental to innovation. Intuitively this happens because the incentives to innovate derived from the prospect of becoming a troll (if the innovation turns out to be obvious) do not exceed the disincentives coming from the reduction in the expected payoff from genuine innovation caused by trolling.[22]

Proposition 5. *Under imperfect enforcement ($\mu_1 < 1$) and licensing frictions ($\varphi > 0$), the innovation flow when trolls can extract rents from future genuine innovators is lower than in the baseline model.*

What is the impact of patent screening in this setup? Figure 7 shows that the relationship between the innovation flow m and the leniency rate λ can still be non-monotonic—although in the depicted example $m(\lambda)$ has much less curvature than in the baseline model without trolling. In the benchmark model, a more lenient patent screening, by decreasing the proportion of monopolized niches, removes some of the hurdles to profitable entry faced by genuine innovators. When patent trolling occurs, this benefit no longer exists since trolls pose an obstacle to profitable entry in competitive niches. The force that makes $m(\lambda)$ potentially increasing when λ is low is the direct effect of leniency on the prospect of profiting from innovation by becoming a troll when the innovation is obvious—this effect was zero in the baseline model but is clearly positive in (17), for given values of v and v^T. The force that makes $m(\lambda)$ decreasing for larger values of λ is

[22] Notice that when a genuine innovator faces a troll there is a probability $(1-\mu_1)\varphi$ that the troll wins in court and licensing negotiations fail, in which case the innovation goes to waste.



Note: Parameter values are $\alpha = 0.1$, $\beta = 0.98$, $\pi = 0.5$, $\varphi = 0$, $\mu_1 = 0.45$, and $\mu_0 = 0.25$.

Figure 7: Impact of screening leniency on innovation with trolling. This figure shows that, even when trolling occurs, innovation is maximized with some screening leniency (solid line), and that innovation is less than in the benchmark model (dashed line).

of the same nature as in the baseline model: by reducing the expected duration of the incumbency status, leniency decreases both v and v^T.

Welfare is also affected in this environment. Without trolling, genuine innovators could enter unopposed when they landed on a competitive niche. Now, they face the opposition of a troll, leading to the waste of innovations whenever their owners are defeated in court and the licensing negotiations fail. Thus, the proportion of monopolized niches is no longer directly relevant for welfare since genuine innovations go to waste with the same probability in all niches. Hence, social welfare is now given by

$$W(\lambda) = m(\lambda) \{\alpha [1 - \varphi (1 - \mu_1)] \Pi - 1\}, \tag{20}$$

where the effect of the leniency rate is just the one occurring through the innovation rate. Since, as shown by Figure 7, the innovation rate can be maximized at some strictly positive value of λ, it follows that some screening leniency may still be welfare-enhancing under patent trolling. Numerical results indicate that this is more likely to be the case when μ_1 is small.

4.4 Licensing under asymmetric information

Throughout our analysis, we have assumed there was no asymmetric information regarding the quality of the innovations. We now consider the case in which litigation and the potential negotiation about licensing around such a litigation occurs under asymmetric information, that is, with the quality of the innovation being private information of the entering innovator.

Having assumed that legal costs are negligible—compared to the value of being a monopolist—implied that it was a dominant strategy for incumbents to go to court regardless of the kind of innovation they faced. Our results for the case in which licensing negotiations occur after the court ruling are unchanged by the presence of asymmetric information since an obvious infringer (whose innovation yields $\varepsilon \to 0$ profits) is never willing to pay more than zero to become the new incumbent, while the existing incumbent monopolist would require a licensing fee of at least v_t to give up its incumbency. At the same time, v_t is the maximum licensing fee that the infringer would accept if holding a genuine innovation. Hence, the equilibrium with asymmetric information involves a take-it-or-leave offer by the incumbent requesting a licensing fee v_t, and this offer will only be accepted by the infringer if the innovation is genuine. This implies the same payoffs, niche turnover, and welfare implications as in our baseline analysis.

Our results are also robust to the possibility of licensing happening prior to litigation, a procedure extensively discussed in the literature as a way to avoid the costs of going to court. Prior literature also discusses asymmetric information as a potential obstacle to efficient ex-ante licensing (e.g., Llobet, 2003, in a vertical differentiation framework), the strategic use of litigation when information is asymmetric (Figueroa and Lemus, 2024; Pease *et al.*, 2023; Righi and Simcoe, 2023), and how some informational asymmetries (e.g., regarding the gains of an incumbent from going to court) might provide additional incentives to litigate (Choi, 1998).

To discuss ex-ante licensing and the implications of asymmetric information in our

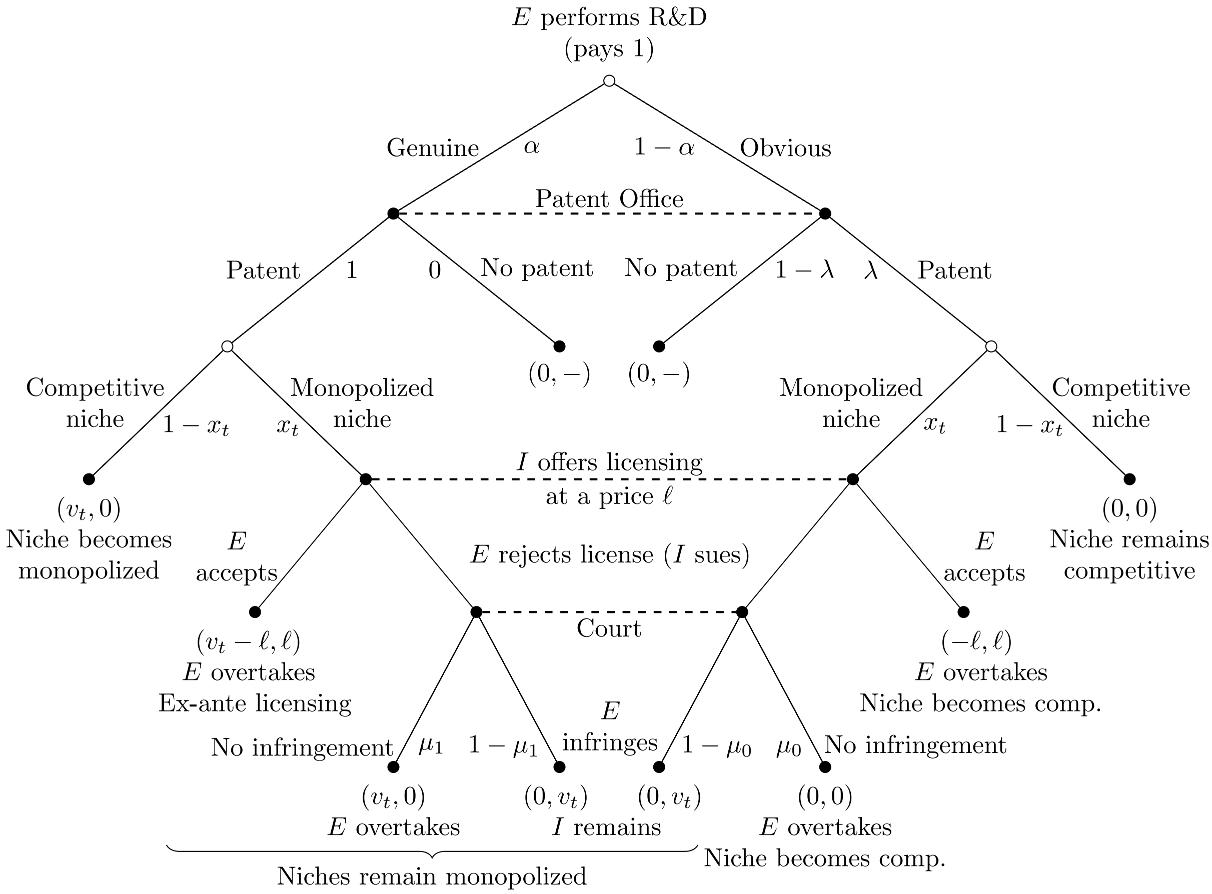


Figure 8: Game tree with ex-ante licensing. The incumbent, without having observed the quality of the entering innovation, can offer a license with fee l to the innovator to avoid litigation.

setup, we extend the baseline model to allow the incumbent, facing an entering innovator, to make an ex-ante take-it-or-leave-it licensing offer of its patent to the entrant, which, if accepted, avoids litigation. At this point, the incumbent ignores whether the entering innovation is genuine or obvious, while the innovator knows. By accepting the offer, the entrant becomes the new incumbent. Otherwise, both firms go to court, and the game unfolds as in the baseline setup. Reflecting the same frictions as in the baseline model, we maintain the assumption that there is an exogenous probability φ that the licensing negotiations fail and the case goes to court even if the innovator is willing to accept the offer. The timing of the model is summarized in Figure 8.

In general, the incumbent's optimal licensing offer depends on its beliefs about the quality of the entering innovation and on the expected response of the innovator. Similarly to the case with ex-post licensing, an obvious innovator's expected value from becoming the new incumbent is zero, so making an offer that an obvious innovator may accept would be unprofitable for the incumbent, since it would be strictly dominated by the expected payoff from going to court and having a chance to remain a monopolist incumbent. This means that it is optimal for the incumbent to offer the highest licensing fee that a genuine innovator is willing to accept, which is such an innovator's expected payoff $(1-\mu_1)v_t$ from going to court and having the probability $1-\mu_1$ that the incumbent wins and preserves the incumbency value v_t.

When the innovation is genuine, this offer leads to a licensing agreement with probability $1-\varphi$, and to litigation otherwise. When the innovation is obvious, the challenger rejects the offer, and the case goes to court. As a result, ex ante licensing produces the same industry dynamics and (expected) payoffs for both the incumbent and the innovator as the baseline model. Thus, our results under ex-ante licensing are equivalent to those with ex-post licensing and are not affected by asymmetric information.

This equivalence relies on having assumed the same probability of failure of the licensing negotiations in both scenarios. However, if the licensing frictions captured by φ are directly related to some heterogeneous incidence of financial constraints among the entering innovators (e.g. different levels of available liquid wealth or exogenous collateral), the scenario with ex-ante licensing might involve an effectively lower value of φ than the baseline scenario with ex-post licensing. Why? Because in the former, the licensing fee that the incumbent requires is $(1-\mu_1)v_t$, which is lower than the licensing fee v_t required when the negotiation happens after having defeated the innovator in court. The lower ex-ante licensing fee might be financially feasible to a strictly larger proportion of genuine innovators, implying a lower probability of failing to reach an agreement than under ex-post licensing. This would imply a lower probability that the genuine innovation goes to waste and, other things equal, it would reduce the welfare-maximizing level of patent screening leniency, bringing it closer to the innovation-maximizing level (Proposition 2).

5 Concluding Remarks

This paper explores how patent systems affect innovation when technological progress occurs sequentially along a quality ladder. Patents not only reward inventors but can also be strategically used by their incumbent holders to block or extract rents from future innovators. In an ideal legal system, high-quality innovations should be patentable and protected from such behavior, while low-quality (or "obvious") innovations might justifiably be excluded. However, under imperfect enforcement, courts may wrongly block genuine innovations, causing a misallocation of innovation incentives. We show that relaxing patentability standards for obvious innovations —granting them some protection— can reduce the risk faced by genuine innovators because their presence brings competition to the market and diminishes litigation threats on further entrants.

Our analysis suggests that some leniency towards obvious innovation may improve social welfare, particularly when patent licensing is subject to frictions. By making the corresponding industry more competitive, leniency fosters innovation and prevents valuable innovations from being stifled. This result is robust across different settings, including scenarios with standard welfare losses and alternative informational and licensing assumptions. We also consider variations where obvious innovators may engage in rent-seeking behavior ("patent trolling") and show that leniency is still optimal, although for slightly different reasons.

Our results entail empirically testable implications for the relationship between patent leniency and innovation outcomes across different industries, particularly in sectors with high litigation activity or where sequential innovation is prominent (e.g., pharmaceuticals or software). Such a relationship has greater welfare implications in the presence of licensing frictions such as financial constraints. Specifically, our analysis predicts that the net gains from such patent screening leniency (and hence the optimal leniency rate) are higher in industries or economies in which innovators face tighter financial constraints. Our analysis also predicts that when financial frictions are an obstacle to efficient licensing, having innovators negotiating licensing agreements with incumbents before going to court may be superior to doing so after the innovator loses in court, as the implied licensing payment is lower (and hence potentially more financial affordable) before the uncertainty about the court's ruling is resolved.

From a policy perspective, our analysis sheds new light on the relationship between ex-ante patent screening (or patent standards) and the enforcement of patent rights in court. It takes us to the novel implication that imperfections whereby courts excessively protect incumbents against subsequent genuine innovators may justify having looser patenting standards since the access to the market of obvious innovations may strategically facilitate the future entry of genuine innovations —by making them less likely to confront an incumbent willing to fiercely defend its monopoly position in court.

References

Aghion, P. and Howitt, P. (1992). A model of growth through creative destruction. *Econometrica*, **60** (2), 323–351.

Allison, J. R. and Lemley, M. A. (1998). Empirical evidence on the validity of litigated patents. *AIPLA Quarterly Journal*, **26** (3), 185.

Atal, V. and Bar, T. (2014). Patent quality and a two-tiered patent system. *Journal of Industrial Economics*, **62** (3), 503–540.

Bessen, J. and Maskin, E. (2009). Sequential innovation, patents, and imitation. *Rand Journal of Economics*, **40** (4), 611–635.

Caillaud, B. and Duchêne, A. (2011). Patent office in innovation policy: Nobody's perfect. *International Journal of Industrial Organization*, **29** (2), 242–252.

Choi, J. P. (1998). Patent litigation as an information-transmission mechanism. *American Economic Review*, pp. 1249–1263.

Daughety, A. F. and Reinganum, J. F. (1995). Keeping society in the dark: On the admissibility of pretrial negotiations as evidence in court. *Rand Journal of Economics*, **26** (2), 203–221.

Denicolò, V. (2000). Two-stage patent races and patent policy. *Rand Journal of Economics*, **31** (3), 488–501.

Denicolò, V. and Zanchettin, P. (2002). How should forward patent protection be provided? *International Journal of Industrial Organization*, **20** (6), 801–827.

Denicolò, V. and Zanchettin, P. (2014). What causes over-investment in R&D in endogenous growth models? *Economic Journal*, **124** (581), 1192–1212.

Dixit, A. K. and Stiglitz, J. E. (1977). Monopolistic competition and optimum product diversity. *American Economic Review*, **67** (3), 297–308.

Farrell, J. and Shapiro, C. (2008). How strong are weak patents. *American Economic Review*, **98** (4), 1347–1369.

Figueroa, N. and Lemus, J. (2024). The benefit of the doubt: Patent examination under strategic obfuscation. *Management Science*, **70** (9), 5853–5870.

Frakes, M. D. and Wasserman, M. F. (2019). Irrational ignorance at the patent office. *Vanderbilt Law Review*, **72** (3), 975–1030.

Gilbert, R. J. and Newbery, D. M. G. (1982). Preemptive Patenting and the Persistence of Monopoly. *American Economic Review*, **72** (3), 514–526.

Green, J. R. and Scotchmer, S. (1995). On the division of profit in sequential innovation. *Rand Journal of Economics*, **26** (1), 20–33.

Grossman, G. M. and Helpman, E. (1991). Quality ladders in the theory of growth. *Review of Economic Studies*, **58** (1), 43–61.

Heller, M. A. and Eisenberg, R. S. (1998). Can patents deter innovation? The anticommons in biomedical research. *Science*, **280** (5364), 698–701.

Hopenhayn, H., Llobet, G. and Mitchell, M. (2006). Rewarding sequential innovations: Prizes, patents and buyouts. *Journal of Political Economy*, **104** (6), 1041–1068.

Jacobsmeyer, B. (2018). Forum shopping in patent cases: Lessons for the unified patent court. *Michigan Technology Law Review*, **25** (1), 131–162.

Landeo, C. M., Nikitin, M. and Baker, S. (2007). Deterrence, lawsuits, and litigation outcomes under court errors. *Journal of Law, Economics, & Organization*, **23** (1), 57–97.

Lee, T. and Wilde, L. L. (1980). Market structure and innovation: A reformulation. *Quarterly Journal of Economics*, **94** (2), 429–36.

Lemley, M. A. and Sampat, B. (2008). Is the patent office a rubber stamp? *Emory Law Journal*, **58**, 181.

— and Shapiro, C. (2005). Probabilistic patents. *Journal of Economic Perspectives*, **19** (2), 75–98.

Lemley, M. D. (2001). Rational ignorance at the patent office. *Northwestern University Law Review*, **95**, 1495–1532.

— (2008). Ignoring patents. *Michigan State Law Review*, **19**, 19–34.

Llobet, G. (2003). Patent litigation when innovation is cumulative. *International Journal of Industrial Organization*, **21** (8), 1135–1157.

Loury, G. C. (1979). Market structure and innovation. *Quarterly Journal of Economics*, **93** (3), 395–410.

Marshall, G. and Parra, Á. (2019). Innovation and competition: The role of the product market. *International Journal of Industrial Organization*, **65**, 221–247.

O'Donoghue, T. (1998). A patentability requirement for sequential innovation. *Rand Journal of Economics*, **29** (4), pp. 654–679.

—, Scotchmer, S. and Thisse, J.-F. (1998). Patent breadth, patent life, and the pace of technological progress. *Journal of Economics & Management Strategy*, **7** (1), 1–32.

Parra, Á. (2019). Sequential innovation, patent policy, and the dynamics of the replacement effect. *Rand Journal of Economics*, **50** (3), 568–590.

Pease, M., Andrews, M. and Vasu, R. (2023). In the interest of full disclosure: Consequences of the grand bargain in patenting. *Available at SSRN 4552164*.

Picard, P. M. and van Pottelsberghe de la Potterie, B. (2013). Patent office governance and patent examination quality. *Journal of Public Economics*, **104**, 14–25.

Righi, C. and Simcoe, T. (2023). Patenting inventions or inventing patents? continuation practice at the uspto. *Rand Journal of Economics*, **54** (3), 416–442.

Schankerman, M. and Schuett, F. (2021). Patent Screening, Innovation, and Welfare. *Review of Economic Studies*, **89** (4), 2101–2148.

Scotchmer, S. (2004). *Innovation and Incentives*. Cambridge, MA: MIT Press.

— and Green, J. (1990). Novelty and disclosure in patent law. *Rand Journal of Economics*, **21** (1), pp. 131–146.

Segal, I. and Whinston, M. D. (2007). Antitrust in innovative industries. *American Economic Review*, **97** (5), 1703–1730.

Spier, K. E. (1994). Pretrial bargaining and the design of fee-shifting rules. *Rand Journal of Economics*, **25** (2), 197–214.

Appendix

Proof of Lemma 1. To show existence, we start by imposing the steady condition $x_t = x$ for all t in eq. (2), obtaining eq. (5). Replacing x into (3) we obtain eq. (6). Using the free-entry condition (4), we obtain $v = p^{-1}$. Finally, using condition (4) again and in the steady state of (1), we find the expression for the innovation flow (7). Because all these equations are linear, we have a unique solution.

To determine when the innovation flow is positive, observe that (7) is positive if and only if $\pi > \underline{\pi}$ and lower than one if and only if $\pi < \underline{\pi} + \beta x^{-1}$. ■

Proof of Corollary 1. Comparative statics follow from differentiating expressions (5–8) with respect to the respective variables. ■

Proof of Proposition 1. The derivative of the innovation flow with respect to λ is

$$m'(\lambda) = \frac{(1-\beta)(1-\alpha)\mu_0}{\beta(\alpha\mu_1 + (1-\alpha)\lambda\mu_0)^2}\left[1 - \alpha\Pi\left(\frac{p(\lambda)}{\alpha}\right)^2\right].$$

Since the terms outside the square brackets are positive, the sign of this derivative depends on the term in square brackets.

Consider first the case with $\mu_1 = 1$. In this scenario, $p(\lambda) = \alpha$, and the square brackets becomes $[1 - \alpha\Pi]$. The condition for positive innovation, $\pi > \bar{\pi}$, simplifies to $\alpha\Pi > 1$, resulting in $m'(\lambda) < 0$ and with an optimal solution $\lambda_m = 0$.

Consider now $\mu_1 < 1$. If innovation is maximized at interior value of λ, then $1 = \alpha\Pi(p(\lambda)/\alpha)^2$ must hold. We show that this point determines a maximum. The second derivative of innovation with respect to λ is

$$m''(\lambda) = -\frac{2(1-\beta)(1-\alpha)^2\mu_0^2}{\beta(\alpha\mu_1 + (1-\alpha)\mu_0\lambda)^3}\left[1 - \alpha\Pi\left(\frac{p(\lambda)}{\alpha}\right)^3\right].$$

The terms outside the square brackets are negative; the sign of $m''(\lambda)$ is the opposite of the sign in square brackets. Interior leniency, $1 = \alpha\Pi(p(\lambda)/\alpha)^2$, automatically implies

$$1 - \alpha\Pi\left(\frac{p(\lambda)}{\alpha}\right)^3 = \alpha\Pi\left(\frac{p(\lambda)}{\alpha}\right)^2 - \alpha\Pi\left(\frac{p(\lambda)}{\alpha}\right)^3 = \alpha\Pi\left(\frac{p(\lambda)}{\alpha}\right)^2\left(1 - \frac{p(\lambda)}{\alpha}\right) > 0$$

as $p(\lambda) < \alpha$ whenever $\mu_1 < 1$. Thus, if there exists a λ_m that leads to an interior leniency, then the second-order condition holds. More generally, because $p'(\lambda) > 0$, $m''(\lambda) < 0$ for every $\lambda \geq \lambda_m$. We will use this fact in other proofs. This fact also means that the derivative is always negative after an interior maximum (if one exists) and the function $m(\lambda)$ is, therefore, single-peaked.

We now derive conditions for which an interior λ_m exists. Solving $1 = \alpha\Pi(p(\lambda^*)/\alpha)^2$ as a function of $p(\lambda^*)$, we obtain two solutions. One solution leads to a negative probability; the other leads to a positive, well-defined, probability $p(\lambda^*) = \sqrt{\alpha\Pi}/\Pi \in (0,1)$. Since $p(\lambda)$ is one-to-one with λ, we can solve for for λ^*, obtaining

$$\lambda^* = \frac{\alpha\big((1-\mu_1)\sqrt{\alpha\Pi} - (\alpha\Pi\mu_1 - 1)\big)}{(1-\alpha)\mu_0(\alpha\Pi - 1)}.$$

It is easy to show that $\lambda^* > 0$ if $\mu_1 < \hat{\mu}_1 = 1/\sqrt{\alpha\Pi} < 1$. In this case we set $\lambda_m = \lambda^* > 0$. Otherwise, when $\mu_1 \geq \hat{\mu}_1$, $\lambda^* < 0$, which can not be. Since $m(\lambda)$ is single peaked, $m'(\lambda) < 0$ for every $\lambda > \lambda^*$ and thus, $\lambda_m = 0$, proving the result. ■

Proof of Proposition 2. First notice that when $\varphi = 0$ or when $\mu_1 = 1$, welfare problem (10) simplifies to $W = m(\lambda)(\alpha\Pi - 1)$. Since, by assumption 1, $\alpha\Pi - 1 > 0$, and the optimal leniency coincides with the one that maximizes innovation and $\lambda_w = \lambda_m$.

Suppose now that $\varphi \in (0,1]$ and $\mu_1 \in (0,1)$. The first derivative is characterized in (11). Evaluate the derivative at λ^* from Proposition 1, then it becomes:

$$W'(\lambda) = -m(\lambda^*)\Pi\alpha(1-\mu_1)\varphi x'(\lambda^*) > 0,$$

as $x'(\lambda) < 0$. This implies that the welfare-maximizing λ_w is larger than λ^*. But, since λ^* might be negative, in that case, $\lambda_m = 0$ and λ_w can be larger or equal to zero. When $\lambda_m = \lambda^* > 0$, then $\lambda_w > \lambda_m$. By single-peakeness of $m(\lambda)$, the last result also implies that, $m'(\lambda_w) < 0$ and $m''(\lambda_w) < 0$.

We now show that $W(\lambda)$ has a unique maximizer. Call λ° any point satisfying $W'(\lambda^\circ) = 0$. If λ° doesn't exist, $W(\lambda)$ is uniquely maximized at a corner. Computing the second derivative of the welfare function at λ°

$$\begin{aligned} W''(\lambda^\circ) &= [\alpha\Pi(1 - x(\lambda^\circ)(1-\mu_1)\varphi) - 1]m''(\lambda^\circ) \\ &\quad -2\alpha\Pi(1-\mu_1)\varphi\, m'(\lambda^\circ)x'(\lambda^\circ) - \alpha\Pi(1-\mu_1)\varphi\, m(\lambda^\circ)x''(\lambda^\circ) < 0, \end{aligned}$$

where the negative sign stems from the fact that: $[\alpha\Pi(1 - x(\lambda^\circ)(1-\mu_1)\varphi) - 1] > 0$ by Assumption 1, we already showed $m'(\lambda^\circ) < 0$, $m''(\lambda^\circ) < 0$ and $x'(\lambda) < 0$, and it easy to verify that $x''(\lambda) > 0$. Hence, if interior, λ° characterizes a maximum. Moreover, since $W''(\lambda^\circ) < 0$ at every candidate λ°, the maximum is unique. Call it $\lambda'_w(\varphi) > 0$, which may not be interior.

To show $\lambda'_w(\varphi) > 0$ when not at a corner, we use the implicit function theorem. Compute the cross-derivative between λ and φ:

$$\frac{\partial W'(\lambda_w)}{\partial \varphi} = -\alpha\Pi(1-\mu_1)\left[m'(\lambda_w)x(\lambda_w) + m(\lambda_w)x'(\lambda_w)\right] > 0$$

Then, by the implicit function theorem,

$$\lambda'_w(\varphi) = -\frac{\partial W'(\lambda_w)/\partial\varphi}{W''(\lambda_w)} > 0,$$

implying that, in an interior solution, the derivative is positive. ■

Proof of Lemma 2. Let $i > 1$ be an arbitrary number of innovations in the quality ladder and $z_i = \delta^i z$ for any $z > 0$. In a monopolized niche, due to the unit elasticity of demand, the incumbent wants to charge the highest price feasible. In this case $p = z_{i-1}$. Then, the incumbent profits are given by $\pi = (p - z_i)q = a(1-\delta)$ which is independent of the number of innovations and the baseline cost z. The dead-weight loss in the market is given by

$$\int_{z_i}^{z_{i-1}} q(p)dp - \pi = a\left(\ln\left(\delta^{-1}\right) - (1-\delta)\right)$$

also independent of i and z. ■

Proof of Lemma 3. The gross welfare originated from an additional innovation can be written as

$$w = \alpha(1-x)\Pi + \alpha x[\mu_1 + (1-\mu_1)(1-\varphi)](\Pi + L) + (1-\alpha)\lambda x \mu_0 L$$

This expression distinguishes three situations where surplus can be created as a result of innovation: i) when genuine innovation occurs in a competitive niche, it creates a surplus Π that is appropriated by the firm; ii) when a genuine innovator replaces a monopolist, it appropriates Π, but it also frees the static loss L generated by the previous incumbent; and iii) when an obvious innovator succeeds in court against a monopolist incumbent, L is freed.

The previous expression can be rewritten as $w(\lambda) = \alpha(\Pi+L)[1-(1-\mu_1)x(\lambda)\varphi]$ which once we account for the size of the innovation flow m and its cost, results in the expression for social welfare function

$$W(\lambda) = m(\lambda)(w(\lambda)-1) = m(\lambda)\{\alpha[1-x(\lambda)(1-\mu_1)\varphi](\Pi+L)-1\},$$

as reported in (12). ■

Proof of Proposition 3. When $L = 0$, the expression for welfare in (12) converges to that is (10), leading to the same solution: $\lambda_w(\varphi)$. Similarly, when $\varphi = 0$ or when $\mu_1 = 1$, the expression becomes $W(\lambda) = m(\lambda)[\alpha(\Pi+L)-1]$, which reaches a maximum at the innovation-maximizing leniency rate λ_m. More generally, the derivative of W with respect to λ can be written as

$$W'(\lambda) = \{[\alpha[1-x(\lambda)(1-\mu_1)\varphi](\Pi+L)-1\}m'(\lambda) - \alpha(1-\mu_1)\varphi(\Pi+L)m(\lambda)x'(\lambda).$$

Similar to the proof of Proposition 2, evaluating the derivative at λ^* from Proposition 1, then it becomes:

$$W'(\lambda) = -\alpha(1-\mu_1)\varphi(\Pi+L)m(\lambda^*)x'(\lambda^*) > 0,$$

implying that the welfare-maximizing $\lambda_d(\varphi, L) \geq \lambda_m$ and strictly larger when $\lambda_m > 0$.

The steps to show that $W(\lambda)$ has a unique maximizer are identical to those in Proposition 2. For completeness, we present the second derivative of the welfare function at the candidate maximum λ_d

$$\begin{aligned} W''(\lambda_d) = {} & \{\alpha[1-x(\lambda_d)(1-\mu_1)\varphi](\Pi+L)-1\}m''(\lambda_d) \\ & - 2\alpha(1-\mu_1)\varphi(\Pi+L)\ m'(\lambda_d)x'(\lambda_d) - \alpha(1-\mu_1)\varphi(\Pi+L)m(\lambda_d)x''(\lambda_d) < 0, \end{aligned}$$

where the negative sign stems from the same argument given in Proposition 2. Hence, since $W''(\lambda_d) < 0$ at every candidate λ_d, the maximum is unique (possibly a corner).

To show $\partial\lambda_d(\varphi, L)/\partial L < 0$ when not at a corner, we use the implicit function theorem, which tells us $\partial\lambda_d(\varphi, L)/\partial L = -(\partial W'(\lambda_d)/\partial L)/W''(\lambda_d)$. We already showed $W''(\lambda_d) < 0$. We show that the other derivative is also negative. Computing:

$$\frac{\partial W'(\lambda_d)}{\partial L} = \alpha[1-x(\lambda_d)(1-\mu_1)\varphi]m'(\lambda_d) - \alpha(1-\mu_1)\varphi m(\lambda_d)x'(\lambda_d),$$

where the first term is negative and the second term is positive. Observe that $W(\lambda_d) = 0$ implies

$$\alpha\left(1-\mu_1\right)\varphi m\left(\lambda_d\right)x'\left(\lambda_d\right) = m'\left(\lambda_d\right)\left(\alpha\left[1-x\left(\lambda_d\right)\left(1-\mu_1\right)\varphi\right] - \frac{1}{(\Pi+L)}\right)$$

replacing back into $\partial W'(\lambda_d)/\partial L$ and re-arranging we obtain

$$\frac{\partial W'(\lambda_d)}{\partial L} = \frac{1}{(\Pi+L)}m'\left(\lambda_d\right) < 0$$

proving the result. ■

Proof of Proposition 4. Let's denote by $m_B(\lambda)$ the innovation flow under innovator buyouts. We need to show $m_B(\lambda) > m(\lambda)$ for every λ. Using equations (7) and (15) the condition becomes $\sqrt{\Pi(\Pi\alpha^2 + 4(p-\alpha))} > \Pi(2p-\alpha)$. Assumption 1 guarantees that the root on the left-hand side is well defined. The right-hand side can be negative as $\alpha > p$ for every $\mu_1 < 1$. If $\mu_1 = 1$, then $p = \alpha$ and $m_B(\lambda) = m(\lambda)$. Suppose $\mu_1 < 1$, when the right-hand side is negative, the result follows. So, we just need to prove the inequality in the case when the right-hand side is positive. Squaring on both sides and rearranging the expression becomes $4\Pi(\alpha - p)(p\Pi - 1) > 0$, which is true for every λ, as $p\Pi > 1$ under Assumption 1. ■

Proof of Proposition 5. Using (18), the free entry condition in (17) can be written as

$$[\alpha\mu_1 + (1-\alpha)\lambda\mu_0\beta^T(m)]v(m) = 1. \tag{21}$$

This condition contrasts with that of the baseline model where, in an interior solution, we have that $pv = 1$ with p given by (6) and v given by (8). Notice that $v(m)$ is decreasing in m and equals v when m takes its baseline value in (7). Hence, to show that the equilibrium innovation flow is lower under trolling, it is enough to show that the probability of profitable entry in the benchmark model, p, is greater than the expression in parenthesis in (21). When courts are perfect ($\mu_1 = 1$), there is no opportunistic behavior, $p = \alpha$, $\beta^T(m) = 0$, and the innovation flow in both scenarios coincides. When $\mu_1 < 1$, we need to show that

$$\alpha\frac{\alpha\mu_1 + (1-\alpha)\lambda\mu_0}{\alpha + (1-\alpha)\lambda\mu_0} > \alpha\mu_1 + (1-\alpha)\lambda\mu_0\left(\frac{\beta m\alpha\left(1-\mu_1\right)\left(1-\varphi\right)}{1-\beta+\beta m\left\{\alpha[1-\varphi(1-\mu_1)] + (1-\alpha)\lambda\mu_0\right\}}\right).$$

In this inequality, the right-hand side is increasing in β and m. So a sufficient condition for it to hold is that it holds for $\beta = m = 1$. After some algebra, this sufficient condition simplifies to

$$\frac{1}{\alpha + (1-\alpha)\lambda\mu_0} > \frac{1-\varphi}{\alpha[1-\varphi(1-\mu_1)] + (1-\alpha)\lambda\mu_0},$$

which, in turn, simplifies to $\varphi[\alpha\mu_1 + (1-\alpha)\lambda\mu_0] > 0$, which is true. ■