Dynamic competition and intellectual property rights in a model of product development

Billette de Villemeur, Etienne and Ruble, Richard and Versaevel, Bruno

LEM, Université de Lille, F-59655, France., Emlyon business School, F-69134 (primary affiliation) and CNRS, GATE Lyon Saint-Etienne, Ecully F-69130, France., Emlyon business School, F-69134 (primary affiliation) and CNRS, GATE Lyon Saint-Etienne, Ecully F-69130, France.

September 2017

Dynamic competition and intellectual property rights in a model of product development

Etienne Billette de Villemeur, Richard Ruble and Bruno Versaevel*

March 29, 2018

Abstract

We study innovation timing and socially optimal intellectual property rights (IPRs) when firms facing market uncertainty invest strategically in product development. If demand growth and volatility are high, attrition occurs and IPRs should ensure the cost of imitation attains a lower bound we identify. If demand growth and volatility are low then provided that entry is business-stealing, IPRs should set the cost of imitation high enough to induce preemption, and possibly winner-take-all preemption. Moreover, the welfare achieved with optimal IPRs is greater with endogenous innovation than if firm roles are predetermined, illustrating the importance of fostering dynamic competition. In extensions we show that firms benefit from open standards, that takeovers have ambiguous welfare effects and that simple licensing schemes are welfare improving.

JEL classification: G31, L13, O33

Keywords: cost of imitation, dynamic competition, patent policy, winner-take-all preemption

*Billette de Villemeur: LEM, Université de Lille, F-59655, France. Ruble and Versaevel: emlyon business School, F-69134 (primary affiliation) and CNRS, GATE Lyon Saint-Etienne, Ecully F-69130, France. Corresponding author: ruble@em-lyon.com. We are grateful to Xavier Boutin, Yann Braouezec, Benoît Chevalier-Roignant, David Encaoua, Thilo Meyer-Brandis, Dean Paxson, Régis Renault, Jacco Thijssen, Mihkel Tombak, and Lenos Trigeorgis as well as seminar participants at Paris-Dauphine university, IESEG business school, LMU and TU München, and participants at the 2016 AFSE meeting and Northwestern Annual Conference on Innovation Economics, 2015 International Conference on Real Options and 2014 EARIE meeting for valuable suggestions.

1 Introduction

When developing an invention into a commercial product requires significant resources, just a few firms may compete for positions in an industry either as a first-mover or as a second entrant. In these circumstances the timing of product introductions is determined by investment strategies that are driven by the relative costs of innovation and imitation. These strategies accordingly respond to policy variables that impact the cost of imitation, and whose choice should therefore account for the dynamics of industry investments.

In a model of dynamic competition to develop a new product, we therefore study the effect of the relative costs of innovation and imitation on the investment strategies of firms when their roles as innovator or imitator are endogenous and characterize the regulator's choice of optimal intellectual property right (IPR) levels. By identifying the role played by the drift and volatility of product market demand on the timing of innovative and imitative investments and hence on economic welfare, we contribute novel insights concerning minimum IPR levels, the necessity of strong IPRs in mature industries, and the importance of fostering dynamic competition between firms.

If innovation has positive spillovers for an imitator attrition may arise, and it is all the more likely when demand growth and volatility are high. A socially optimal level of IPRs involves a minimum cost of imitating that we identify. If demand growth and volatility are low, as typically occurs in mature industries, we show that a high level of IPR protection which induces a preemption race constitutes a second-best from a welfare standpoint in a broad range of situations. If demand growth and volatility are sufficiently low, it is even socially desirable to provide innovators with complete protection and have strategic investment take the form of winner-take-all preemption, so that dynamic competition is intense enough that the firms invest at the net present value threshold as would occur in a competitive industry. The endogeneity of innovation timing plays a key role in establishing these results, and we show that a regulator who did not account for the full range of dynamic competition would run the risk of concluding erroneously that it is optimal not to enforce any IPRs.

To establish these conclusions, we study the exercise of strategic growth options by two initially identical firms pursuing the development of a product for a new market in which they are potential horizontal competitors.[1] Development of innovative and imitative products requires differing stationary levels of irreversible investment which may or may not be related to the exploitation

[1] Our focus on these industries is therefore complementary to research on cumulative innovation such as Green and Scotchmer [13].

of a single patent, and occurs in a context of market uncertainty as the scale of demand follows a geometric Brownian motion.[2] Both firms independently choose thresholds that determine the timing of their investment in product development, which once performed yields an immediate and perpetual profit flow whose level at any moment depends on the number of active firms. We thus study a real option game,[3] but depart from existing models by introducing an *ex-post* asymmetry through the differing fixed costs of innovation and imitation that firms face and by considering the full range of relative fixed costs. Our model therefore allows just as well for broad IPRs implying a relatively high cost of imitation and preemption between firms as well as for significant spillovers resulting in a comparatively low cost of imitation and attrition between firms.

The imitation cost provides us with a way to parametrize first- and second-mover advantage parsimoniously and to nest within a single framework two important timing games, the war of attrition (Hendricks et al. [14]) and preemption (Fudenberg and Tirole [10]).[4] The timing game we study involves firms choosing investment thresholds, or hurdle rates, that determine stochastic investment times, and has a straightforward normal form. We characterize the unique symmetric equilibrium in investment threshold choices. This analysis provides the foundation for the subsequent welfare results which are the main focus of this paper, and we complement it with a more technical discussion of closed-loop strategies in continuous time in the appendix.

After characterizing investment timing and optimal imitation cost levels in a benchmark case where firms invest according to a predetermined sequence instead of engaging in dynamic competition (Proposition 1), we derive the equilibrium timing of innovation for different levels of the imitation cost and identify a critical imitation cost, $\widehat{K}$, which determines whether strategic competition between firms takes the form of attrition or preemption (Proposition 2). For extreme values of the cost of imitation, we find that dynamic competition has the form of a standard timing game. A very low imitation cost leads to a situation of attrition as firms seek to enter second, delaying product introduction and inducing immediate imitation.[5] Conversely a very high imitation cost leads to a situation of preemption as firms seek to enter first and enjoy a phase

[2] We focus on market uncertainty rather than R&D uncertainty, which has been extensively studied by the patent race literature (see *e.g.* Denicolò [5]).

[3] See Chevalier-Roignant and Trigeorgis [3] for a presentation of these games where firms balance the value of retaining flexibility in the face of uncertainty with the strategic incentive to invest early.

[4] The extension of these games to the stochastic case itself presents a number of challenges (Thijssen et al. [26], Steg and Thijssen [24]).

[5] Our model thus encompasses the dynamics described by Scherer (quoted in Fudenberg and Tirole [10]) as "each industry member holding back initiating its R&D effort in the fear that rapid imitation by others will be encouraged, more than wiping out its innovative profits."

of monopoly profit before imitation occurs. Intermediate values of the imitation cost result in hybrid forms of dynamic competition: a waiting game in which firm investment thresholds are continuously distributed over a disconnected support if the imitation cost is moderately low, and a preemption race in which an attrition phase occurs off the equilibrium path if the imitation cost is moderately high.

Provided that innovation has positive spillovers attrition may occur, and it is more likely if there is a low degree of product market competition or if market growth and volatility are high (Proposition 3). This is because high growth and volatility raise the option value of delaying investment, eventually compensating for the lost monopoly profit phase if a firm enters second and imitates instead of innovating. A key additional result concerns the optimal balance between first- and second-mover advantage from the standpoint of the industry. Under both attrition and preemption, positional rents are dissipated in the symmetric equilibrium and expected industry value is therefore maximized if the imitation cost attains the critical level $\widehat{K}$ at which there is neither a war of attrition nor a preemption race, so that firms do not compete for positional rents by either unduly waiting or rushing to innovate (Proposition 4).

Because of the tractability of the equilibrium we characterize, we are able to study socially optimal IPR levels if a regulator adjusts the cost incurred by an imitator through either legislative measures or enforcement. With dynamic competition the welfare trade-off associated with raising the imitation cost is more involved than a straight balancing of the incentive to innovate against the deadweight loss of monopoly, as the effect of higher imitation cost on the timing of imitation is ambiguous under attrition. We identify a lower bound on the socially optimal imitation cost (Proposition 5), which must provide sufficient quasi-rents for firms to avoid the Schererian dynamics described above (*cf.* footnote 5).

Even if it is generally challenging to draw broad conclusions regarding optimal IPR levels, we are able to show that if the static entry incentive is socially excessive, as occurs in the presence of a business-stealing effect, an imitation cost that induces preemption is optimal when market growth and volatility are sufficiently low (Proposition 6). The model therefore provides an argument for strong IPRs in such industries based on objective characteristics of market uncertainty. In passing we obtain closed-form expressions for the optimal threshold for innovation and the resulting level of welfare under preemption (Lemma 1), establishing that a limit imitation cost level which results in winner-take-all preemption is socially optimal when there is sufficient discounting. Moreover we provide specific economic circumstances where the optimal imitation cost is consistent either with attrition or preemption, such as a low consumer surplus from innovation or collusion in the product market (Proposition 7).

These welfare results take on particular relevance when they are compared with the optimal welfare levels obtained without accounting for the endogeneity of innovation or for the full range of dynamic competition, as is often the case in the economic literature on patents. A regulator following this kind of approach could be led to set the level of IPR protection much too low, when in fact competition between firms to innovate plays a vital role and is best incentivized with levels of imitation cost in the preemptive range (Proposition 8). Moreover the comparison of welfare levels that a regulator achieves with and without allowing for dynamic competition casts doubt on the merit of any policy that might involve picking an industrial champion to invest first, even if is complemented by efficient IPR levels.

Finally we discuss several extensions of the model. First, we endogenize the cost of imitation by allowing the innovator to pursue patent protection more aggressively or to make reverse engineering of its product more difficult. A higher baseline cost of imitation reduces the effort exerted by innovators to raise entry barriers, and firms are shown to gain from coordinating *ex-ante* not to introduce subsequent complexity, a policy that may be thought of as an open standard (Proposition 9). We also discuss contracting between innovator and imitator that can take the form either of a takeover or of a license agreement, and show that efficiency always increases in the latter case (Proposition 10).

Our paper is related to early research on innovation incentives and optimal patents, and in particular to Gallini [11] who introduces a cost of imitation that the regulator may use as a policy instrument. We similarly emphasize the role of measures like patent breadth in determining the cost of inventing around an existing innovation, but in contrast with this earlier work we account for the endogenous timing of innovation and thus allow firms to wait before investing rather than assuming that product development occurs when its net present value is positive. Denicolò [5]'s model of optimal IPR protection in a patent race is therefore closer to our work, as it formalizes innovation and imitation as the outcome of a non-cooperative interaction that precedes market competition, though in contrast our model allows for second-mover advantage and attrition, which likely arises in industries with high growth and volatility.

Our work is therefore also related to papers which study the effect of second-mover advantage on investment decisions, most often as a result of explicit informational spillovers. Hoppe [17] allows for uncertainty regarding the success of new technology adoption to benefit a rival's innovation decision whereas in Thijssen et al. [25] information regarding the value of a project arrives continuously over time. Femminis and Martini [9] allow for a disclosure lag of random duration before the follower receives the information. In these models, both preemption and attrition can occur depending on the level of spillovers, but the welfare analysis is either based on pure strategy equilibrium or restricted to preemption regimes. Our analysis characterizes the welfare properties of symmetric mixed strategy equilibrium over a complete range, providing intuitive analytic results regarding optimal IPRs that are related in our model to the dynamic properties of demand.

Our welfare results can be related to several papers that compare welfare across two key alternative policy regimes, a strict winner-take-all regime where only the first firm to innovate receives a patent, and a more permissive regime where late investors are allowed to compete with the first before its patent expires. In La Manna, Macleod, and de Meza [19] firms spend a fixed initial amount in R&D that determines a probability of inventing at a future date. Simple cost and demand conditions, such as constant returns to scale and a linear demand, are identified for the permissive regime to be welfare superior. Henry [15] introduces a mechanism whereby a late inventor can share the patent with the innovator within a given time period. When adjusted, together with other policy instruments, this mechanism is socially beneficial under mild conditions, notably with a linear demand and quantity competition. However, in a model where firms incur a flow cost, in Denicolò and Franzoni [7] it is the strict patent regime that is found to be optimal in a broad set of circumstances, in particular when demand is linear, product market competition is weak, and duplication flow costs are large. Our approach is broadly consistent with these contributions, but we characterize an optimal degree of IPRs with the winner-take-all regime occurring as a limit case rather than evaluating a discrete set of regimes. Moreover, the model of investment under market uncertainty allows us to identify determinants of optimal protection that are not considered in this stream of literature related to measurable properties of the dynamics of markets.

Section 2 describes the model. Section 3 studies a benchmark case in which the roles of firms are predetermined. Section 4 characterizes the symmetric equilibrium when firms engage in dynamic competition. Section 5 studies welfare when a regulator determines the cost of imitation. Section 6 discusses two extensions of the model. Section 7 concludes.

2 A model of new product development

This section sets up a model of strategic investment in product development that reflects the characteristic features of innovation and imitation discussed in the introduction. The main assumptions are presented in Section 2.1, the continuation payoffs that firms obtain once innovation occurs in Section 2.2, and Section 2.3 describes the relationship between the configuration payoffs and imitation cost that provides the intuition for much of the analysis in the paper.

2.1 Assumptions

Two identical firms engage in dynamic competition to introduce their version of a novel product in an evolving market. Organizational constraints prevent a firm from developing two variants of the novel product and entry barriers shield both firms from other competitors.

Introduction of the product immediately generates a perpetual profit flow whose baseline value is denoted π_M if a single firm is active in the market and π_D if both are, with $0 < 2\pi_D \leq \pi_M$. A firm that introduces its product first is said to be an *innovator* and if a firm introduces its product second it is referred to as the *imitator*. The baseline profit flow is scaled by a measure of market size Y_t, $t \geq 0$, so flow profits to active firms at a given time are either $\pi_M Y_t$ or $\pi_D Y_t$. To capture the idea that demand for the new product evolves in a context of market uncertainty, the measure of market size is assumed to follow a geometric Brownian motion $dY_t = \alpha Y_t dt + \sigma Y_t dW_t$ where W_t is a standard Wiener process and α and $\sigma \geq 0$ are the drift and volatility. Both firms have a common discount rate r.

Product development involves an irreversible investment which encompasses standard setup costs associated with bringing a product to market such as dedicated plant, equipment and marketing expenses as well as the cost of developing the firm's product variant. The fixed costs of innovation and imitation are respectively denoted I and K. While I is assumed to be positive and finite, the extreme cases $K = 0$ and $K = \infty$ are allowed and we are agnostic about the relative magnitudes of I and K.[6] If the second firm can develop an equivalent product completely independently then $K = I$, whereas if there are positive spillovers $K < I$, and with scarce inputs or IPR protection that compels imitators to invent around any intellectual property held by the innovator $K > I$ can hold.[7]

[6] Within the biopharmaceutical industry for instance, the cost of imitation varies across business segments. The conditions of imitation for drugs strongly differ from those for vaccines. Pharmaceutical firms rely on intellectual property rights in order to increase the costs of imitators for new drugs "which otherwise could be copied more easily than products whose production processes can be kept secret, or for which the time and relative expense needed to copy the invention are much higher" (Scherer and Watal [27], p. 4). If such patent protection is not available, a generic product can be introduced at a much lower fixed cost than incurred by the branded product supplier. However, this ease of imitation is not found in the case of vaccines, which are made from living micro-organisms, and unlike drugs "are not easily reverse-engineered, as the greatest challenges often lie in details of production processes that cannot be inferred from the final product," implying that "there is technically no such thing as a generic vaccine" (Wilson [28], p. 13).

[7] Imperfect competition in the input market can also lead to asymmetric fixed costs for initially identical firms. Billette de Villemeur et al. [1] show for example that if the cost of investment is determined endogenously by a monopoly input supplier, the fixed cost is lower for the first firm that invests.

To derive welfare results we suppose that like profits, the baseline consumer surplus flows under monopoly and duopoly, S_M and S_D with $0 < S_M \leq S_D$, are scaled by the market size parameter Y_t, and that the social discount rate is equal to r. The static welfare gain from imitation, $(S_D + 2\pi_D) - (S_M + \pi_M)$, plays a key role in our normative analysis. A standard result in industrial economics is that in a broad range of oligopoly models, the static welfare gain is lower than the private entry incentive π_D,[8] and we appeal to this result further below to identify certain welfare effects.

Finally in order for the investment problems we study to be economically meaningful we assume that $r > \alpha$, and to focus on cases where firms to prefer to delay initially we assume $Y_0 \leq (r - \alpha) I / \pi_M$.

2.2 Continuation payoffs

Firms obtain continuation payoffs once innovation occurs which are determined by their position in the investment sequence. These payoffs are defined for a given value $y = Y_0$ of the market size process as functions of the market size at which innovation occurs, which is denoted Y and satisfies $Y \geq y$. They thus represent the current values of anticipated rather than instantaneous payoffs, and as in the literature they are denoted F, L and M according to whether a firm invests as a follower, as a leader or if investments are simultaneous.

The continuation payoff of a follower is obtained by studying the decision problem of a firm once its rival has innovated. It then holds a growth option on a duopoly market and its optimal policy is to develop the imitative product whenever the market reaches an optimal threshold, denoted Y_F. Standard arguments (see Section A.1) establish that the instantaneous value at the threshold Y of this option is

$$V_D(Y) = \begin{cases} A_D Y^\beta, & Y < Y_F \\ \frac{\pi_D}{r-\alpha} Y - K, & Y \geq Y_F \end{cases} \tag{1}$$

where β is shorthand for the function of parameters

$$\beta(\alpha, \sigma, r) := \frac{1}{2} - \frac{\alpha}{\sigma^2} + \sqrt{\left(\frac{1}{2} - \frac{\alpha}{\sigma^2}\right)^2 + \frac{2r}{\sigma^2}}, \tag{2}$$

$Y_F := (\beta(r-\alpha)K) / ((\beta - 1)\pi_D)$ is the optimal duopoly investment threshold and $A_D := \beta^\beta \pi_D^\beta / (\beta - 1)^{\beta-1} (r-\alpha)^\beta K^{\beta-1}$ is a constant. The discounting parameter (2) satisfies $\beta > 1$

[8] Specifically this holds in industries where entry raises industry output, there is a business-stealing effect and firms do not price below marginal cost (Mankiw and Whinston [20]).

and $\partial\beta/\partial\alpha, \partial\beta/\partial\sigma, -\partial\beta/\partial r < 0$.[9] The expected discounted value at time $t = 0$ of obtaining the duopoly option $V_D(Y)$ at Y when rival innovation occurs is then

$$\begin{aligned} F(Y) &= \mathbb{E}_y\left[\int_{\tau(\max\{Y,Y_F\})}^{\infty} \pi_D Y_s e^{-rs} ds - e^{-r\tau(\max\{Y,Y_F\})} K\right] \\ &= \begin{cases} A_D y^{\beta}, & Y < Y_F \\ \left(\frac{\pi_D}{r-\alpha} Y - K\right)\left(\frac{y}{Y}\right)^{\beta}, & Y \geq Y_F. \end{cases} \end{aligned} \tag{3}$$

Given that the optimal policy of the follower is to invest once the threshold Y_F is reached, the leader payoff at time $t = 0$ from innovating at a threshold Y is

$$\begin{aligned} L(Y) &= \mathbb{E}_y\left[\int_{\tau(Y)}^{\tau(\max\{Y,Y_F\})} \pi_M Y_s e^{-rs} ds - e^{-r\tau(Y)} I + \int_{\tau(\max\{Y,Y_F\})}^{\infty} \pi_D Y_s e^{-rs} ds\right] \\ &= \left(\frac{\pi_M}{r-\alpha} Y - I\right)\left(\frac{y}{Y}\right)^{\beta} - \frac{\pi_M - \pi_D}{r-\alpha} \frac{y^{\beta}}{[\max\{Y, Y_F\}]^{\beta-1}} \end{aligned} \tag{4}$$

where in the first line $\tau(Z) = \inf\{t \geq 0 | Y_t \geq Z\}$ denotes the stochastic time at which the market size process first hits a threshold Z (see Section A.2). The first summand in (4) corresponds to the discounted value of perpetual monopoly profits for an innovator and the second corrects for the anticipated reduction in profit flow stemming from the rival firm's entry. Finally if both firms invest together at a market size threshold Y the continuation payoff from simultaneous investments is

$$\begin{aligned} M(Y) &= \mathbb{E}_y\left[\int_{\tau(Y)}^{\infty} \pi_D Y_s e^{-rs} ds\right] \\ &= \left(\frac{\pi_D}{r-\alpha} Y - I\right)\left(\frac{y}{Y}\right)^{\beta}. \end{aligned} \tag{5}$$

Note that this formulation assumes that the fixed cost is I for both firms, which can be due to a lag for information spillovers to take place that is not modeled explicitly for example.[10]

[9] A standard property of β which is apparent in the continuation payoffs in the text is that the expected discounted value of a monetary unit received at the stochastic time $\tau(Y)$ when the process first hits a threshold $Y \geq y$ is $\mathbb{E}_y e^{-r\tau(Y)} = (y/Y)^{\beta}$. Note that $\lim_{\sigma\to 0} \beta = r/\alpha$.

[10] An alternative assumption would be that fixed costs are attributed randomly when investments are simultaneous so the expected fixed cost at the moment of investment is $(I + K)/2$. In this case the first intersection of L and M is no longer at Y_F and M and F do not overlap beyond $\max\{Y_F, Y_M\}$ where the threshold Y_M is defined further below in the text.

2.3 Continuation payoff configurations and critical imitation costs $\widetilde{K}$ and $\widehat{K}$

The relative positions of F, L and M are determined by the level of K, and in turn determine key features of the model such as whether firms enjoy a first or a second-mover advantage upon investment. To characterize these payoff functions, three critical levels of the imitation cost, $\widetilde{K}, \widehat{K}$ and I with $\widetilde{K} < \widehat{K} < I$ delimit four typical configurations of the continuation payoffs that are represented in Figures $1-4$.[11] The first two critical values, $\widetilde{K} = \left(\beta\left(\left(\pi_M/\pi_D\right)-1\right)/\left(\left(\pi_M/\pi_D\right)^{\beta}-1\right)\right)^{1/(\beta-1)} I$ and $\widehat{K} = \left(\left(1+\beta\left(\left(\pi_M/\pi_D\right)-1\right)\right)/\left(\pi_M/\pi_D\right)^{\beta}\right)^{1/(\beta-1)} I$, are solutions to the conditions $L\left(Y_L\right) = M\left(Y_M\right)$ and $L\left(Y_L\right) = F\left(Y_L\right)$ respectively which are derived in Section A.2.3 where $Y_L := \left(\beta\left(r-\alpha\right)I\right)/\left(\left(\beta-1\right)\pi_M\right)$ and $Y_M := \left(\beta\left(r-\alpha\right)I\right)/\left(\left(\beta-1\right)\pi_D\right)$.

For $K < \widetilde{K}$ (Figure 1),[12] $F > L$ for all $Y \geq y$ and L has a global maximum at Y_M. For $\widetilde{K} < K < \widehat{K}$ (Figure 2), $F > L$ for all $Y \geq y$ but the global maximum of L is at Y_L. Observe that L is not monotonic to the right of its global maximum in this case and that there is a local maximum at Y_S. For $\widehat{K} < K < I$ (Figure 3), $F < L$ over an interval $\left(Y_P, \overline{Y}_P\right)$ that includes Y_L. Finally, for $K > I$ (Figure 4), $F < L$ for all $Y \geq y$.

Although the exact relationship between these functions is involved, there is a general intuition regarding the effect of imitation cost on payoffs that is apparent in the figures. Starting at the lower bound of the cost of imitation $K = 0$, the follower payoff is globally higher than the leader payoff ($F > L$). Raising the level of K increases the follower investment threshold Y_F and shifts F downward, but conversely this higher investment threshold lengthens the monopoly phase for a firm that innovates before the follower threshold is reached and raises the leader payoff, shifting L upward over the range (y, Y_F). As the cost of imitation becomes large enough ($K > I$), the leader payoff eventually becomes globally higher than the follower payoff ($F < L$).

3 Benchmark: predetermined investment sequence

To get a first insight about the dynamics of innovation and imitation in our model we begin by examining the case of exogenous firm roles. One way this situation arises is through industrial policy, for example if a former state-run monopolist enjoys priority market access in a deregulated industry. This case is also a useful first step from a theoretical standpoint because its

[11] The pivotal cases $K \in \left\{\widetilde{K}, \widehat{K}, I\right\}$ are not shown but are straightforward to obtain from the cases that are shown in the figures by continuity of F, L and M in K.

[12] Figure 1 is drawn assuming $K > \underline{K} := (\pi_D/\pi_M)I$. Below this value, $Y_F < Y_L$ and F is decreasing over (y, ∞) but the key properties of F, L and M described in the text still hold.

equilibrium, whose derivation is straightforward, provides a useful comparison with the endogenous firm roles case that we study afterward, particularly with respect to the contrasting industry profit-maximizing and socially optimal imitation cost levels we obtain.

3.1 Industry equilibrium

Suppose that the order of investments is exogenously fixed with firm 1 developing its product first whereas firm 2 must wait for firm 1's variant to be introduced before developing its own product. Let Y_i, $i \in \{1, 2\}$ denote firm i's investment threshold with firm 2's threshold constrained to satisfy $Y_2 \geq Y_1$. With these assumptions the equilibrium pattern of investments is found by backward induction as in a standard Stackelberg duopoly model. We therefore first identify firm 2's optimal investment threshold, denoted $Y_2^F(Y_1)$, and then firm 1's optimal forward-looking investment threshold, denoted Y_1^L.

Firm 2 solves the follower problem described at the beginning of Section 2.2. As it is constrained by $Y = Y_1$, its optimal investment threshold is $Y_2^F = \max\{Y_F, Y_1\}$ and it obtains the payoff $F(Y_1)$. The leader payoff $L(Y_1)$ that firm 1 gets is given by (4), which directly incorporates the reaction of firm 2. Its decision problem is therefore $\max_{Y_1 \geq y} L(Y_1)$. The unique solution is $Y_1^L = Y_L$, and firm 1 obtains the payoff $L(Y_L)$.

With exogenous firm roles and hence sequential threshold choices, innovation and imitation therefore occur at Y_L and $\max\{Y_F, Y_L\}$ respectively. If $K \leq \underline{K} := (\pi_D/\pi_M)\, I$ the equilibrium thresholds are $\{Y_L, Y_L\}$ and involve *clustering,* whereas if $K > (\pi_D/\pi_M)\, I$ the equilibrium thresholds are $\{Y_L, Y_F\}$ and involve *diffusion.*

3.2 Efficient imitation cost and IPR levels

We next examine efficiency from the standpoint of the industry and of the regulator successively.

3.2.1 Industry optimum

As seen above, the level of the imitation cost determines the type of equilibrium outcome in the industry. This cost in turn ultimately depends on technological conditions, but it can also be affected by such choices as *ex-ante* agreements regarding the pooling of resources or common standards. It therefore makes sense to think of K as a decision variable in certain industries and to inquire as to what level is optimal from the industry's perspective.

Answering this question amounts to identifying the cost of imitation that maximizes industry profit, *i.e.* to solving $\max_{K\in\mathbb{R}_+}[L(Y_L)+F(Y_L)]$. Substituting equilibrium threshold values into the payoffs, equilibrium industry profit is

$$L(Y_L)+F(Y_L)=\frac{I}{\beta-1}\left(\frac{y}{Y_L}\right)^{\beta}-\frac{\pi_M-2\pi_D}{r-\alpha}\frac{y^{\beta}}{[Y_F]^{\beta-1}}.$$

With exogenous roles the innovator's investment threshold Y_L is independent of K, whereas increasing K raises the imitation threshold Y_F which has either a neutral effect on industry profit if $2\pi_D=\pi_M$ or a positive effect if $2\pi_D<\pi_M$. In this latter case the optimal imitation cost for the industry is $K^{\pi}=\infty$ and only firm 1 invests. In such an industry, one would not therefore expect development expenses to be pooled or measures such as common standards to be adopted.

3.2.2 Social optimum

The imitation cost can also be affected by policy variables and in particular by the level of IPR protection chosen by regulators. We consider a second-best welfare benchmark, in which firms are free to select their entry thresholds according to the predetermined investment sequence defined above.

For given thresholds Y_1 and Y_2, social welfare is

$$\left(\frac{S_M+\pi_M}{r-\alpha}Y_1-I\right)\left(\frac{y}{Y_1}\right)^{\beta}+\left(\frac{(S_D+2\pi_D)-(S_M+\pi_M)}{r-\alpha}Y_2-K\right)\left(\frac{y}{Y_2}\right)^{\beta}. \tag{6}$$

Substituting for the equilibrium values of Y_1 and Y_2, the welfare second-best can be expressed as the following function of the regulator's instrument:

$$W(K)=\begin{cases}\left(\left(\frac{\beta}{\beta-1}\frac{S_D+2\pi_D}{\pi_M}-1\right)I-K\right)\left(\frac{y}{Y_L}\right)^{\beta}, & K\leq\underline{K}\\ \left(\beta\frac{S_M}{\pi_M}+1\right)\frac{I}{\beta-1}\left(\frac{y}{Y_L}\right)^{\beta}+\left(\beta\frac{(S_D+\pi_D)-(S_M+\pi_M)}{\pi_D}+1\right)\frac{K}{\beta-1}\left(\frac{y}{Y_F}\right)^{\beta}, & K>\underline{K}.\end{cases}$$

$W(K)$ is clearly decreasing for $K\leq\underline{K}$. For $K>\underline{K}$, its behavior depends the magnitude of the second summand which is of the form $A_W K^{-(\beta-1)}$. The sign of the constant A_W depends on the sign of $(\beta((S_D+\pi_D)-(S_M+\pi_M))/\pi_D)+1$. If this latter term is positive, then social welfare is globally decreasing and the socially optimal imitation cost is $K^W=0$. However in many oligopoly models the private entry incentive is socially excessive because of the business-stealing effect (see Section 2.1), which implies that $S_D+\pi_D<S_M+\pi_M$. The constant A_W is then negative if $\beta>\pi_D/((S_M+\pi_M)-(S_D+\pi_D))(\geq 1)$, in which case $W(K)$ increases over $(\underline{K},\infty)$. Even then however,

$$\lim_{K\to\infty} W\left(K\right)=\left(\beta\frac{S_M}{\pi_M}+1\right)\frac{I}{\beta-1}\left(\frac{y}{Y_L}\right)^{\beta}\leq\left(\frac{\beta}{\beta-1}\frac{S_D+2\pi_D}{\pi_M}-1\right)I\left(\frac{y}{Y_L}\right)^{\beta}=W(0),$$

with strict inequality if $S_D + 2\pi_D > S_M + \pi_M$, *i.e.* if the product market does not function as a cartel with both firms active. We conclude that the socially optimal imitation cost level is $K^W = 0$ for all parameter values. As might be expected therefore, if the order of investments is predetermined then while a monopoly is optimal for the industry the social optimum favors competition and avoiding duplication of development expenditures.

To summarize these results,

Proposition 1 *If the investment sequence is predetermined, investments are clustered at Y_L if $K \leq \underline{K}$ and diffused over Y_L and Y_F if $K > \underline{K}$. Provided that entry raises product market output, the efficient imitation costs for the industry and for society are respectively $K^\pi = \infty$ and $K^W = 0$.*

4 Endogenous innovation and imitation

Suppose the roles of firms as innovator or imitator are not predetermined but instead result from dynamic competition. A non-cooperative timing game therefore determines the sequence of investments. Firms choose innovative investment thresholds that they can update as imitators in a non-strategic continuation phase if rival innovation occurs. Industry dynamics typically consist of a period of inaction before either firm has developed the product over which the strategic interaction plays out, followed possibly by a monopoly phase and a duopoly phase once both firms have developed their own variants of the product. The game is described in Section 4.1, equilibrium in Section 4.2 and the effect of imitation cost on equilibrium thresholds and payoffs is discussed in Section 4.3.

4.1 Firm strategies and payoffs

The strategy of firm i, $i = 1, 2$, consists of a threshold $Y_i \in [y, \infty]$ that triggers its investment when reached for the first time. Strategies are chosen at time zero to determine the stochastic time at which each firm plans to invest, assumed to be a first hitting time, and thus the timing of innovation in the industry. Once innovation occurs any remaining firm revises its investment threshold in the continuation phase which determines when imitation occurs.[13]

To describe the investment game the strategies Y_1 and Y_2 must be mapped into outcomes. For $Y_1 \neq Y_2$ this is straightforward, since one of the firms is the leader and obtains the payoff $L(\min\{Y_1, Y_2\})$ while the other is therefore the follower and obtains the payoff $F(\min\{Y_1, Y_2\})$. For $Y_1 = Y_2$ however, taking the outcome to consist of simultaneous investments with payoffs $M(Y_i)$ for each firm does not correctly represent economic behavior if payoffs satisfy $L(Y_i) \geq F(Y_i) > M(Y_i)$, $i = 1, 2$, *i.e.* if both firms seek to invest whereas it would be optimal for only one to do so. Such cases arise typically in preemption games, and in the discrete time mixed strategy equilibrium that continuous time approximates, investments in fact turn out to be partially coordinated. A standard solution in the literature is to have players use extended mixed strategies (Fudenberg and Tirole [10], Thijssen et al. [26], see also Section B), but we follow an alternative approach in this section for simplicity. This approach consists in positing a probabilistic tie-breaking rule for such simultaneous investment attempts. This rule is calibrated to yield payoffs that are consistent with the symmetric equilibrium using extended mixed strategies, and notably satisfy the same rent-dissipation property.

If $Y_1 = Y_2 = Y$ we therefore assume that either firm innovates (leads in investment) with probability[14]

$$p(Y) = \begin{cases} \frac{F(Y)-M(Y)}{L(Y)+F(Y)-2M(Y)}, & Y < Y^F \text{ and } L(Y) \geq F(Y) \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

and simultaneous investments accordingly occur with probability $1 - 2p(Y)$.

With these assumptions the payoff of firm i is

$$V(Y_i, Y_{-i}) = \begin{cases} L(Y_i), & Y_i < Y_{-i} \\ p(Y_i) L(Y_i) + p(Y_i) F(Y_i) + (1 - 2p(Y_i)) M(Y_i), & Y_i = Y_{-i} \\ F(Y_{-i}), & Y_i > Y_{-i} \end{cases} \tag{8}$$

and the normal form of the investment game is

$$(\{1, 2\}, [y, \infty) \times [y, \infty), (V, V)).$$

[13] This ability of firms to update their thresholds when rival investment occurs is the main difference with Reinganum [21]'s technology adoption game with open-loop strategies in which firms remain committed to their initial threshold choice as outcomes unfold.

[14] Observe that in the first line the investment probability solves the condition $pL(Y) + pF(Y) + (1 - 2p) M(Y) = F(Y)$ so firms are indifferent between the expected payoff from investing at the threshold Y and the follower payoff from postponing investment.

A Nash equilibrium of the investment game is a pair of strategies $\left(\widehat{Y}_1, \widehat{Y}_2\right)$ such that $V\left(\widehat{Y}_i, \widehat{Y}_{-i}\right) \geq V\left(Y_i, \widehat{Y}_{-i}\right)$ for all $Y_i \geq y$, $i \in \{1,2\}$. In the next subsection, we describe the unique symmetric Nash equilibrium, which can involve pure or mixed strategies, and establish the necessary formal structure for the subsequent welfare results. Because by construction the payoff function $V\left(Y_i, Y_{-i}\right)$ encapsulates outcomes of an equilibrium with extended mixed strategies, the equilibrium obtained in the static investment game of this section is consistent with the continuous time games in closed-loop strategies in the literature. We verify this by studying a dynamic version of the game in the appendix, which does not pose novel difficulties but is more notationally costly.

4.2 Equilibrium

The investment game has several Nash equilibria involving either pure or mixed strategies. We assume that there are no coordinating mechanisms available, so that the firms, being symmetric *ex-ante,* play the same strategies.[15] Moreover, the resulting equilibrium leads to a compelling relationship between industry outcomes and imitation cost. We therefore focus on the unique symmetric equilibrium, consistently with Fudenberg and Tirole [10]'s study of preemption and with the discussion of attrition in Hendricks *et al.* [14].

To grasp the nature of the timing game for different levels of the imitation cost more easily, the reader may refer again to Figures $1-4$. For $K \leq \widetilde{K}$ (Figure 1), the leader payoff $L\left(Y\right)$ lies below the follower payoff $F\left(Y\right)$ for all $Y \geq y$. Innovating at any threshold $Y_i < Y_M$ is dominated by innovating at Y_M, and from Y_S onward $L\left(Y\right)$ is decreasing so the investment game constitutes a standard war of attrition. For $\widetilde{K} < K < \widehat{K}$ (Figure 2), the leader payoff $L\left(Y\right)$ also lies below the follower payoff $F\left(Y\right)$ for all $Y \geq y$. In this case however, Y_L is the global maximum of the leader payoff and innovating at Y_M is dominated by thresholds in $\left[Y_L, \overline{Y}_L\right)$. Firms therefore engage in a nonstandard war of attrition, with innovation thresholds continuously distributed over the support $\left[Y_L, \overline{Y}_L\right] \cup [Y_M, \infty)$ where $L\left(Y\right)$ decreases. For $\widehat{K} < K < I$ (Figure 3), the leader payoff $L\left(Y\right)$ lies above the follower payoff $F\left(Y\right)$ for $Y \in \left(Y_P, \overline{Y}_P\right)$, where Y_P and $\overline{Y}_P$ are the lower and upper roots of the condition $L\left(Y\right) = F\left(Y_F\right)$. There is therefore preemption over this range, though firms engage in a war of attrition off the equilibrium path if the threshold $\overline{Y}_P$ is reached and no firm has yet invested. Finally if $I \leq K$ (Figure 4), the leader payoff $L\left(Y\right)$ lies above the follower payoff $F\left(Y\right)$ for $Y \in (Y_P, Y_F)$, which corresponds to a standard preemption game.

[15] In a similar model of investment with spillovers Hoppe [17] focuses on asymmetric pure strategy equilibria under attrition. This analysis applies for instance if the firms have multimarket contact that allows investments to be coordinated across markets.

The innovation threshold that arises in symmetric equilibrium, $Y_I := \min\left\{\widehat{Y}_1, \widehat{Y}_2\right\}$, is described in the next proposition. This threshold takes different values $\widetilde{Y}_A$, Y_P or Y_L depending upon whether attrition or preemption occur (cases (*i*) and (*ii*)) or the imitation cost attains a pivotal level $\widehat{K}$ at which neither preemption nor attrition occur (case (*iii*)) (see Section A.3 for proof and derivation of $\widetilde{Y}_A$).

Proposition 2 *In the symmetric equilibrium of the investment game,*

(i) (attrition) if $K < \widehat{K}$ equilibrium is in mixed strategies with an innovation threshold $\widetilde{Y}_A$ distributed continuously over $[Y_M, \infty)$ if $K \leq \widetilde{K}$ and $\left[Y_L, \overline{Y}_L\right] \cup [Y_S, \infty)$ if $\widetilde{K} < K < \widehat{K}$;

(ii) (preemption) if $K > \widehat{K}$ equilibrium is in pure strategies and the innovation threshold is Y_P;

(iii) if $K = \widehat{K}$ equilibrium is in pure strategies and the innovation threshold is Y_L.

Given the equilibrium threshold for innovation, the optimal follower behavior described in Section 2.2 determines the imitation threshold. Imitation therefore occurs immediately in case (*i*) if $\widetilde{Y}_A \geq Y_M$, or with a lag at Y_F in case (*i*) if $\widetilde{Y}_A < Y_M$ and in cases (*ii*) and (*iii*).

The pivotal imitation cost level that separates attrition and preemption satisfies $\widehat{K} < I$, so a lower imitation cost is necessary but not sufficient for a second mover advantage to exist and attrition to occur. Too see why, consider an industry in which $K = I$. In such an industry a first-mover that invests optimally earns additional monopoly profits until the market size process hits Y_F. In order for a second-mover advantage to arise and firms to be willing to wait, the cost of imitation must be sufficiently low to compensate the second mover for forgoing this monopoly rent. In practice therefore a lower cost for imitators, which constitutes the most likely situation absent IPRs, does not itself ensure that firms have a second-mover advantage or that they will find it desirable to pursue so-called imitation strategies.

Because the strategic interaction depends on the position of the imitation cost K relative to $\widehat{K}$, the comparative statics of this threshold reveal the effect that the main industry parameters, the intensity of product market competition (π_M/π_D) and the characteristics of the demand process (α and σ), have on the nature of the timing game.

Proposition 3 *The more intense product market competition is and the lower are drift and volatility, the more likely it is that preemption occurs, and conversely for attrition:*

$$\frac{\partial \widehat{K}}{\partial\left(\pi_M/\pi_D\right)} < 0,\ \frac{\partial \widehat{K}}{\partial \alpha}, \frac{\partial \widehat{K}}{\partial \sigma} > 0.$$

To provide intuition for the last inequality, recall that because of market uncertainty there is an option value from waiting. So long as that there is an inherent advantage to imitation ($K < I$), then for large enough levels of drift and volatility such that $K < \widehat{K}$, this option value outweighs any preemptive motive to secure monopoly rents. That is to say, *an attrition regime is more likely in industries with greater trend growth and demand volatility.*[16] This is an important observation in our framework, because it identifies a countervailing force to several mechanisms that are highlighted in the rest of the paper. As the next sections show, firm choices regarding technology or licensing and regulator choices of IPR levels generally make dynamic competition more preemptive. One therefore expects attrition to occur relatively rarely, except in those industries in which market uncertainty is significant.

4.3 Imitation cost and industry outcomes

The equilibrium described in Proposition 2 involves several intuitive relationships between the cost of imitation and the dynamic pattern of investments and firm profitability which we describe here successively.

4.3.1 Thresholds

From the expression of Y_F it directly follows that a higher imitation cost raises the standalone duopoly threshold. However raising Y_F does not by itself imply that imitation is delayed in equilibrium at least in an attrition regime, since innovation and hence imitation occur at a stochastic threshold beyond Y_F with positive probability. The effect of imitation cost on innovation and innovation timing must therefore be studied more carefully.

To begin with, there is an inverse relationship between the cost of imitation and the innovation threshold. Under preemption, it is straightforward to establish that $\partial Y_P / \partial K < 0$. Under attri-

[16] One example of real-world conditions that might fit such a framework is the following. Pharmaceutical firms face market conditions that impact product introductions and that can vary significantly across geographic areas. In low- and middle-income countries, economic and demographic drivers often imply high demand growth, but political instability can also result in less demand predictability than in high-income economies, and thus discourage the industry from introducing new treatments or preventives. Managers of big pharmaceutical companies are very aware of such market characteristics, and emphasize that although "pharmaceutical markets in key emerging economies, such as China, India, and Brazil, are expanding at rates of more than 12 percent per year (...) uncertain demand, and political and economic instability in some countries have deterred private investors for decades" (Witty [29], pp. 118 and 124).

tion, the distribution of innovator entry thresholds is shifted leftward so the innovation threshold $\widetilde{Y}_A$ decreases stochastically.

The relationship between the cost of imitation and the imitation threshold, which is monotone under preemption since $\partial Y_F / \partial K > 0$, is again more involved under attrition. Indeed in the waiting game a higher imitation cost results in a higher imitation threshold if the innovation threshold realization is low so that there is a positive lag between innovation and imitation ($\widetilde{Y}_A \leq Y_F$). However if the realization of the innovation threshold is high ($\widetilde{Y}_A > Y_F$), imitation occurs right after innovation, so its distribution is accordingly also shifted leftward by an increase in the cost of imitation to the right of Y_F. Therefore a higher imitation cost delays imitation if innovation occurs early, but indirectly hastens imitation if innovation occurs late. A better measure of the speed of imitation is therefore the gap between innovation and imitation thresholds, which under attrition is equal to $\max \left\{ Y_F, \widetilde{Y}_A \right\} - \widetilde{Y}_A$ and is nondecreasing in K.

In the model, changes in imitation cost therefore have a monotonic effect on both the innovation threshold and the imitation lag. An increase in imitation cost thus accelerates innovation and delays the arrival of imitation conditional upon innovation having occurred.

4.3.2 Equilibrium payoffs

There also exists a useful equilibrium relationship between imitation cost and industry performance. Observe first that because under attrition and preemption competition between firms to secure second- or first-mover advantage results in the dissipation of positional rents, equilibrium firm values have straightforward expressions. In an attrition regime, the equilibrium firm value is $M(Y_M)$ over $\left[0, \widetilde{K}\right]$, $L(Y_L)$ over $\left[\widetilde{K}, \widehat{K}\right]$ and $F(Y_F)$ over $\left[\widehat{K}, \infty\right)$. Moreover viewed as functions of the cost of imitation, $M(Y_M)$ is constant, $L(Y_L)$ is increasing and $F(Y_F)$ is decreasing over the interiors of the relevant ranges. Industry value is therefore nondecreasing in K up until $\widehat{K}$ and decreasing thereafter, and we can conclude that it is at this level where neither attrition nor preemption occur that industry value is maximized, as firms do not have an incentive to dissipate resources by seeking a positional advantage of either sort.

Proposition 4 *The expected industry value is* $\min \{F(Y_F), \max \{L(Y_L), M(Y_M)\}\}$. *Viewed as a function of* K *it is quasiconcave, constant over* $\left(0, \widetilde{K}\right)$ *and has a unique maximum at* $K^\pi = \widehat{K}$.

In economic terms, Proposition 4 establishes that starting from a zero imitation cost, firms benefit *ex-ante* from raising the fixed cost of imitation above $\widetilde{K}$ so as to shield an innovator with positive probability from instantaneous imitation. Moreover, and despite possibly wasteful duplicative fixed costs *ex-post*, raising imitation cost for the second firm even further to $\widehat{K}$ is beneficial for the industry when the endogenous timing of investments is accounted for.

Aside from providing an intuitive characterization of industry value, Proposition 4 is instrumental in the next section in establishing several of the welfare results.

5 Normative analysis

The previous section showed how the nature of the timing game (attrition or preemption) and the dynamics of innovation and imitation are related to the cost of imitation. The social welfare generated by the innovative and imitative products is therefore determined by regulatory choices which affect imitation cost. In this section, we seek to identify socially optimal levels of imitation cost in a second-best framework where regulators set the cost of imitation while firms freely determine the timing of their investments. At first glance this question might seem to just involve a classic trade-off between the reward of innovation and the dynamic deadweight loss from monopoly, since a higher imitation cost *prima facie* raises the optimal threshold for imitation. However as seen in Section 4.3.1, imitation does not necessarily occur at an optimal threshold so an increase in imitation cost has an ambiguous effect on the timing of imitation in an attrition regime, and the optimal imitation cost therefore needs to be studied more carefully.

For our welfare analysis, it is useful to express the social welfare function (6) in terms of producer surplus and the consumer surpluses due to innovation and imitation. Its value in a free entry equilibrium is

$$W(K) = 2\min\left\{F(Y_F), \max\left\{L(Y_L), L(Y_M)\right\}\right\} \\ + \frac{S_M y^\beta}{r-\alpha}\mathbb{E}\left[\widetilde{Y}_I\right]^{-(\beta-1)} + \frac{(S_D - S_M)\, y^\beta}{r-\alpha}\mathbb{E}\left[\max\left\{\widetilde{Y}_I, Y_F\right\}\right]^{-(\beta-1)}. \quad (9)$$

The first summand in (9) is expected industry value, which by Proposition 4 is constant for $K < \widetilde{K}$ and strictly quasiconcave with a maximum at $\widehat{K}$. The second term is the expected consumer surplus due to innovative investment. This term is monotonically increasing in K since a higher imitation cost shifts the distribution of innovation thresholds leftward. The third term is the expected consumer surplus due to the imitator's entry whose relationship with K is ambiguous under attrition, since it is the lag between innovation and imitation and not the timing of imitation itself that increases monotonically with K.

The function $W(K)$ does not have a closed-form expression over its entire range. However its value over $\left[0, \widehat{K}\right]$ has an intuitive bound and there is a semi-closed form over $\left[\widehat{K},\infty\right)$ whose maximum value we are able to calculate, so several properties of the social optimum can be derived (See appendix for proofs of the following propositions and lemmas).

First, the socially optimal imitation cost has a positive lower bound:

Proposition 5 *If the order of investments is endogenous, the socially optimal cost of imitation satisfies* $K^W \geq \widetilde{K}$.

The intuition for this result is straightforward. The first term (producer surplus) in the welfare function is constant over $\left[0, \widetilde{K}\right]$ (Proposition 4) and the second term (consumer surplus from innovation) is increasing since a higher imitation cost accelerates innovation, so only the third term (consumer surplus from imitation) requires more careful consideration. However if $K \leq \widetilde{K}$ the innovation threshold is distributed over $[Y_M, \infty)$, and imitation occurs immediately after innovation. Within this range an increase in imitation cost therefore actually accelerates imitation indirectly, which unambiguously increases the consumer surplus from imitation.

It is therefore never optimal to set the cost of imitation at zero. Rather, it must be sufficiently high so that an innovator expects a phase of monopoly profits with positive probability provided that he innovates at a low enough threshold. Conversely a firm that "wins" the timing game by being more patient than its rival should accordingly pay a minimum price to develop its imitative product, so that the industry avoids the Schererian dynamics described in the introduction. Even if one adheres to the view that IPRs should be abolished altogether (Boldrin and Levine [2]), it is nevertheless important to ascertain that the cost of imitation meets such a threshold.[17]

Second, the positive lower bound of the socially optimal imitation cost can be tightened further if an intuitive condition is met and there is sufficient discounting. The next proposition thus provides a rigorous foundation for strong IPRs based on their dynamic characteristics, under

[17] In practice the lower bound $\widetilde{K}$ can be used to assess initiatives like those in the pharmaceutical industry to reduce the relative cost of imitation and encourage generic competition. In low- and middle-income countries, often characterized by a rapidly expanding and highly uncertain demand which makes attrition more likely (Proposition 3), optimal social welfare in the local market may indeed involve low IPRs and attrition but in all cases requires that a sufficient level of protection be maintained so that there remains a window of market sizes in which an innovator entering sufficiently early is incentivized by a period of monopoly profits.

For example, in order to increase access to antiretroviral drugs to treat HIV infection in the developing world, over the last decades political mobilization has facilitated the production of generic versions of the medicines patented in developed countries (Hoen et al. [16]), which is consistent with our analysis above.

the assumption that the static private entry incentive is socially excessive ($\pi_D > (S_D + 2\pi_D) - (S_M + \pi_M)$) which characterizes many standard oligopoly models (see Section 2.1).

Proposition 6 *If the order of investments is endogenous, the socially optimal cost of imitation satisfies* $K^W \geq \widehat{K}$ *provided that the static private entry incentive is socially excessive and* β *is sufficiently large.*

The proposition establishes that the drift and volatility of market size, through their effect on the discounting parameter β, play a key role in identifying which type of dynamic competition is socially optimal. Specifically, it is in those industries for which drift and volatility are not too large that IPR protection should be set sufficiently high for competition between firms to be preemptive, whereas the issue of optimal IPR levels remains an open question if the drift and volatility are significant.[18]

The proposition is established by means of two lemmas.

Lemma 1 $W(K)$ *has a unique maximum over* $\left[\widehat{K}, \infty\right]$ *and there exists* $\beta_0 > 1$ *such that the socially optimal innovation threshold is*

$$Y_P^W = \begin{cases} \frac{1}{1+\psi} Y_L, & \beta < \beta_0 \\ Y_{NPV}, & \beta \geq \beta_0 \end{cases}, \text{ where } \psi := \frac{1}{1 + \frac{\frac{S_M}{\pi_M}\left(\frac{\pi_M}{\pi_D} - \frac{\beta-1}{\beta}\right)}{\frac{2}{\beta} + \frac{S_D - S_M}{\pi_D}}}.$$

Although we do not have an expression for the socially optimal imitation cost itself, the lemma gives an exact expression for the optimal innovation threshold if firms play a game of preemption (even though preemption thresholds do not themselves have a closed form generally). This optimal innovation threshold lies between the break-even threshold Y_{NPV} and the standalone monopoly threshold Y_L, and since ψ is increasing in β, when β is large enough the optimum is a corner solution that involves setting an arbitrarily high imitation cost $K^W = \infty$. Imitation then never occurs and instead firms race to enter in winner-take-all preemption, the timing of the monopoly innovation having been driven to the competitive threshold by the threat of potential entry.

The exact form of the preemption threshold given in Lemma 1 allows the local maximum of $W(K)$ over $\left[\widehat{K}, \infty\right]$ to be evaluated. There is no corresponding expression for the maximum value

[18] For instance for orphan drugs and rare disease development, the U.S. Food and Drug Administration enacted an enhanced form of IPR protection (Orphan Drug Exclusivity) together with a tax credit that lowers the costs of clinical trials (Grabowski et al. [12]). To the extent that such markets are characterized by low growth and volatility, our analysis offers theoretical support to such regulatory measures.

of $W(K)$ over $\left[0, \widehat{K}\right]$, but if the static private entry incentive is socially excessive then the third term in the welfare function has an intuitive bound involving the profits of the imitating firm, and the values of social welfare over the two different ranges can be compared, establishing the proposition.

Lemma 2 *If the static private entry incentive is socially excessive, then $K^W \geq \widehat{K}$ if*

$$\frac{S_M}{\pi_M} \geq \Omega(\beta)\,,\ \textit{where}\ \Omega(\beta) := \frac{3}{2\beta\left(\left(\frac{\beta}{\beta-1}\right)^{\beta-1} - 1\right)}.$$

The function $\Omega(\beta)$ is decreasing in β with $\lim_{\beta\to\infty} \Omega(\beta) = 0$, so Lemma 2 establishes that with sufficient discounting (*i.e.* sufficiently low drift and volatility) the optimal level of social welfare lies in the range over which dynamic competition is preemptive.

The previous two lemmas provide sufficient conditions, both for high IPRs to be socially optimal ($S_M/\pi_M \geq \Omega(\beta)$) and, provided that this is the case, for optimal IPRs to result in winner-take-all preemption ($\beta \geq \beta_0$).[19] For several standard product market specifications that satisfy the excess static private entry incentive restriction in the proposition, these conditions are straightforward to verify.

Example 1 *(linear demand)* Suppose that the product market is characterized by a linear inverse demand $P = A - BQ$, $A, B > 0$ and that firms have constant unit variable cost c. Then after normalizing by $(A-c)^2/B$, product market outcomes are $(S_M, S_D, \pi_M, \pi_D) = (1/8, 2/9, 1/4, 1/9)$. Solving $S_M/\pi_M = 0.5 = \Omega(\beta)$ numerically gives the threshold for $K^W \geq \widehat{K}$ as $\beta \approx 2.5692$. With these values, β_0 is the upper root of $5\beta^2 - 8\beta - 20 = 0$ which gives $\beta_0 \approx 2.9541$ as the threshold for winner-take-all preemption.

Example 2 *(isoelastic demand)* Suppose that the product market is characterized by an isoelastic inverse demand $P = AQ^{-1/\varepsilon}$, $A > 0, \varepsilon > 1$, and that firms have constant unit variable cost c. Then product market outcomes satisfy $S_M/\pi_M = \varepsilon/(\varepsilon-1)$, $S_D/\pi_D = 4\varepsilon/(\varepsilon-1)$ and $\pi_M/\pi_D = 4(2(\varepsilon-1)/(2\varepsilon-1))^{\varepsilon-1}$. The threshold for the condition $S_M/\pi_M = \varepsilon/(\varepsilon-1) = \Omega(\beta)$ to hold for all ε is $\beta \approx 1.7201$. With these values, β_0 is the upper root of

$$\beta^2\left(4\left(\frac{2(\varepsilon-1)}{2\varepsilon-1}\right)^{\varepsilon-1} - 1\right) - 2\beta - \left(3 - \frac{2}{\varepsilon}\right) = 0,$$

[19] It is useful for the examples in the text to note that β_0 is the upper root of

$$\beta^2 \frac{S_M}{\pi_M}\left(\frac{\pi_M}{\pi_D} - 1\right) + \beta\left(2\frac{S_M}{\pi_M} - \frac{S_D}{\pi_D}\right) - \frac{S_M}{\pi_M} - 2 = 0$$

(see Section A.6).

which gives

$$\beta_0 = \frac{1 + \sqrt{1 + \left(3 - \frac{2}{\varepsilon}\right)\left(4\left(\frac{2(\varepsilon-1)}{2\varepsilon-1}\right)^{\varepsilon-1} - 1\right)}}{4\left(\frac{2(\varepsilon-1)}{2\varepsilon-1}\right)^{\varepsilon-1} - 1}.$$

This is an increasing function of ε, with $\lim_{\varepsilon\to\infty}\beta_0 = \left(1 + \sqrt{1 + 3\left(4\mathrm{e}^{-1/2} - 1\right)}\right) / \left(4\mathrm{e}^{-1/2} - 1\right) \approx 2.3122$, which therefore gives the threshold for winner-take-all preemption in the social optimum for all ε.

Further welfare results can be obtained in specific cases:

Proposition 7 *If the order of investments is endogenous then*

(i) if the consumer surplus from innovation is sufficiently small ($S_M \approx 0$), then $\widetilde{K} < K^W < \widehat{K}$;

(ii) if there is collusion in the product market ($S_D + 2\pi_D = S_M + \pi_M$), then $K^W \geq \widehat{K}$.

Part (*i*) complements the previous results of the section by showing that there exist conditions under which values of the imitation cost in the attrition range constitute a social optimum. This is not obvious *a priori* since there is no closed form expression for $W(K)$ over this range, but it is nevertheless possible to show that $\lim_{K\to\widehat{K}^-} W'(K) < 0$. Intuitively, in this case the innovator does not contribute measurably to the consumer surplus and producer surplus is locally insensitive to imitation cost at $\widehat{K}$. Therefore, decreasing imitation cost from $\widehat{K}$ incentivizes imitation and improves welfare. Part (*ii*) reflects the opposite situation, where imitator entry does not affect consumer surplus because of collusion in the product market. Then only innovation contributes to social welfare, and the role of imitation cost is to incentivize innovation in a preemption regime.

We complete our analysis by comparing welfare levels with endogenous innovation with those obtained in Section 3 where the investment sequence is predetermined. Consider a regulator who can choose both whether firms compete dynamically and the level of imitation cost. This situation might arise, for example, if a regulator seeks to accelerate innovation in a particular market by contracting with one firm directly. A regulator then sets the optimal imitation cost, which is $K^W = 0$ if firm roles are predetermined and $K^W > \widehat{K}$ if roles are endogenous and β is sufficiently large. Because we can evaluate the innovation threshold and the optimal welfare level in this case (Lemma 1), it is possible to compare both situations. The upshot is that fostering dynamic competition to innovate first is a valuable policy instrument in such industries, and can be shown to result in a higher level of welfare:

Proposition 8 *If product market competition is sufficiently intense, the welfare optimum with endogenous innovation under the conditions of Proposition 6 is greater than with predetermined roles.*

Observe moreover that the optimal imitation cost levels differ significantly depending upon whether firm roles are predetermined or endogenous. This is particularly relevant because the patent design literature generally takes the value of innovation as given and does not endogenize its timing,[20] which results in a starkly different conclusion and a suboptimal prescription regarding imitation cost, illustrating the importance of accounting for industry dynamics rather than relying on static competition in the product market.

6 Extensions

In this section, we discuss how some other aspects of innovation and imitation fit into our model. One of these is the ability of an innovating firm to raise the imitator's entry barrier by making its product more costly to reverse engineer or by strengthening its patentability. Another is contracting between the innovator and the imitator, which can consist in a buyout of the rival firm or a technology transfer. From a formal standpoint these different extensions add an intermediate decision to the investment game during the monopoly phase, either at the moment of the innovator's entry or when the imitator invests. By raising the value of innovating, these extensions favor first-mover advantage and the emergence of preemption regimes, with contrasting implications for imitation timing and welfare.

6.1 Endogenous imitation cost

Suppose that the innovating firm may affect the cost of imitation by varying the amount of either technical or legal protection. In case of technical protection, the cost of reverse engineering can be raised by increasing product complexity. For example an innovating firm can render its product more difficult to disassemble, or even add misleading complexity (Samuelson and Scotchmer [23]). In the case of legal protection, wider patents imply higher costs for inventing

[20] Admittedly with either perfect competition or monopoly instead of duopoly, the timing of innovation follows a straightforward investment rule. As Rockett [22] observes, accordingly "most [models] take the identity of the innovator as given." Denicolò [5] is an exception, but his patent race model does not allow for attrition and second-mover advantage.

around to develop a non-infringing imitation, and firms may decide to pursue patent protection more or less aggressively (Encaoua et al. [8]).

Such choices are incorporated into the model by supposing that when it invests at threshold Y_i, an innovating firm chooses how much additional cost, ρ, to incur in order to raise the imitation cost by an amount $f(\rho)$.[21] The imitation cost increase is instantaneous and the function f is taken to be twice differentiable, increasing and concave with $f(0) = 0$ and $\lim_{\rho \to 0} f'(\rho) = \infty$. The fixed costs of the innovator and imitator are accordingly $I(\rho) := \overline{I} + \rho$ and $K(\rho) := \overline{K} + f(\rho)$ where $\overline{I}$ and $\overline{K}$ represent baseline values where no cost-raising expenditure is undertaken.

Proceeding by backward induction, the cost-raising effort affects the imitator payoff $F(Y)$ and standalone threshold Y_F through $K(\rho)$. At the moment of innovation therefore, an innovator entering at the threshold Y_i faces the decision problem

$$\max_{\rho \in \mathbb{R}_+} L_{\mathrm{E}}(Y_i, \rho) := \left(\frac{\pi_M}{r-\alpha} Y_i - \overline{I} - \rho\right)\left(\frac{y}{Y_i}\right)^{\beta} - \frac{\pi_M - \pi_D}{r-\alpha} \frac{y^{\beta}}{\left[\max\{Y_i, Y_F(\rho)\}\right]^{\beta-1}}.$$

Let $\rho^*(Y_i)$ denote the solution to this problem. At an interior solution, $Y_F(\rho^*) > Y_i$ and ρ^* satisfies

$$\beta\left(\frac{\pi_M}{\pi_D} - 1\right) f'(\rho^*) = \left(\frac{Y_F(\rho^*)}{Y_i}\right)^{\beta}.$$

A straightforward comparative static argument establishes that the optimal cost-raising effort is increasing in the investment threshold and decreasing in the baseline imitation cost.[22]

To proceed further we focus on the situation where $\overline{K} \geq \widehat{K}$ so the dynamic competition is preemptive.[23] Allowing the cost of imitation to be endogenous results in a higher leader payoff $L_{\mathrm{E}}(Y, \rho^*(Y))$ and a lower follower payoff $F_{\mathrm{E}}(Y, \rho^*(Y)) = F(Y)|_{K=\overline{K}+\rho^*(Y)}$ than when this cost is exogenous. This makes the investment game even more preemptive. Since equilibrium payoffs are pegged to the follower value under preemption, firms have a lower expected value in equilibrium. To avoid this penalizing outcome firms would prefer to both commit *ex-ante* not to exert any cost-raising effort if they innovate. One way to achieve such a commitment is by agreeing to an open

[21] See Huisman and Kort [18] for a model of preemption with firms competing on both the timing and magnitude of investment.

[22] The latter property is in line with the situation of biopharmaceutical firms (see footnote 6 above) where greater reliance is placed on patenting in the medications segment in which natural entry barriers are low than in the vaccines segment.

[23] This restriction relates specifically to optimal stopping and is not necessary for the economic analysis. For high values of the innovation threshold Y_i, corner solutions $\rho^* = 0$ arise that result in a kink of L_{E}. In such cases the threshold strategies firms are assumed to use needn't be optimal investment policies. Under preemption however such thresholds occur only off the equilibrium path.

or common technological standard, a measure which is not desirable if roles are predetermined (Proposition 1).

Proposition 9 *If the cost of imitation is endogenous and the baseline investment game is preemptive firms benefit from agreeing ex-ante to a common standard.*

6.2 Takeover and licensing

Contracts ranging from acquisitions and pay-for-delay agreements to joint ventures and licensing contracts typically play an important role in innovation decisions. These arrangements have contrasting effects on investment incentives that can be incorporated into our model. Assume that firms can contract once to transfer either productive assets or technology in exchange for a lump sum transfer, φ, from the innovator firm to the imitator, and that the contract is written by the innovator who holds all the bargaining power.

Because of the efficiency effect ($\pi_M > 2\pi_D$) it is profitable for an innovator to pay its rival not to subsequently enter the market if it can, by taking over its assets or engaging in some equivalent measures.[24] Proceeding by backward induction, in the continuation phase that begins when innovation occurs at a threshold Y_i, the remaining firm's expected payoff is $F(Y_i)$. This continuation payoff constitutes a participation constraint in any contract that the innovator offers. The innovating firm therefore offers a transfer $\varphi^*(Y_i) = V_D(Y_i)$ at the moment when it invests. The leader payoff with takeovers is therefore

$$L_{\mathrm{T}}(Y) := \left(\frac{\pi_M}{r-\alpha}Y - I\right)\left(\frac{y}{Y}\right)^{\beta} - F(Y).$$

The effect on the investment game is straightforward. The follower payoff is unchanged whereas the leader payoff is larger than without buyouts, rendering preemption more likely. All else equal the effect of the takeover option on the leader payoff depends on the strength of the efficiency effect, and if it is sufficiently strong or in industries with sufficiently high demand growth or volatility (if $\pi_M/\pi_D \geq \beta + 1$) attrition does not occur for any level of K.

Whether the possibility of takeovers runs in the interest of the industry or not depends on the cost of imitation. Under preemption expected profits are pegged to the follower value and therefore unaffected by takeovers, whereas if $K < \widehat{K}$, the industry functions naturally under attrition and

[24] In the pharmaceutical industry pay-for-delay agreements can arise, generally in the context of a patent infringement suit brought by a brand-name company against a generic producer that challenges the innovator's IPRs (see Danzon [4]).

firm values are pegged to the leader value. Industry profit then increases if takeovers are allowed, so one would expect an active market for acquisitions to develop in such industries, and all the more so if demand growth and volatility are high.

If a takeover is not possible an innovator must contend with follower entry but can recoup revenue from the imitator's investment through a license fee. Suppose that $K = \overline{K} + K_I$ where $\overline{K}$ is an incompressible level of imitation cost reflecting such items as distribution and marketing expenses and K_I denotes the part of the product development cost that can be eliminated by a technology transfer. Because licensing does not allow the innovator to push back the moment of imitation, the optimal policy is to set the maximum license fee consistent with the participation constraint at the moment of imitation, $\varphi^* = -K_I$.

Proceeding by backward induction, the expected revenue from licensing adds a positive term to the leader payoff which becomes

$$L_{\mathrm{L}}(Y) := \left(\frac{\pi_M}{r-\alpha}Y - I\right)\left(\frac{y}{Y}\right)^{\beta} - \left(\frac{\pi_M - \pi_D}{r-\alpha}\max\{Y, Y_F\} - K_I\right)\left(\frac{y}{\max\{Y, Y_F\}}\right)^{\beta}.$$

As the leader payoff shifts up to the left of Y_F while leaving the follower payoff function unchanged, the investment game is more preemptive with licensing than with takeovers. However whereas the welfare consequences of takeovers are ambiguous (firms weakly benefit and the consumer surplus from innovation increases because product innovation occurs earlier, but the consumer surplus from imitation is eliminated), the welfare consequences of licensing are unambiguously positive. The visible effect of licensing is the reduction in the duplication of R&D efforts as in Gallini [11], and an additional indirect benefit stems from the acceleration of innovation which raises consumer surplus.

Thus,

Proposition 10 *With contracting between the innovator and the imitator* (*i*) *takeovers are the preferred instrument of an innovator and raise industry profit if* $K < \widehat{K}$ *whereas* (*ii*) *licensing is Pareto-improving.*

7 Conclusion

We have sought to model the dynamic allocation of resources to innovation and imitation, explicitly incorporating the interrelated investment decisions under uncertainty of imperfectly competitive firms. As compared with the classic literature on innovation and patents, endogenizing the time at which innovation and imitation occur allows us to highlight a novel policy channel, in which IPR levels act upon welfare through their effect on dynamic competition.

The main message that emerges from our analysis is a broadly familiar one, insofar as we find that IPRs must be important enough to provide a sufficient incentive for innovation. By integrating the theory of investment under uncertainty into the analysis of innovation incentives, we are able to sharpen this general perspective by pinpointing the role of specific market characteristics which act as key determinants of investment, and thus to provide a grounding for strong IPRs in circumstances that seem most likely to be present in mature industries. In such industries we then find that the barriers to imitation should be sufficiently high so as to render dynamic competition between firms preemptive, and if discounting is important enough competition should take the form of a winner-take-all contest. Moreover, a regulator who would attempt to accelerate innovation by contracting with one firm directly would lose out on the social benefits of dynamic competition, as the welfare achieved with optimal IPRs is lower with predetermined roles than with endogenous innovation.

In those industries in which growth and volatility are relatively high on the other hand, which are those most typically associated with vibrant innovation, attrition may be effective in ensuring that additional benefits of imitation resulting from greater product market competition do not arrive excessively late. Even then some degree of IPR protection can be needed if the cost of imitation is extremely low, in order to ensure that a firm that develops an imitative product as the winner of the attrition game nevertheless pays a high enough entry cost so an industry does not become mired in inefficient dynamics.

In practice, antitrust and industrial policy decisions commonly focus on static product market characteristics. The demand characteristics that we have highlighted, demand growth and volatility, play at least as significant a role in determining the investment incentives and product development decisions, and as such should naturally underlie any determination of optimal IPRs. As we have argued throughout in the footnotes, the policy measures taken in at least one emblematic innovation-intensive industry seem to have been broadly consistent with such an analysis.

References

[1] Billette de Villemeur E, Ruble R, Versaevel B (2014) Investment timing and vertical relationships, International Journal of Industrial Organization 33:110-123.

[2] Boldrin M, Levine D (2013) The case against patents, Journal of Economic Perspectives 27(1):3-22.

[3] Chevalier-Roignant B, Trigeorgis L (2011) Competitive Strategy: Options and Games, (Cambridge: MIT Press).

[4] Danzon P (2014) Competition and antitrust cases in the pharmaceutical industry, mimeo.

[5] Denicolò V (1996) Patent races and optimal patent breadth and length, Journal of Industrial Economics 44(3):249-265.

[6] Dixit A, Pindyck R (1994) Investment under Uncertainty, (Princeton: Princeton University Press).

[7] Denicolò V, Franzoni L (2010) On the winner-take-all principle in innovation races, Journal of the European Economic Association 8(5):1133-1158.

[8] Encaoua D, Guellec D, Martinez C (2006) Patent systems for encouraging innovation: Lessons from economic analysis, Research Policy 35(9):1423-1440.

[9] Femminis G, Martini G (2011) Irreversible investment and R&D spillovers in a dynamic duopoly, Journal of Economic Dynamics and Control 35(7):1061-1090.

[10] Fudenberg D, Tirole J (1985) Preemption and rent equalization in the adoption of new technology, Review of Economic Studies 52(3):383-401.

[11] Gallini N (1992) Patent policy and costly imitation, RAND Journal of Economics 23(1):52-63.

[12] Grabowski HG, DiMasi JA, Long G, (2015) The roles of patents and research and development incentives in biopharmaceutical innovation, Health Affairs 34(2):302-310.

[13] Green J, Scotchmer S (1995) On the division of profit in sequential innovation, RAND Journal of Economics 26(1):20-33.

[14] Hendricks K, Weiss A, Wilson C (1988) The war of attrition in continuous time with complete information, International Economic Review 29(4):663-680.

[15] Henry E (2010) Runner-up patents: is monopoly inevitable?, Scandinavian Journal of Economics 112(2):417-440.

[16] Hoen E, Berger J, Calmy A, Moon S (2011) Driving a decade of change: HIV/AIDS, patents and access to medicines for all, Journal of the International AIDS Society 14:15-26.

[17] Hoppe H (2000) Second-mover advantages in the strategic adoption of new technology under uncertainty, International Journal of Industrial Organization 18:315-338.

[18] Huisman K, Kort P (2015) Strategic capacity investment under uncertainty, RAND Journal of Economics 46(2):376-408.

[19] La Manna M, MacLeod R, de Meza D (1989) The case for permissive patents, European Economic Review 33(7):1427-43

[20] Mankiw G, Whinston M (1986) Free entry and social inefficiency, RAND Journal of Economics 17(1):48-58.

[21] Reinganum J (1981) On the diffusion of a new technology: a game theoretic approach, Review of Economic Studies 48(3):395-405.

[22] Rockett K (2010) Property rights and invention in Bronwyn H and Rosenberg N (eds) Handbook of the Economics of Innovation (Amsterdam: North-Holland).

[23] Samuelson P, Scotchmer S (2002) The law and economics of reverse engineering, Yale Law Journal 111:1575-1663.

[24] Steg J-H, Thijssen J (2015) Quick or persistent? Strategic investment demanding versatility, Bielefeld: Center for Mathematical Economics working paper 541.

[25] Thijssen J, Huisman K, Kort P (2006) The effects of information on strategic investment and welfare, Economic Theory 28(2):399-424.

[26] Thijssen J, Huisman K, Kort P (2012) Symmetric equilibrium strategies in game theoretic real option models, Journal of Mathematical Economics 48(4):219-225.

[27] Scherer F, Watal J (2002) Post-TRIPS options for access to patented medicines in developing countries, Journal of International Economic Law 5(4):913-939

[28] Wilson P (2010) Giving developing countries the best shot: An overview of vaccine access and R&D, Oxfam International, 28 pages.

[29] Witty A (2010) New strategies for innovation in global health: A pharmaceutical industry perspective, Health Affairs 30(1):118-126.

A Appendix

A.1 Duopoly option value V_D

Once the rival firm has invested at the market size threshold Y any remaining firm holds a standard growth option (Dixit and Pindyck [6]) whose value is obtained by solving the optimal stopping problem

$$V_D\left(Y\right) = \sup_{\tau \geq t} \mathbb{E}_Y \left(\int_{\tau}^{\infty} \pi_D Y_s e^{-rs} ds - K e^{-r\tau} \right).$$

From the Hamilton-Jacobi-Bellman equation, the value function $V_D(Y)$ satisfies

$$rV_D\left(Y\right) dt = \mathbb{E}_Y dV_D\left(Y\right)$$

and expanding the right-hand side using Itô's lemma yields the ordinary differential equation that $V_D(Y)$ solves in the continuation region,

$$rV_D = \alpha Y V_D' + \frac{1}{2}\sigma^2 Y^2 V_D'',$$

along with the boundary and smooth pasting conditions

$$\begin{aligned} V_D\left(0\right) &= 0 \\ V_D\left(Y_F\right) &= \frac{\pi_D}{r-\alpha} Y_F - K \\ V_D'\left(Y_F\right) &= \frac{\pi_D}{r-\alpha}. \end{aligned}$$

The function $A_1 Y^{\beta_1} + A_2 Y^{\beta_2}$ is a candidate solution. The associated fundamental quadratic is $0.5\sigma^2\beta\left(\beta-1\right) + \beta\alpha - r = 0$ which has two roots of which only

$$\beta = \frac{1}{2} - \frac{\alpha}{\sigma^2} + \sqrt{\left(\frac{1}{2} - \frac{\alpha}{\sigma^2}\right)^2 + \frac{2r}{\sigma^2}}$$

is positive. Setting $A_2 = 0$ to satisfy the first boundary condition, it follows from the other conditions that

$$Y_F = \frac{\beta}{\beta-1}\frac{r-\alpha}{\pi_D}K$$

and that A_1 is $\left(\pi_D/\left(\beta\left(r-\alpha\right)\right)\right)\left[Y_F\right]^{-(\beta-1)}$ which yields the expression for A_D in the text.

A.2 Leader, follower and simultaneous investment payoff functions

A.2.1 Derivation of F, L and M

The expression for the follower payoff (3) is simply the discounted value of $V_D(Y)$. The expression (4) is obtained from the definition of $L(Y)$ by:

$$
\begin{aligned}
&\mathbb{E}_y\left[\int_{\tau(Y)}^{\tau(\max\{Y,Y_F\})} \pi_M Y_s e^{-rs} ds - e^{-r\tau(Y)} I + \int_{\tau(\max\{Y,Y_F\})}^{\infty} \pi_D Y_s e^{-rs} ds\right] \\
&= \mathbb{E}_y\left[\int_{\tau(Y)}^{\infty} \pi_M Y_s e^{-rs} ds - e^{-r\tau(Y)} I - \int_{\tau(\max\{Y,Y_F\})}^{\infty} (\pi_M - \pi_D) Y_s e^{-rs} ds\right] \\
&= \left(\frac{\pi_M}{r-\alpha} Y - I\right)\left(\frac{y}{Y}\right)^{\beta} - \frac{\pi_M - \pi_D}{r-\alpha} \frac{y^{\beta}}{[\max\{Y, Y_F\}]^{\beta-1}}.
\end{aligned}
$$

Finally the derivation of the simultaneous payoff (5) is similar to (4), noting that the flow profit upon investment is π_D instead of π_M and that there is no follow-on entry so the second integral dropped altogether.

A.2.2 Behavior of F, L and M

The most straightforward continuation payoffs to study are M and F, so we start with these before turning to L.

First, $M(Y) = ((\pi_D/(r-\alpha))\,Y - I)\,(y/Y)^{\beta}$ is quasiconcave in Y over (y, ∞) with a maximum at $Y_M := (\beta(r-\alpha)I)/((\beta-1)\pi_D)$ such that $Y_M > y$ given the assumption $Y_0 \leq Y_{NPV}$, and $\lim_{Y\to\infty} M(Y) = 0$. Observe that $M(Y)$ is independent of K.

Next, consider the follower payoff F. If $K < ((\beta-1)\pi_D y)/(\beta(r-\alpha))$, then $Y_F < y$ and F is a strictly decreasing function of Y over $[y, \infty)$ whereas if $K \geq ((\beta-1)\pi_D y)/(\beta(r-\alpha))$ then F is constant over $[y, Y_F]$ and decreasing over (Y_F, ∞), with $\lim_{Y\to\infty} F(Y) = 0$. An increase in K has two effects on $F(Y)$: first Y_F increases so the range of values over which the graph of F is horizontal weakly increases, and second the value of $F(Y)$ decreases for all K so the entire graph is shifted downward.

The comparison of $M(Y)$ and $F(Y)$ is straightforward: $Y_M < (>)\, Y_F$ if and only if $I < (>)\, K$ and if $K < I$ then $M(Y) < F(Y)$ for all Y whereas if $K > I$ then $M(Y) > F(Y)$ over a half-line $(\underline{Y}_M, \infty)$ where $\underline{Y}_M \in (y, Y_M)$.

Finally, the leader payoff $L(Y)$ is generally defined in parts according to whether Y is greater or smaller than Y_F. If $K < ((\beta-1)\pi_D y)/(\beta(r-\alpha))$, necessarily $Y_F < Y$ so that $L(Y) = M(Y)$ whereas if $K \geq ((\beta-1)\pi_D y)/(\beta(r-\alpha))$ then

$$L(Y) = \begin{cases} \left(\frac{\pi_M}{r-\alpha}Y - I\right)\left(\frac{y}{Y}\right)^\beta - \frac{\pi_M-\pi_D}{r-\alpha}Y_F\left(\frac{y}{Y_F}\right)^\beta, & Y < Y_F \\ M(Y), & Y \geq Y_F. \end{cases}$$

L is therefore continuous with a possible kink at Y_F. The first part of L is quasiconcave in Y with a maximum at $Y_L := (\beta(r-\alpha)I)/((\beta-1)\pi_M)$ such that $Y_L > y$ given the assumption on Y_0.

An increase in K has two effects on $L(Y)$: first Y_F increases which reduces the range over which the graphs of L and M overlap, and second the value of L over its first part increases by a constant inducing a parallel shift upward in the graph of L. Because $\pi_M > \pi_D$, $Y_L < Y_M$ for all K whereas $Y_L < (>)\, Y_F$ if and only if $I < (>)\, (\pi_M/\pi_D)\, K$. Given the ranking of Y_M and Y_F discussed above it follows that whenever $K \leq I$, L has two local maxima, Y_L and Y_M. It is shown in the next subsection that there exists a critical threshold $\widetilde{K} < I$ such that Y_L (Y_M) is the global maximum of L if and only if $K > \widetilde{K}$ $(K < \widetilde{K})$.

The relationship between L and M is straightforward, as L is greater than M over (y, Y_F) and the two functions coincide over $[Y_F, \infty)$. To compare L and F first note that for sufficiently low values of K, $L < F$ for all Y, but otherwise the two functions can intersect. It is also shown in the next subsection that there exists a critical threshold $\widehat{K} < I$ at which *i*) Y_L is a global maximum of L and *ii*) both L and F attain the same maximum value at Y_L and Y_F respectively. Above this threshold, direct calculations establish that either $K < I$ in which case $F > L$ over an interval $\left(\overline{Y}_P, \underline{Y}_F\right) \subseteq (y, Y_F)$ or $K \geq I$ in which case $L \geq F$ for all Y.

A.2.3 Critical thresholds $\widetilde{K}$ and $\widehat{K}$

The threshold $\widetilde{K}$ solves $L(Y_L) = M(Y_M)$. If it were the case that $\widetilde{K} < (\pi_D/\pi_M)I$ so that $\max\{Y_L, Y_F\} = Y_L$, then after substitution $L(Y_L) = ((\pi_D/(r-\alpha))\, Y_L - I)\,(y/Y_L)^\beta = M(Y_L) < M(Y_M)$ since Y_M is the unique global maximizer of M. Hence it must be that $\widetilde{K} \geq (\pi_D/\pi_M)I$, and therefore satisfies

$$\left(\frac{\pi_M}{r-\alpha}Y_L - I\right)\left(\frac{y}{Y_L}\right)^\beta - \frac{\pi_M-\pi_D}{r-\alpha}Y_F\left(\frac{y}{Y_F}\right)^\beta = \left(\frac{\pi_D}{r-\alpha}Y_S - I\right)\left(\frac{y}{Y_S}\right)^\beta.$$

Multiplying both sides by $(Y_F/y)^\beta$ (note that $Y_F/Y_L = (\pi_M/\pi_D)\left(\widetilde{K}/I\right)$ and $Y_F/Y_M = \widetilde{K}/I$ here) and substituting for remaining Y_L, Y_F and Y_M terms yields

$$\frac{1}{\beta-1}\left(\frac{\pi_M}{\pi_D}\right)^\beta \frac{\left[\widetilde{K}\right]^\beta}{I^{\beta-1}} + \frac{\beta}{\beta-1}\left(1-\frac{\pi_M}{\pi_D}\right)\widetilde{K} = \frac{1}{\beta-1}\frac{\left[\widetilde{K}\right]^\beta}{I^{\beta-1}},$$

which has a unique positive solution,

$$\widetilde{K} = \left(\frac{\beta\left(\frac{\pi_M}{\pi_D}-1\right)}{\left(\frac{\pi_M}{\pi_D}\right)^\beta - 1}\right)^{\frac{1}{\beta-1}} I.$$

The threshold $\widehat{K}$ solves $L(Y_L) = F(Y_L)$. If it were the case that $K \leq \widetilde{K}$ then as shown above $L(Y_L) \leq M(Y_M)$ would hold, whereas $M(Y_M) < F(Y_L)$. Hence it must be that $\widehat{K} > \widetilde{K}$ and $\max\{Y_L, Y_F\} = Y_F$. The threshold $\widehat{K}$ therefore solves

$$\left(\frac{\pi_M}{r-\alpha}Y_L - I\right)\left(\frac{y}{Y_L}\right)^\beta - \frac{\pi_M - \pi_D}{r-\alpha}Y_F\left(\frac{y}{Y_F}\right)^\beta = \frac{y^\beta}{\beta\left[Y_F\right]^{(\beta-1)}}.$$

Multiplying both sides by $(Y_F/y)^\beta$ (note that here $Y_F/Y_L = (\pi_M/\pi_D)\left(\widehat{K}/I\right)$) and substituting for remaining Y_L and Y_F terms yields

$$\frac{I}{\beta-1}\left(\frac{\pi_M}{\pi_D}\right)^\beta \frac{\left[\widehat{K}\right]^\beta}{I^{\beta-1}} + \frac{\beta}{\beta-1}\left(1-\frac{\pi_M}{\pi_D}\right)\widehat{K} = \frac{\widehat{K}}{\beta-1}.$$

which has a unique positive solution,

$$\widehat{K} = \left(\frac{1+\beta\left(\frac{\pi_M}{\pi_D}-1\right)}{\left(\frac{\pi_M}{\pi_D}\right)^\beta}\right)^{\frac{1}{\beta-1}} I. \tag{11}$$

A.3 Proposition 2 (endogenous investment symmetric equilibrium)

A.3.1 Part (i)

There are two subcases for this part, $K \leq \widetilde{K}$ and $\widetilde{K} < K < \widehat{K}$ (see Figures 1 and 2).

$K \leq \widetilde{K}$ **subcase** This is a standard war of attrition (in thresholds) over (Y_M, ∞). Therefore provided that firms do not move with positive probability in $[y, Y_M]$ the equilibrium distribution follows from the argument in Hendricks et al. [14]. In a nondegenerate (and symmetric) mixed strategy equilibrium, firms randomize investment triggers continuously over $[Y_M, \infty)$. To derive the equilibrium distribution assume that firm $j \neq i$ randomizes her investment trigger at $t = 0$ according to a cumulative distribution function G. Firm i's expected payoff from investing at a threshold $Y_i \geq Y_M$ is

$$\int_{Y_M}^{\infty} V(Y_i, s) g(s) ds = \int_{Y_M}^{Y_i} F(s) g(s) ds + (1 - G(Y_i)) M(Y_i).$$

Firm i will randomize if $\partial \left(\int_{Y_M}^{\infty} V(Y_i, s) g(s) ds \right) / \partial Y_i = 0$ over the support, that is if G satisfies $\left[F(Y) - M(Y) \right] g(Y) = -M'(Y) \left[1 - G(Y) \right]$ for all $Y \in (Y_M, \infty)$. As the same condition holds for firm $-i$, the equilibrium distribution for each firm is

$$G_\alpha(Y) = 1 - \exp \int_{Y_S}^{Y} \frac{M'(s)}{F(s) - M(s)} ds$$

and substituting for F and M gives

$$G_\alpha(Y) = \begin{cases} 0, & Y < Y_M \\ 1 - \left(\frac{Y}{Y_M} \right)^{\beta \frac{I}{I-K}} e^{-\beta \frac{I}{I-K} \left(\frac{Y}{Y_M} - 1 \right)}, & Y \geq Y_M. \end{cases}$$

Then, the distribution of the equilibrium innovation threshold for the industry, $\widetilde{Y}_A$, is just that of the minimum of the firm thresholds, $G_A(Y) = 1 - (1 - G_\alpha(Y))^2$.

It is claimed in the text (Section 4.3.1) that an increase in imitation cost accelerates innovation. As for $Y > Y_M$

$$\frac{\partial G_\alpha}{\partial K} = \left(\frac{Y}{Y_M} \right)^{\beta \frac{I}{I-K}} e^{-\beta \frac{I}{I-K} \left(\frac{Y}{Y_M} - 1 \right)} \beta \frac{I}{(I-K)^2} \left(\frac{Y}{Y_M} - 1 - \ln \frac{Y}{Y_M} \right) > 0$$

(the last bracketed term is positive by the logarithm inequality), the distribution of each firm's innovation threshold (and therefore of their minimum $\widetilde{Y}_A$) is shifted leftward by an increase in imitation cost.

$\widetilde{K} < K < \widehat{K}$ **subcase** This is also a war of attrition but since $L(Y)$ is decreasing over (Y_L, Y_F), increasing over (Y_F, Y_M) and decreasing over (Y_M, ∞) its form is nonstandard. To identify the support of the mixed strategies, observe that any $Y_i \in (y, Y_L)$ is a strictly dominated strategy and no player puts positive probability on Y_L in a symmetric equilibrium. Similarly any $Y_i \in \left(\overline{Y}_L, Y_M\right)$ is strictly dominated by Y_M, and no player puts positive probability on Y_M in a symmetric equilibrium. Investment thresholds are therefore continuously distributed over $\left[Y_L \overline{Y}_L\right] \cup [Y_M, \infty)$.

Letting G_β denote the equilibrium distribution, assume that firm $j \neq i$ randomizes her investment trigger. Firm i randomizes if $\partial \left(\int_{Y_S}^{\infty} V(Y_i, s) g_\beta(s) ds\right) / \partial Y_i = 0$ over the support, that is if G_β satisfies $[F(Y) - L(Y)]\, g_\beta(Y_i) = -L'(Y)\,[1 - G_\beta(Y)]$ for all $Y \in \left(Y_L, \overline{Y}_L\right)$ and $[F(Y) - M(Y)]\, g_\beta(Y) = -M'(Y)\,[1 - G_\beta(Y)]$ for all $Y \in (Y_M, \infty)$. The former condition holds for

$$
\begin{aligned}
G_{\alpha'}(Y) &= 1 - \exp \int_{Y_L}^{Y} \frac{L'(s)}{F(Y_F) - L(s)} ds \\
&= \frac{L(Y_L) - L(Y)}{F(Y_F) - L(Y)}
\end{aligned}
$$

while the latter condition is satisfied by G_α so that the equilibrium distribution is

$$
G_\beta(Y) = \begin{cases} 0, & Y < Y_L \\ G_{\alpha'}(Y), & Y_L \leq Y \leq \overline{Y}_L \\ G_{\alpha'}\left(\overline{Y}_L\right), & \overline{Y}_L < Y < Y_M \\ G_{\alpha'}\left(\overline{Y}_L\right) + \left(1 - G_{\alpha'}\left(\overline{Y}_L\right)\right) G_\alpha(Y), & Y \geq Y_M. \end{cases}
$$

To establish that an increase in imitation cost accelerates innovation as claimed in the text, observe that $L(Y_L) - L(Y)$ is independent of K while $F(Y_F) - L(Y)$ is decreasing in K so $\partial G_{\alpha'}/\partial K > 0$ and that $\partial \overline{Y}_L / \partial K > 0$ so that, given that $\partial G_\alpha / \partial K > 0$ as seen above, $\partial G_\beta / \partial K > 0$ over the relevant range.

A.3.2 Part *(ii)*

There are two subcases for this part, $\widehat{K} < K < I$ and $K \geq I$ (see Figures 3 and 4). It is simpler to begin with the subcase $K \geq I$ which is standard.

$K \geq I$ subcase For $K \geq I$, $L(Y_F) \geq F(Y_F)$ so there exists a unique $Y_P < Y_F$ such that $L(Y_P) = F(Y_F)$. The preemption range is (Y_P, Y_F), and in this range firms seek to innovate before their rival for any $Y_{-i} \in (Y_P, Y_F)$. In equilibrium both firms set $Y_i = Y_P$ which by the tie-breaking rule results in either firm innovating at Y_P with equal probability.

As joint investment equilibria also arise in preemption models, it must be verified that this is not the case here. Investment at the optimal simultaneous investment threshold Y_M by both firms results in a payoff $M(Y_M)$ whereas for $K \geq I$ (for $K > \widetilde{K}$ in fact), $L(Y_L) > M(Y_M)$, so joint investment cannot be an equilibrium.

$\widehat{K} < K < I$ **subcase** For $\widehat{K} < K < I$, the condition $L(Y) = F(Y)$ has two roots $Y_P, \overline{Y}_P$ with $Y_P < Y_L$ and $\overline{Y}_P \in (Y_L, Y_F)$. For a given $Y > \overline{Y}_P$, $L(Y) \leq F(Y_F)$ so playing beyond the preemption range $\left(Y_P, \overline{Y}_P\right)$ is a dominated strategy. Over the preemption range $\left(Y_P, \overline{Y}_P\right)$ firms preempt one another as in the previous subcase and in equilibrium both firms invest at Y_P, which by the tie-breaking rule (7) results in either firm investing at Y_P with equal probability.

A.3.3 Part (iii)

If $K = \widehat{K}$, then $L(Y_L) = F(Y_F)$ and the only symmetric equilibrium is (Y_L, Y_L).

A.4 Proposition 3 (comparative statics of $\widehat{K}$)

For the comparative statics in (π_M/π_D) evaluating the relevant partial derivatives and rearranging yields

$$\frac{\partial \widehat{K}}{\partial (\pi_M/\pi_D)} = -\beta \frac{\frac{\pi_M}{\pi_D} - 1}{\frac{\pi_M}{\pi_D}\left(1 + \beta\left(\frac{\pi_M}{\pi_D} - 1\right)\right)} \widehat{K}$$

so $\partial \widehat{K}/\partial(\pi_M/\pi_D) < 0$ directly, whereas for the comparative statics in β

$$\frac{\partial \widehat{K}}{\partial \beta} = \left(-\frac{\ln\left(\frac{1+\beta\left(\frac{\pi_M}{\pi_D}-1\right)}{\left(\frac{\pi_M}{\pi_D}\right)^{\beta}}\right)}{\beta - 1} + \frac{\left(\left(\frac{\pi_M}{\pi_D} - 1\right)\left(\frac{\pi_M}{\pi_D}\right)^{\beta} - \left(1 + \beta\left(\frac{\pi_M}{\pi_D} - 1\right)\right)\left(\frac{\pi_M}{\pi_D}\right)^{\beta} \ln \frac{\pi_M}{\pi_D}\right)\left(\frac{\pi_M}{\pi_D}\right)^{\beta}}{\left(1 + \beta\left(\frac{\pi_M}{\pi_D} - 1\right)\right)\left(\frac{\pi_M}{\pi_D}\right)^{2\beta}} \right) \frac{\widehat{K}}{\beta - 1}$$

$$= -\frac{1}{(\beta - 1)^2}\left(\ln\left(\frac{1 + \beta\left(\frac{\pi_M}{\pi_D} - 1\right)}{\frac{\pi_M}{\pi_D}}\right) - \frac{(\beta - 1)\left(\frac{\pi_M}{\pi_D} - 1\right)}{1 + \beta\left(\frac{\pi_M}{\pi_D} - 1\right)}\right)\widehat{K}.$$

The sign of $\partial \widehat{K}/\partial \beta$ is the opposite of that of the (bracketed) middle term. Applying the logarithm inequality $\ln x > (x - 1)/x$ for $x > 0, x \neq 1$ with $x = (1 + \beta((\pi_M/\pi_D) - 1))/(\pi_M/\pi_D)$ yields

$$\ln\left(\frac{1 + \beta\left(\frac{\pi_M}{\pi_D} - 1\right)}{\frac{\pi_M}{\pi_D}}\right) > \frac{(\beta - 1)\left(\frac{\pi_M}{\pi_D} - 1\right)}{1 + \beta\left(\frac{\pi_M}{\pi_D} - 1\right)}$$

so $\partial \widehat{K}/\partial \beta < 0$ and hence $\partial \widehat{K}/\partial \alpha, \partial \widehat{K}/\partial \sigma > 0$.

A.5 Proposition 5 ($K^W \geq \widetilde{K}$)

If $K < \widetilde{K}$, firms randomize investment triggers over $[Y_M, \infty)$ according to the distribution $G_\alpha(Y)$ and imitator entry is immediate. As discussed in the text, producer surplus is constant over $\left[0, \widetilde{K}\right]$, whereas since $\widetilde{Y}_A$ is stochastically decreasing in K (see Section A.3) consumer surplus from both innovation and imitation is increasing in K over this range.

A.6 Lemma 1 (characterization of Y_P^W)

Suppose $K > \widehat{K}$ so in equilibrium the innovation threshold is $Y_P \in [Y_{NPV}, Y_L)$ and imitation occurs at Y_F. The social welfare function (9) is

$$\begin{aligned} W(K) &= \left(\frac{\pi_M + S_M}{r-\alpha} Y_P - I\right)\left(\frac{y}{Y_P}\right)^\beta + \left(\frac{(2\pi_D + S_D) - (\pi_M + S_M)}{r-\alpha} Y_F - K\right)\left(\frac{y}{Y_F}\right)^\beta \\ &= \left(\frac{\pi_M + S_M}{r-\alpha} Y_P - I\right)\left(\frac{y}{Y_P}\right)^\beta + \left(\beta\left(\frac{S_D - (\pi_M + S_M)}{\pi_D}\right) + \beta + 1\right)\frac{\pi_D y}{\beta(r-\alpha)}\left(\frac{y}{Y_F}\right)^{\beta-1} \end{aligned}$$

and the value $W(\widehat{K})$ results by continuity since $\lim_{K\to\widehat{K}} Y_P = Y_L$.

A preliminary step is to obtain an expression for dY_P/dK which is used subsequently in the computation of $W'(K)$. Recall that Y_P is defined implicitly by the condition $L(Y_P) = F(Y_F)$. Dividing this identity by y^β and grouping the Y_P and Y_F terms yields a more compact form,

$$\begin{aligned} \frac{\pi_M}{r-\alpha}\frac{1}{[Y_P]^{\beta-1}} - \frac{I}{[Y_P]^\beta} &= \frac{\pi_M}{r-\alpha}\frac{1}{[Y_F]^{\beta-1}} - \frac{K}{[Y_F]^\beta} \\ &= \left(\frac{\pi_M}{r-\alpha} - \frac{\beta-1}{\beta}\frac{\pi_D}{r-\alpha}\right)\frac{1}{[Y_F]^{\beta-1}}. \end{aligned}$$

Observe that $[Y_P]^{-\beta} = \left[\left(\frac{\pi_M}{r-\alpha}Y_F - K\right)/\left(\frac{\pi_M}{r-\alpha}Y_P - I\right)\right][Y_F]^{-\beta}$. The above condition has the form $f(Y_P) = g(K)$, so $dY_P/dK = g'(K)/f'(Y_P)$ with

$$\begin{aligned} f'(Y_P) &= -(\beta-1)\frac{\pi_M}{r-\alpha}\frac{1}{[Y_P]^\beta} + \beta\frac{I}{[Y_P]^{\beta+1}} \\ &= (\beta-1)\frac{\pi_M}{r-\alpha}\frac{Y_L - Y_P}{[Y_P]^{\beta+1}} > 0 \end{aligned}$$

and

$$g'(K) = -\frac{\beta-1}{Y_F}\frac{dY_F}{dK} g(K) < 0.$$

Substituting $g(K) = f(Y_P)$ into $g'(K)$ and then developing yields

$$\frac{dY_P}{dK} = -\frac{\beta-1}{K}\frac{f(Y_P)}{f'(Y_P)} = -\frac{r-\alpha}{\pi_M}\frac{\frac{\pi_M}{r-\alpha}Y_P - I}{Y_L - Y_P}\frac{Y_P}{K}.$$

Evaluating $W'(K)$ gives

$$W'(K) = \left(-(\beta-1)\frac{S_M+\pi_M}{r-\alpha}Y_P+\beta I\right)\frac{y^\beta}{[Y_P]^{\beta+1}}\frac{dY_P}{dK} - \left(\beta\left(\frac{S_D-(\pi_M+S_M)}{\pi_D}\right)+\beta+1\right)\left(\frac{y}{Y_F}\right)^\beta.$$

Substitute the expression for dY_P/dK above to get

$$W'(K) = \left(-(\beta-1)\frac{S_M+\pi_M}{r-\alpha}Y_P+\beta I\right)\left(-\frac{r-\alpha}{\pi_M}\frac{\frac{\pi_M}{r-\alpha}Y_P-I}{Y_L-Y_P}\frac{1}{K}\right)\left(\frac{y}{Y_P}\right)^\beta$$
$$-\left(\beta\left(\frac{S_D-(\pi_M+S_M)}{\pi_D}\right)+\beta+1\right)\left(\frac{y}{Y_F}\right)^\beta.$$

Substituting for $[Y_P]^{-\beta}$ in the first term, rearranging, and factoring $(y/Y_F)^\beta$ yields

$$W'(K) = \left((\beta-1)\left(\left(\frac{S_M}{\pi_M}+1\right)Y_P-Y_L\right)\frac{\frac{\pi_M}{r-\alpha}Y_F-K}{Y_L-Y_P}\frac{1}{K} - \left(\beta\left(\frac{S_D-(\pi_M+S_M)}{\pi_D}\right)+\beta+1\right)\right)\left(\frac{y}{Y_F}\right)^\beta$$

$$= \left(\left((\beta-1)\left(\left(\frac{S_M}{\pi_M}+1\right)(Y_P-Y_L)+\frac{S_M}{\pi_M}Y_L\right)\right)\frac{\frac{\beta}{\beta-1}\frac{\pi_M}{\pi_D}-1}{Y_L-Y_P}\right.$$
$$\left.-\left(\beta\left(\frac{S_D-(\pi_M+S_M)}{\pi_D}\right)+\beta+1\right)\right)\left(\frac{y}{Y_F}\right)^\beta$$

$$= \left(-\left(\frac{S_M}{\pi_M}+1\right)\left(\beta\frac{\pi_M}{\pi_D}-(\beta-1)\right)+\frac{S_M}{\pi_M}\left(\beta\frac{\pi_M}{\pi_D}-(\beta-1)\right)\frac{Y_L}{Y_L-Y_P}\right.$$
$$\left.-\left(\beta\left(\frac{S_D-(\pi_M+S_M)}{\pi_D}\right)+\beta+1\right)\right)\left(\frac{y}{Y_F}\right)^\beta.$$

Regrouping the constant (non-Y) terms gives

$$-\left(\frac{S_M}{\pi_M}+1\right)\left(\beta\frac{\pi_M}{\pi_D}-(\beta-1)\right)-\left(\beta\left(\frac{S_D-(\pi_M+S_M)}{\pi_D}\right)+\beta+1\right)=(\beta-1)\frac{S_M}{\pi_M}-\beta\frac{S_D}{\pi_D}-2.$$

Therefore

$$W'(K) = \left(\left(\beta\frac{S_M}{\pi_D}-(\beta-1)\frac{S_M}{\pi_M}\right)\frac{Y_L}{Y_L-Y_P}+(\beta-1)\frac{S_M}{\pi_M}-\beta\frac{S_D}{\pi_D}-2\right)\left(\frac{y}{Y_F}\right)^\beta.$$

Since $\beta(S_M/\pi_D) > (\beta-1)(S_M/\pi_M)$ and $\lim_{K\to\widehat{K}} Y_P = Y_L$, $\lim_{K\to\widehat{K}} W'(K) = +\infty$. Moreover $Y_L/(Y_L-Y_P)$ and y/Y_F are both decreasing functions of K, so W' is decreasing over $\left(\widehat{K},\infty\right)$. Provided that $\lim_{K\to\infty} W'(K) < 0$ therefore, the first-order condition has a unique root in $\left(\widehat{K},\infty\right)$.

Since $\lim_{K\to\infty} Y_P = (r-\alpha) I/\pi_M = ((\beta-1)/\beta) Y_L$, the sign of $\lim_{K\to\infty} W'(K)$ is the same as that of

$$\begin{aligned} & \beta \left(\beta \frac{S_M}{\pi_D} - (\beta-1) \frac{S_M}{\pi_M} \right) + (\beta-1) \frac{S_M}{\pi_M} - \beta \frac{S_D}{\pi_D} - 2 \\ = \; & \beta^2 \frac{S_M}{\pi_D} - (\beta-1)^2 \frac{S_M}{\pi_M} - \beta \frac{S_D}{\pi_D} - 2. \end{aligned}$$

Taken as a function of β, this is a quadratic $\Delta(\beta)$, with $\Delta(1) = (S_M - S_D - 2\pi_D)/\pi_D < 0$ and $\lim_{\beta\to\infty} \Delta(\beta) = \infty$. Therefore there exists a unique $\beta_0 > 1$ such that $\Delta(\beta_0) = 0$. It follows that the constrained optimization problem $\max_{K\geq\widehat{K}} W(K)$ has a unique optimum, which is finite (interior) if $\beta < \beta_0$ and infinite otherwise. Provided that $\beta < \beta_0$, the first-order condition then yields the expression for Y_P^W given in the proposition.

A.7 Lemma 2 (condition for $K^W \geq \widehat{K}$)

The lemma is established by first deriving an upper bound of W over $\left[0, \widehat{K}\right)$ (attrition) and then comparing this bound with the maximum value over $\left[\widehat{K}, \infty\right]$.

Under attrition, expected social welfare (9) can be expressed

$$\begin{aligned} W(K) = \mathbb{E}\left[\left(\frac{S_M + \pi_M}{r-\alpha} \widetilde{Y}_A - I \right) \left(\frac{y}{\widetilde{Y}_A} \right)^\beta \right] & \\ + \mathbb{E}\left[\left(\frac{(S_D + 2\pi_D) - (S_M + \pi_M)}{r-\alpha} \max\left\{ \widetilde{Y}_A, Y_F \right\} - K \right) \left(\frac{y}{\max\left\{ \widetilde{Y}_A, Y_F \right\}} \right)^\beta \right]. \end{aligned}$$

The first integrand is a quasiconcave function of investment threshold with a maximum at $(\beta(r-\alpha) I)/((\beta-1)(S_M+\pi_M)) \leq Y_L \leq \widetilde{Y}_A$, which therefore is decreasing over (Y_L, ∞). Therefore

$$\begin{aligned} \mathbb{E}\left[\left(\frac{S_M + \pi_M}{r-\alpha} \widetilde{Y}_A - I \right) \left(\frac{y}{\widetilde{Y}_A} \right)^\beta \right] & \leq \left(\frac{S_M + \pi_M}{r-\alpha} Y_L - I \right) \left(\frac{y}{Y_L} \right)^\beta \\ & \leq \left(\beta \frac{S_M}{\pi_M} + 1 \right) \frac{I}{\beta-1} \left(\frac{y}{Y_L} \right)^\beta . \end{aligned} \tag{14}$$

The bound on the second summand uses the assumption that the static entry incentive is excessive,

$$\mathbb{E}\left[\left(\frac{(S_D+2\pi_D)-(S_M+\pi_M)}{r-\alpha}\max\left\{\widetilde{Y}_A,Y_F\right\}-K\right)\left(\frac{y}{\max\left\{\widetilde{Y}_A,Y_F\right\}}\right)^{\beta}\right]$$
$$\leq \mathbb{E}\left[\left(\frac{\pi_D}{r-\alpha}\max\left\{\widetilde{Y}_A,Y_F\right\}-K\right)\left(\frac{y}{\max\left\{\widetilde{Y}_A,Y_F\right\}}\right)^{\beta}\right].$$

The term on the right-hand side is the equilibrium expected payoff of a follower, $\mathbb{E}\left[F\left(\max\left\{\widetilde{Y}_A,Y_F\right\}\right)\right]$, that firm i obtains by setting $Y_i=\infty$. Equilibrium payoffs are constant over the support of mixed strategies, so $\mathbb{E}\left[F\left(\max\left\{\widetilde{Y}_A,Y_F\right\}\right)\right]=\max\left\{L\left(Y_L\right),L\left(Y_S\right)\right\}$. This last term is maximized for $K=\widehat{K}$, by Proposition 4. Therefore

$$\mathbb{E}\left[\left(\frac{(S_D+2\pi_D)-(S_M+\pi_M)}{r-\alpha}\max\left\{\widetilde{Y}_A,Y_F\right\}-K\right)\left(\frac{y}{\max\left\{\widetilde{Y}_A,Y_F\right\}}\right)^{\beta}\right]\leq\frac{\widehat{K}}{\beta-1}\left(\frac{y}{\widehat{Y}_F}\right)^{\beta}$$

where $\widehat{Y}_F:=\left(\beta\left(r-\alpha\right)\widehat{K}\right)/((\beta-1)\pi_D))=\left(\widehat{K}/I\right)\left(\pi_M/\pi_D\right)Y_L$. Using (11) to substitute for $\left(\widehat{K}/I\right)^{-(\beta-1)}$ gives

$$\begin{aligned}\mathbb{E}\left[\left(\frac{(S_D+2\pi_D)-(S_M+\pi_M)}{r-\alpha}\widetilde{Y}_F-K\right)\left(\frac{y}{\widetilde{Y}_F}\right)^{\beta}\right] &\leq \frac{1}{1+\beta\left(\frac{\pi_M}{\pi_D}-1\right)}\frac{I}{\beta-1}\left(\frac{y}{Y_L}\right)^{\beta} \quad (15)\\ &< \frac{0.5I}{\beta-1}\left(\frac{y}{Y_L}\right)^{\beta}. \quad (16)\end{aligned}$$

Combining (14) and (15) yields

$$\max_{K\in\left[0,\widehat{K}\right)}W\left(K\right)<\left(\beta\frac{S_M}{\pi_M}+1.5\right)\frac{I}{\beta-1}\left(\frac{y}{Y_L}\right)^{\beta}$$

(we seek an analytically tractable bound rather than the tightest bound here).

Over $\left[\widehat{K},\infty\right]$, social welfare can be evaluated exactly, yielding[25]

$$\max_{K\in\left[\widehat{K},\infty\right]}W\left(K\right)=\begin{cases}\frac{S_M}{\pi_M}\frac{(1+\psi)^{\beta}}{\psi}\frac{I}{\beta-1}\left(\frac{y}{Y_L}\right)^{\beta}, & \beta<\beta_0\\ \left(\frac{\beta}{\beta-1}\right)^{\beta}\frac{S_M}{\pi_M}I\left(\frac{y}{Y_L}\right)^{\beta}, & \beta\geq\beta_0\end{cases} \quad (17)$$

[25] Derivation is available from the authors.

where ψ is given in Lemma 1. By revealed preference, the optimal welfare level is at least as large as if the regulator sets an infinite cost of imitation, so

$$\max_{K\in\left[\widehat{K},\infty\right]} W\left(K\right) \geq \left(\frac{\beta}{\beta-1}\right)^{\beta}\frac{S_M}{\pi_M}I\left(\frac{y}{Y_L}\right)^{\beta}.$$

A sufficient condition for $\max_{K\in\left[0,\widehat{K}\right)} W\left(K\right) < \max_{K\in\left[\widehat{K},\infty\right]} W\left(K\right)$ is therefore

$$\left(\frac{\beta}{\beta-1}\right)^{\beta}\frac{S_M}{\pi_M}I\left(\frac{y}{Y_L}\right)^{\beta} > \left(\beta\frac{S_M}{\pi_M}+1.5\right)\frac{I}{\beta-1}\left(\frac{y}{Y_L}\right)^{\beta}.$$

Cancelling common terms and rearranging yields

$$\left(\left(\frac{\beta}{\beta-1}\right)^{\beta-1}-1\right)\frac{S_M}{\pi_M} > \frac{3}{2\beta}$$

which establishes the condition in the text.

A.8 Proposition 7 (specific welfare cases)

For part (i), we consider the limiting case $S_M = 0$ (assuming that $S_D > 0$) and then use the continuity of $W(K)$. From the expression of $W(K)$ (9), only producer surplus and the consumer surplus from imitation matter in this case. Both of these are decreasing in K over $\left(\widehat{K},\infty\right]$ so any maximum of K must lie in $\left(\widetilde{K},\widehat{K}\right]$ (the lower bound is the one given by Proposition 5). Moreover, since producer surplus is maximized at $\widehat{K}$ (Proposition 4), the sign of the left derivative of welfare at $\widehat{K}$, $\lim_{K\to\widehat{K}^-} W'(K)$, depends only on the consumer surplus from imitation. Given the innovation threshold under attrition $\widetilde{Y}_A = \min\left\{\widetilde{Y}_1,\widetilde{Y}_2\right\}$, this term is

$$\begin{aligned}
\frac{\left(S_D-S_M\right)y^{\beta}}{r-\alpha}\mathbb{E}\frac{1}{\left[\max\left\{\widetilde{Y}_A,Y_F\right\}\right]^{(\beta-1)}} &= \frac{\left(S_D-S_M\right)y^{\beta}}{r-\alpha}\left(\frac{1}{\left[Y_F\right]^{(\beta-1)}}\Pr\left\{\widetilde{Y}_A\leq Y_F\right\}+\mathbb{E}_{\widetilde{Y}_A>Y_F}\frac{1}{\left[\widetilde{Y}_A\right]^{(\beta-1)}}\right)\\
&= \frac{\left(S_D-S_M\right)y^{\beta}}{\left(r-\alpha\right)\left[Y_F\right]^{(\beta-1)}}\left(G_A\left(\overline{Y}_L\right)+\int_{Y_M}^{\infty}\left(\frac{Y_F}{s}\right)^{\beta-1}dG_A(s)\right)
\end{aligned}$$

where $G_A = 1-\left(1-G_\beta\left(Y\right)\right)^2$ denotes the equilibrium distribution of $\widetilde{Y}_A$. Since $G_\beta\left(\overline{Y}_L\right)\big|_{K=\widehat{K}} = 1$, $G_A\left(\overline{Y}_L\right)\big|_{K=\widehat{K}} = 1$ and the right hand term vanishes at $\widehat{K}$ (note that $dG_A/dY = 2\left(1-G_\beta\right)\left(dG_\beta/dY\right)$). Therefore at $\widehat{K}$ only the direct effect $\partial Y_F/\partial K$ remains, hence

$$\lim_{K\to\widehat{K}^-} W'(K) = -(\beta-1)\frac{S_D-S_M}{r-\alpha}\left(\frac{y}{Y_F}\right)^{\beta}\frac{\partial Y_F}{\partial K}\leq 0$$

with strict inequality if $S_D > S_M$. For $S_M = 0$ therefore, $\lim_{K \to \widehat{K}^-} W'(K) < 0$ so the maximum of W lies in $\left(\widetilde{K}, \widehat{K}\right)$.

The argument for part (ii) is straightforward. From the expression of $W(K)$ (9), only producer surplus and the consumer surplus from innovation matter in this case. The first of these is weakly increasing in K over $\left[0, \widehat{K}\right]$ and the second is increasing over $\mathbb{R}_+$, so the maximum of $W(K)$ lies in $\left(\widehat{K}, \infty\right]$.

A.9 Proposition 8 (predetermined sequence vs. endogenous innovation welfare)

The social welfare assuming predetermined firm roles that results from setting the optimal imitation cost $K^W = 0$ is (see Section 3)

$$W^{\text{exog}} = \left(\frac{\beta}{\beta - 1} \frac{S_D + 2\pi_D}{\pi_M} - 1\right) I \left(\frac{y}{Y_L}\right)^{\beta}.$$

Proposition 6 identifies the optimal welfare achieved assuming endogenous innovation (provided i) the static entry incentive is excessive and ii) $S_M/\pi_M \geq \Omega(\beta)$) as being bounded below by

$$W^{\text{endog}} = \left(\frac{\beta}{\beta - 1}\right)^{\beta} \frac{S_M}{\pi_M} I \left(\frac{y}{Y_L}\right)^{\beta}.$$

Using condition i), *i.e.* $(S_D + 2\pi_D) - (S_M + \pi_D) < \pi_D$,

$$W^{\text{exog}} \leq \left(\frac{\beta}{\beta - 1}\left(\frac{S_M + \pi_D}{\pi_M} + 1\right) - 1\right) I \left(\frac{y}{Y_L}\right)^{\beta}.$$

A sufficient condition for welfare to be greater with endogenous innovation is therefore

$$\left(\frac{\beta}{\beta - 1}\right)^{\beta} \frac{S_M}{\pi_M} \geq \frac{\beta}{\beta - 1}\left(\frac{S_M + \pi_D}{\pi_M} + 1\right) - 1.$$

Regrouping terms and dividing by $\beta/(\beta - 1)$ gives

$$\left(\left(\frac{\beta}{\beta - 1}\right)^{\beta - 1} - 1\right) \frac{S_M}{\pi_M} \geq \frac{\pi_D}{\pi_M} + \frac{1}{\beta}.$$

Using the second condition to substitute a lower bound for S_M/π_M gives the sufficient condition,

$$\frac{3}{2\beta} \geq \frac{\pi_D}{\pi_M} + \frac{1}{\beta}$$

so optimal welfare is higher with endogenous innovation if $\pi_M/\pi_D \geq \beta$, *i.e.* if the degree of product market competition is sufficiently large.

A.10 Proposition 10 (takeovers and licensing)

Before establishing the proposition we verify the claim made in the text, *i.e.* that attrition does not occur if $\pi_M/\pi_D \geq \beta + 1$. Consider the limiting case $K = 0$. The leader payoff $L_{\mathrm{T}}(Y)$ is maximized at the threshold $Y_T = (\beta(r-\alpha)I)/((\beta-1)(\pi_M - \pi_D))$. The investment game is (weakly) preemptive if $L_{\mathrm{T}}(Y_T) \geq F(Y_T)$, that is if

$$\left(\frac{\pi_M - \pi_D}{r-\alpha}Y_T - I\right)\left(\frac{y}{Y_T}\right)^{\beta} \geq \frac{\pi_D}{r-\alpha}Y_T\left(\frac{y}{Y_T}\right)^{\beta}$$

which yields the desired condition on π_M/π_D.

To establish the proposition we first verify that a takeover is the preferred instrument. The condition $L_{\mathrm{T}}(Y) \geq L_{\mathrm{L}}(Y)$ works out to

$$\left(\frac{\pi_M - \pi_D}{r-\alpha}\max\{Y, Y_F\} - K_I\right)\left(\frac{y}{\max\{Y, Y_F\}}\right)^{\beta} \geq \left(\frac{\pi_D}{r-\alpha}\max\{Y, Y_F\} - \overline{K} - K_I\right)\left(\frac{y}{\max\{Y, Y_F\}}\right)^{\beta}$$

which holds because of the efficiency effect $\pi_M > 2\pi_D$.

That takeovers increase firm profit for $K < \widehat{K}$ follows from $L_{\mathrm{T}}(Y) > L(Y)$ and the rent dissipation property of attrition and preemption.

Similarly, licensing (provided $K_I > 0$) increases firm profit while leaving the timing of imitation unchanged. It therefore remains to verify that licensing results in earlier innovation. Let $\widehat{K}_L < \widehat{K}$ denote the critical threshold that separates attrition and preemption in the presence of licensing, which solves $L_{\mathrm{L}}(Y_L) = F(Y_F)$ (as licensing only has a level effect on the leader payoff for $Y < Y_F$ the payoff L_{L} is maximized at Y_L). For $K \geq \widehat{K}_L$, as $L_{\mathrm{L}}(Y) > L(Y)$ allowing licensing results in innovation at a threshold that is either lower than the preemption threshold without licensing or weakly lower than the previous possible innovation thresholds. Otherwise if $K < \widehat{K}_L$, the industry is in an attrition regime both with and without licensing and the distribution of innovation thresholds shifts left with licensing.

B Dynamic representation of the investment game

To represent the investment game whose normal form is studied in Section 4 in continuous time, assume that the feasible investment strategies of firms are first-hitting times $\tau(Y_i) := \inf\{t \geq 0 \,| Y_t \geq Y_i\}$. This applies for instance if managerial decisions consist of hurdle rates $((Y_i\pi_M/(r-\alpha))/I) - 1$ for innovative investment. Then the distributions of investment times are ordered by the investment thresholds Y_i and the investment game is isomorphic that in Fudenberg and Tirole [10]. Their analysis applies verbatim, by defining extended mixed strategies over investment thresholds instead of time.[26]

B.1 Strategies and payoffs

In the dynamic representation of the investment game the continuation payoffs depend on the current state of the stochastic process $y \in \mathbb{R}_+$ and are accordingly denoted $L^y(Y)$, $F^y(Y)$ and $M^y(Y)$.

An *extended* mixed strategy for player $i \in \{1,2\}$ in state y is a pair of real-valued functions $(G_i^y, \alpha_i^y) : [y,\infty) \times [y,\infty) \to [0,1] \times [0,1]$ such that (a) G_i^y is non-decreasing and right-continuous, (b) $\alpha_i^y(Y) > 0 \Rightarrow G_i^y(Y) = 1$, (c) α_i^y is right-differentiable and (d) if $\alpha_i^y(Y) = 0$ and $Y = \inf\{Z \geq Y, \alpha_i^y(Z) > 0\}$ then α_i^y has positive right-derivative at Y.

Let $G_i^{y-}(Y) := \lim_{Z \to Y^-} G_i^y(Z)$ denote the left-hand limit of $G_i^y(Y)$, $a_i^y(Y) = G_i^y(Y) - G_i^{y-}(Y)$ the magnitude of any jump at Y and set $G_i^{y-}(y) = 0$, $i = 1,2$. Let $\mathcal{Y}_i(y) = \infty$ if $\alpha_i^y(Y) = 0$ for all $Y \geq y$ and $\mathcal{Y}_i(y) = \inf\{Z \geq y, \alpha_i^y(Z) > 0\}$ otherwise, so $\mathcal{Y}(y) = \min\{\mathcal{Y}_1(y), \mathcal{Y}_2(y)\}$ denotes the first threshold at which an investment is certain to occur. Finally let

$$\mu_L(u,v) := \frac{u(1-v)}{u+v-uv} \text{ and } \mu_M(u,v) := \frac{uv}{u+v-uv}.$$

Firm payoffs are

$$\begin{aligned} &V^y\left((G_i^y, \alpha_i^y), \left(G_j^y, \alpha_j^y\right)\right) = \\ &\left[\int_y^{\max\{\mathcal{Y}(y)^-, y\}} \left(L^y(s)\left(1 - G_j^y(s)\right) dG_i^y(s) + F^y(s)\left(1 - G_i^y(s)\right) dG_j^y(s)\right) + \sum_{Z<\mathcal{Y}(y)} a_i^y(Z)\, a_j^y(Z)\, M^y(Z)\right] \\ &\qquad + \left(1 - G_i^{y-}(\mathcal{Y}(y))\right)\left(1 - G_j^{y-}(\mathcal{Y}(y))\right) W^{\mathcal{Y}(y)}\left((G_i^y, \alpha_i^y), \left(G_j^y, \alpha_j^y\right)\right), \end{aligned}$$

$i, j \in \{1,2\}$, $i \neq j$ where

$$W^Y\left((G_i^y, \alpha_i^y), \left(G_j^y, \alpha_j^y\right)\right) = \frac{a_j^{y-}(Y)}{1 - G_j^{y-}(Y)}\left((1 - \alpha_i^y(Y))\, F^y(Y) + \alpha_i^y(Y)\, M^y(Y)\right) + \frac{1 - G_j^y(Y)}{1 - G_j^{y-}(Y)} L^y(Y)$$

[26] Steg and Thijssen [24] study an investment game with closed-loop stopping times strategies and obtain similar equilibrium outcomes. Their framework accounts for the process exiting the attrition region, whereas with first-hitting time strategies firms remain within the attrition region once it has been attained even if the value of the process subsequently exits.

if $\mathcal{Y}_i(y) < \mathcal{Y}_j(y)$,

$$= \frac{a_i^{y-}(Y)}{1 - G_i^{y-}(Y)} \left(\left(1 - \alpha_j^y(Y)\right) L^y(Y) + \alpha_j^y(Y) M^y(Y) \right) + \frac{1 - G_i^y(Y)}{1 - G_i^{y-}(Y)} F^y(Y)$$

if $\mathcal{Y}_i(y) > \mathcal{Y}_j(y)$ and

$$= \begin{cases} M^y(Y), & a_i^y(Y) = a_j^y(Y) = 1 \\ \mu_L(a_i^y(Y), a_j^y(Y)) L^y(Y) + \mu_L^y(a_j^y(Y), a_i^y(Y)) F^y(Y) & 0 < a_i^y(Y) + a_j^y(Y) < 2 \\ +\mu_M^y(a_i^y(Y), a_j^y(Y)) M^y(Y), & \\ \frac{\left(a_i^y(Y)\right)' L^y(Y) + \left(a_j^y(Y)\right)' F^y(Y)}{\left(a_i^y(Y)\right)' + \left(a_j^y(Y)\right)'}, & a_i^y(Y) + a_j^y(Y) = 0 \end{cases}$$

if $\mathcal{Y}_i(y) = \mathcal{Y}_j(y)$.

For given y, a pair of simple strategies $((G_1^y, \alpha_1^y), (G_2^y, \alpha_2^y))$ is a Nash equilibrium if (G_i^y, α_i^y) maximizes $V^y\left((G_i^y, \alpha_i^y), \left(G_j^y, \alpha_j^y\right)\right)$, $i, j \in \{1, 2\}$, $i \neq j$. A collection of simple strategies $((G_i^y(Y), \alpha_i^y(Y)))_{y>0}$ is consistent if for $y \leq Y \leq Z$, $G_i^y(Z) = G_i^y(Y) + (1 - G_i^y(Y))\, G_i^Y(Z)$ and $\alpha_i^y(Z) = \alpha_i^Y(Z)$. The consistent strategies $((G_1^y(Y), \alpha_1^y(Y)))_{y>0}$ and $((G_2^y(Y), \alpha_2^y(Y)))_{y>0}$ are a perfect equilibrium if the simple strategies $(G_1^y(Y), \alpha_1^y(Y))$ and $(G_2^y(Y), \alpha_2^y(Y))$ are a Nash equilibrium for every y.

B.2 Equilibrium

In the closure of the attrition range ($K \leq \widehat{K}$) firms do not resort to extended mixed strategies. Equilibrium strategies are therefore obtained from the unconditional strategies $G_\alpha(Y)$ or $G_\beta(Y)$ (see Section A.3) according to whether $K \leq \widehat{K}$ or $\widetilde{K} < K < \widehat{K}$. Therefore, letting $G_\alpha^y(Y) := \frac{G_\alpha(Y) - G_\alpha(y)}{1 - G_\alpha(y)}$ and $G_\beta^y(Y) := \frac{G_\beta(Y) - G_\beta(y)}{1 - G_\beta(y)}$, $(G_i^y(Y), \alpha_i^y(Y)) = (G_\alpha^y(Y), 0)$ and $(G_i^y(Y), \alpha_i^y(Y)) = \left(G_\beta^y(Y), 0\right)$ are symmetric subgame perfect equilibrium strategies in these two subcases.

In the preemption range ($K > \widehat{K}$) the firms do resort to extended mixed strategies and there are two subcases that we consider successively.

B.2.1 $\widehat{K} < K < I$

This is the case represented in Figure 3 whose key features are that the preemption range (over which $L^y(Y) > F^y(Y)$) is the bounded interval $\left(Y_P, \overline{Y}_P\right) \subset (Y_P, Y_F)$, and that if a threshold beyond this range is reached, firms play a waiting game as $F^y(Y) > L^y(Y)$ for $Y > \overline{Y}_P$. In a dynamic representation of the game, subgame perfect equilibrium strategies must account for this possibility.

At any $y > \overline{Y}_P$ the payoff to leading lies below the follower payoff. It is not monotonic however, and there exists a unique threshold $\underline{Y}_L \in \left(\overline{Y}_P, Y_F\right)$ such that $L^y\left(\underline{Y}_L\right) = L^y\left(Y_M\right)$. The leader payoff is decreasing only over $\left(\overline{Y}_P, \underline{Y}_L\right) \cup (Y_M, \infty)$, and it is this range that constitutes the support of mixed strategies. The attrition subgame is then solved similarly to the $\widetilde{K} < K < \widehat{K}$ case in Section A.3 yielding unconditional distributions

$$G_\gamma(Y) = 1 - \exp \int_{\overline{Y}_P}^{Y} \frac{[L^y(s)]'}{F\left(\max\left\{Y, Y_F\right\}\right) - L^y(s)} ds$$

and

$$G_\delta(Y) = \begin{cases} 0, & Y < \overline{Y}_P \\ G_\gamma(Y), & \overline{Y}_P \leq Y \leq \underline{Y}_L \\ G_\gamma\left(\underline{Y}_L\right), & \underline{Y}_L < Y < Y_M \\ G_\gamma\left(\underline{Y}_L\right) + \left(1 - G_\gamma\left(\underline{Y}_L\right)\right) G_\alpha(Y), & Y \geq Y_M \end{cases}$$

so that the conditional distribution is $G_\delta^y(Y) := \frac{G_\delta(Y) - G_\delta(y)}{1 - G_\delta(y)}$.

If y lies in the preemption range, the reasoning is standard and results in firms investing immediately and using the strategy extensions to coordinate simultaneous investment.

Therefore, the symmetric equilibrium strategies are $(G_i^y(Y), \alpha_i^y(Y))$ with

$$\begin{aligned} G_i^y(Y) &= \begin{cases} 0, & Y < Y_P \\ 1, & Y_P \leq Y < \overline{Y}_P \\ G_\delta^y(Y), & Y \geq \overline{Y}_P \end{cases}, \\ \alpha_i^y(Y) &= \begin{cases} 0, & Y < Y_P \\ \frac{L^y(Y) - F^y(Y)}{L^y(Y) - M^y(Y)}, & Y_P \leq Y < \overline{Y}_P \\ 0, & Y \geq \overline{Y}_P \end{cases} \end{aligned}$$

for $i \in \{1, 2\}$.

B.2.2 $\widehat{K} \geq I$

Here $L^y(Y) > F^y(Y)$ over (Y_P, Y_F) so the investment game is a standard preemption game. A specificity of the investment game studied here is that for $K > I$, M^y lies strictly above F^y over $[Y_F, \infty)$. Symmetric equilibrium strategies are nevertheless those of a standard real option game, yielding $(G_i^y(Y), \alpha_i^y(Y))$ with

$$
\begin{aligned}
G_i^y(Y) &= \begin{cases} 0, & Y < Y_P \\ 1, & Y \geq Y_P \end{cases}, \\
\alpha_i^y(Y) &= \begin{cases} 0, & Y < Y_P \\ \frac{L^y(Y)-F^y(Y)}{L^y(Y)-M^y(Y)}, & Y_P \leq Y < Y_F \\ 1, & Y \geq Y_F \end{cases}
\end{aligned}
$$

for $i \in \{1, 2\}$.

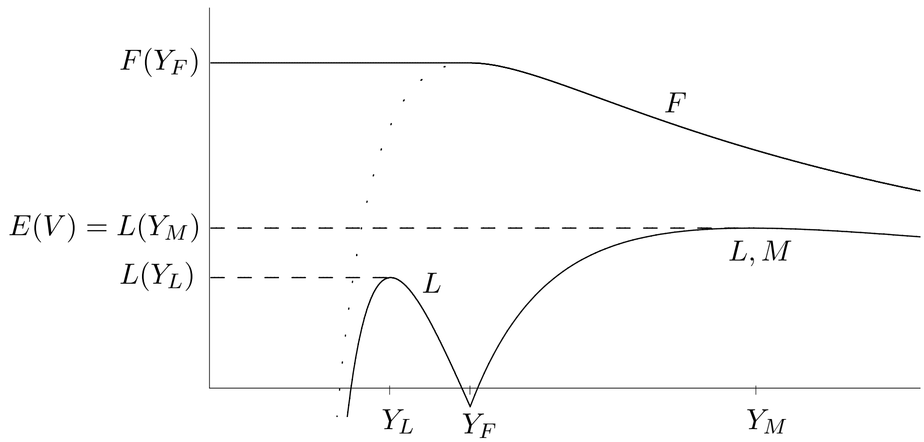

Figure 1 Attrition ($K < \widetilde{K}$): Y_M is a global maximum of the leader payoff, innovation thresholds are distributed over $[Y_M, \infty)$, and imitation occurs immediately after, for an equilibrium value $E(V) = L(Y_M)$.

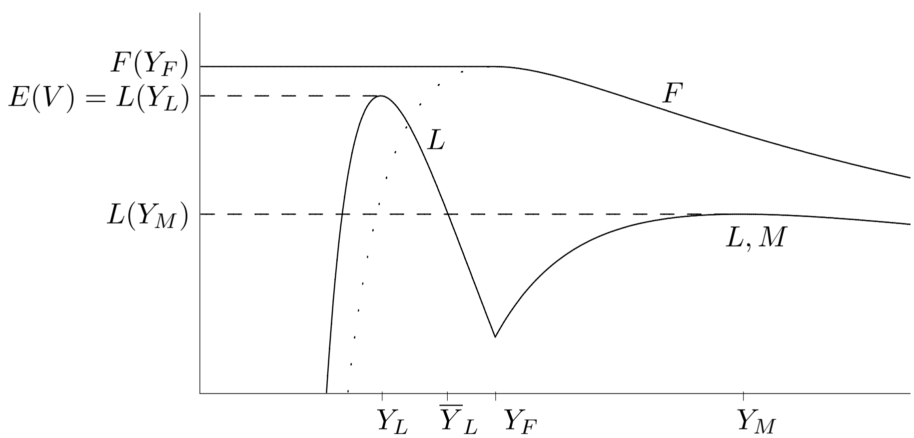

Figure 2 Attrition ($\widetilde{K} \leq K < \widehat{K}$): Y_L is a global maximum of the leader payoff, innovation thresholds are distributed over $\left[Y_L, \overline{Y}_L\right] \cup [Y_M, \infty)$, imitation occurs either at Y_F if the innovation threshold is in $\left[Y_L, \overline{Y}_L\right]$, or immediately otherwise, for an equilibrium value $E(V) = L(Y_L)$.

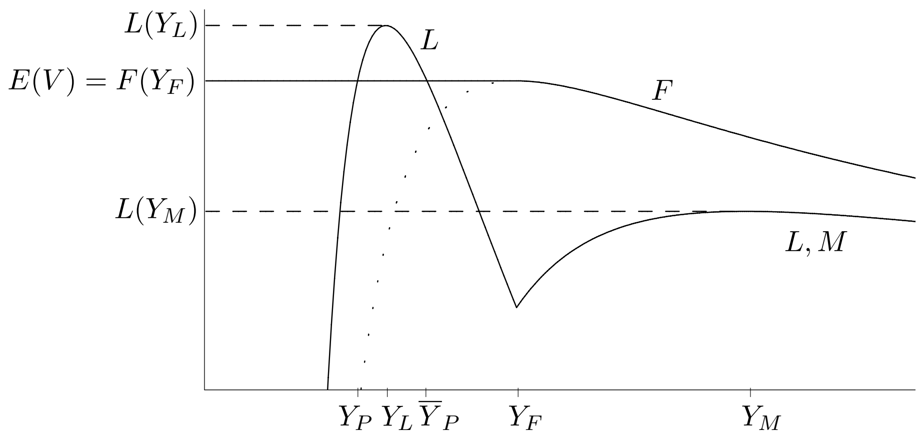

Figure 3 Preemption ($\widehat{K} \leq K < I$): innovation occurs at Y_P and imitation at Y_F, for an equilibrium value $E(V) = F(Y_F)$. There is attrition off the equilibrium path (over $(\overline{Y}_P, \infty)$).

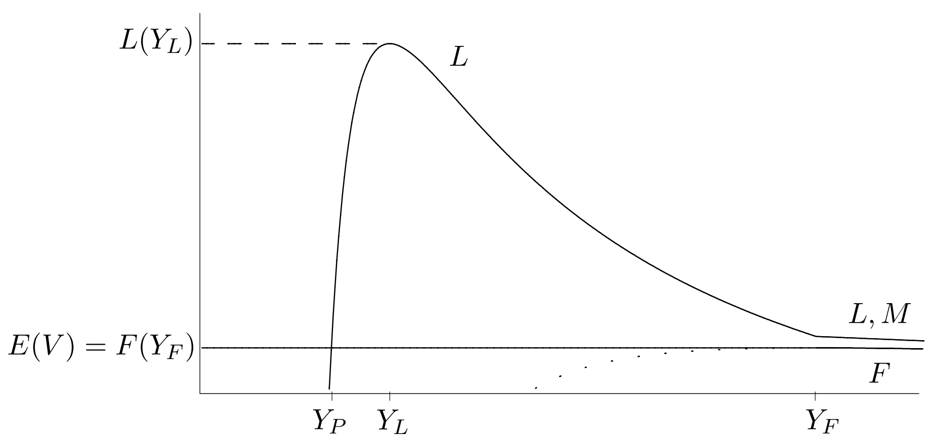

Figure 4 Standard preemption ($I \leq K$): innovation occurs at Y_P and imitation at Y_F, for an equilibrium value $E(V) = F(Y_F)$ (for clarity $M(Y)$ is graphed only on $[Y_F, \infty)$).