# When should a winner take all, or pay some? Innovation and imitation incentives in a dynamic duopoly

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

December 2016

# When should a winner take all, or pay some? Innovation and imitation incentives in a dynamic duopoly

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

December 6, 2016

**Abstract**

We develop a model of investment in duopoly with asymmetric costs of innovating and imitating and endogenous firm roles. Dynamic competition involves either attrition or preemption, the former being likelier with high demand growth and uncertainty. Industry value is maximized when firms neither stall nor hasten entry, and we show that social welfare has local maxima in both the attrition and preemption ranges. In all cases the socially optimal cost of imitation is positive. Attrition is optimal if consumer surplus rises sufficiently under duopoly, whereas with static business-stealing, preemption is optimal if discounting is important enough. Finally we discuss endogenous entry barriers and contracting, finding that firms are more likely to rely on secrecy and patents at low imitation costs and that simple licensing schemes are welfare improving.

JEL Classification: G31 (Capital Budgeting; Fixed Investment and Inventory Studies; Capacity), L13 (Oligopoly and Other Imperfect Markets), O33 (Technological Change: Choices and Consequences; Diffusion Processes)

Keywords: Dynamic oligopoly; Knowledge spillover; Real options

*Billette de Villemeur: LEM, Université de Lille, F-59655, France. Ruble and Versaevel: EM-LYON 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 Alexei Alexandrov, 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, IÉSEG business school, LMU and TU München, and conference participants at the 2014 EARIE (Milan), 2015 International Conference on Real Options (Athens and Monemvasia) 2016 AFSE (Nancy) and Ninth Annual Conference on Innovation Economics (Chicago) for valuable suggestions and help in various forms. All remaining errors are ours.

# 1 Introduction

Profitable investment by a pioneering firm naturally breeds imitation by a subsequent entrant. Developing an invention into a commercial product can require significant enough resources so that only a couple of firms may jockey to secure positions in an industry either as a first-mover or as a second entrant. In such circumstances product development takes the form of a noncooperative timing game in which the nature of competition and equilibrium outcomes are driven by the relative costs of innovation and possible imitation.

We study the effect of differences in the costs of innovating and imitating on strategic investment decisions by developing a theoretical framework in which the roles of firms, as innovators or imitators, are endogenous. How does such an *ex-post* asymmetry in investment costs, together with the endogeneity of roles, affect the initial decisions of *ex-ante* identical firms? At which (relative) cost of imitation is industry value maximized? From the standpoint of policy, how can intellectual property (IP) rights modify the nature of competition and maximize social welfare in the kinds of industries that we examine? Our paper sheds light on these questions and highlights the role played by such industry characteristics as demand growth and uncertainty.

One key set of conclusions that emerges is that with low demand growth and volatility, as may arise in mature industries, a high degree of IP protection leading to preemption is optimal. Moreover, in such industries it can be preferable to afford complete protection to innovators, so that strategic investment has the form of a *winner-take-all* contest.

Conversely, attrition is more likely to arise, and low IP protection can constitute a social optimum, when demand growth and volatility are high, a case that corresponds precisely to the circumstances most commonly associated with innovative industries. However costless imitation by a firm that succeeds in investing second is never socially optimal, so an optimal attrition regime is necessarily *winner-pay-some* with a lower bound on the cost of imitation that we are able to characterize.

## 1.1 Results

Our model leads to three sets of results. First, we characterize the effect of varying imitation cost on strategic competition. A low imitation cost leads to delayed product introduction as firms seek to enter second, a situation of attrition. Conversely a higher imitation cost is associated with accelerated product introduction, a case of preemption. Equilibrium in firm investments involves mostly standard preemption and attrition although the latter can also present a gap in the support of mixed strategies. Preemption is more likely when product market competition is more intense whereas attrition is more likely when discounting is less important, as occurs when demand growth and volatility are high. Under both attrition and preemption, positional rents are dissipated and we are able to identify the optimal relative imitation cost for the industry, which is that cost of imitation at which there is neither a race to preempt nor a war of attrition, *i.e.* at which firms do not compete for positional rents by rushing or waiting unduly to innovate.

Next, we examine the welfare trade-offs associated with raising imitation cost, as may arise in the context of a regulator's choice of broader patent protection. There is a positive lower bound for the optimal level of imitation cost, implying that free imitation is always socially costly. The social welfare function generally has two local maxima in the attrition and preemption range, either of which may constitute a global maximum. We identify different conditions on product market competition or demand primitives for one form of dynamic competition or another to be optimal. Attrition is optimal if a monopoly innovator practices first-degree price discrimination or with a sufficiently elastic demand specification, and preemption is optimal if there is product market collusion under duopoly. If product market competition is characterized by a business-stealing effect, we derive a sufficient condition for preemption to be socially optimal and show that with sufficient demand growth or volatility, the optimum is a corner solution that results in a monopolized industry.

Finally, we discuss the extension of our model in a number of directions by incorporating a broader set of firm decisions. First we endogenize the cost of imitation by allowing the innovator to make reverse engineering of its product more difficult or to pursue patent protection more aggressively. We find that the lower the natural cost of imitation, the greater the effort exerted by innovators to raise entry barriers. In addition, we allow for contracting between innovator and imitator that can take the form either of a buyout or of a license agreement. With the former, attrition may disappear entirely as an equilibrium if discounting is sufficiently large. With the latter, licensing increases welfare if the efficiency effect is present, whereas if there is sufficient product market complementarity the innovator may choose to privately subsidize imitation.

## 1.2 Related literature

Our model of innovation and imitation builds upon a rich literature dating back to Reinganum [29] who provides a foundation for dynamic game-theoretic models of duopoly adoption of a new technology. In a deterministic environment in which one of the firms can commit as a first investor, she identifies a diffusion equilibrium in which investments occur sequentially and result in a first-mover advantage. Fudenberg and Tirole [12] study investment decisions when leader and follower roles are endogenous. In a preemption race, the first investment occurs earlier than under diffusion, dissipating rents to the first investor so that firm values are equalized in equilibrium. In a similar deterministic framework but with asymmetric firms, Katz and Shapiro [23] allow for post-investment licensing or imitation and find that a second-mover advantage can arise so that investment decisions take the form either of a preemption race or of a waiting game, although their focus is on asymmetric equilibria of the attrition game. Subsequently, Hoppe [21] introduces uncertainty regarding the success of new technology adoption into a similar model which leads to the possibility of both preemption and attrition, although her focus is also on asymmetric equilibrium in pure strategies rather than the type of symmetric mixed strategy equilibrium that we characterize here.

Early research on the relative incentives of innovators and imitators addressed the issues of optimal patent length and breadth, as in Gilbert and Shapiro [15] and Gallini [14], whose notation we follow for the cost of inventing around. Denicolò [8] discusses optimal patents in an innovation race that bears some similarities with the model presented here, although his specification does not allow for second-mover advantage and attrition as we do, and we do not seek to characterize patent duration but rather the optimal degree of IP protection. Our focus is on industries in which firms are horizontal competitors rather than on the distribution of rents between basic and applied research, whereas imitation has often been examined in the context of cumulative innovation, notably by Green and Scotchmer [17] and Denicolò [9]. Some more recent work dealing with innovation and imitation such as Mukherjee and Pennings [28] and Henry and Ruiz-Aliseda [20] identifies the importance of the patenting, licensing, and reverse engineering decisions that we examine, but do so after innovation occurs with models in which one of the firms is an incumbent so that their focus is on inherently asymmetric firms.

Because it takes demand to follow a stochastic dynamic process, real options theory seems to us to provide the right framework within which to cast a discussion of innovation incentives. A pioneering application of this type of strategic investment model[1] to patent races is that of Weeds [37], although she more closely describes

[1] Azevedo and Paxson [1] is a recent survey of this field, which draws from game theory and

the invention stage of innovation than the development or product introduction stage that constitute our focus. We depart from existing work on strategic investment by parsimoniously parametrizing first- and second-mover advantage through the relative fixed costs of innovating and imitating firms and by characterizing a symmetric equilibrium in Markovian strategies that lends itself to a novel welfare analysis. Huisman [22] is a complementary contribution which studies the effect of *ex-post* (rather than *ex-ante*) firm asymmetry in a duopoly investment game. Research on innovation dynamics within this literature has addressed on informational spillovers, which are one of the important determinants of second-mover advantage. For instance, Femminis and Martini [11] model a disclosure lag of random duration before the follower benefits from a spillover. The effect of informational spillovers on investment incentives has also been studied in models of learning by Décamps and Mariotti [7] and Thijssen et al. [34]. Through these different contributions runs a common thread which also forms a basis for our work: to the extent an innovator's investment has positive spillovers for its rival, competition between otherwise symmetric firms takes the form either of a preemption race or of a war of attrition. Thus an economic model of these phenomena should in principle account for both types of cases.

### 1.3 An example: imitation cost in the biopharmaceutical industry

The questions we address were originally motivated by real-world situations in which the same firms can face contrasting technological conditions with respect to ease of imitation over the different business segments in which they operate. In the biopharmaceutical industry, typically, whereas medications are easily imitated thus justifying the industry's systematic recourse to patent protection, in the vaccine segment technological conditions render imitation much more costly.[2]

On the one hand, pharmaceutical firms typically 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

continuous time finance in order to incorporate strategic and payoff uncertainty into models of investment. Typical applications are to capacity investment, as in Boyer et al. [4], as well as investment in R&D, as in the present paper. For a thorough and pedagogical presentation, see also Chevalier-Roignant and Trigeorgis [5].

[2]Another characteristic of the pharmaceutical industry is the uncertainty that is introduced by late-stage clinical trials regarding the outcome of an R&D project, most often after significant costs have already been sunk, but we do not seek to represent this specific feature in our model.

higher" (Scherer and Watal [31], 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. In India, after the passage of the Patents Act 1970, and before the TRIPs (Trade Related aspects of Intellectual Property rights) agreements were enforced, pharmaceutical products became unpatentable, "allowing innovations patented elsewhere to be freely copied" (Lanjouw [24], p. 3). By reducing imitation costs, the absence of legal protection fostered the domestic production of generic formulations.

This ease of imitation is not found in the vaccine segment, as vaccines 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 [36], p. 13). The regulatory implication is that a me-too vaccine supplier must pay for clinical trials to demonstrate the safety and efficacy of its product. There is no short-cut toward the bio-equivalence of a copied candidate vaccine, whose design and delivery require investments in technological capabilities and manufacturing facilities that comply with demanding regulatory standards. In the case of recent complex vaccines (*e.g.*, a tetravalent dengue virus vaccine), a follower must catch up with leading-edge R&D and manufacturing approaches (the technological challenges for the design a dengue virus vaccine are reviewed in Guey Chuen et al. [18]). The fixed cost that must be incurred by a new entrant for the delivery of a follow-on vaccine can thus be prohibitively high.[3]

## 2 A model of new product development

This section describes a model of strategic investment in line with the characteristic features of innovation and imitation identified above. Assumptions regarding industry structure and firm conduct are presented in Sections 2.1 and 2.2 and equilibrium is characterized in Section 2.3.

### 2.1 Assumptions

Two identical firms seek to enter a market by introducing their version of a novel product. Organizational constraints preclude a firm from selling two variants of the product and technological or regulatory barriers shield both firms from further entry.

[3] We return to this example in light of the theoretical model in Section 4.1.

Development of the new product occurs in the face of uncertainty regarding future demand levels and involves irreversible investment as described by Dixit and Pindyck [10].

The introduction of the product generates a baseline profit flow $\pi_M$ when a single firm $i$ is active and $\pi_D$ when both are. These values may reflect either standard duopoly competition ($0 < 2\pi_D \leq \pi_M$) or competition with complementary product differentiation ($\pi_M \leq 2\pi_D \leq 2\pi_M$). Flow profit at time $t$ is scaled by a multiplicative component representing market size ($Y_t$) so as to take the form $\pi_M Y_t$ or $\pi_D Y_t$, and this state variable is assumed to follow a geometric Brownian motion ($dY_t = \alpha Y_t dt + \sigma Y_t dZ_t$ where $(Z_t)_{t\geq 0}$ is a standard Wiener process and $\alpha, \sigma \geq 0$) reflecting the idea that demand for a new product evolves in a context of uncertainty. Profit flows begin instantaneously and with certainty once investment has occurred. Firms have a common and constant discount rate assumed to be large enough that the investment problem is economically meaningful ($r > \alpha$), and information is symmetric.

Introducing the new product involves an irrecuperable fixed cost ($I$) for the the first firm that invests to serve demand, *i.e.* for the *innovator*. A firm that observes its rival's innovation can invest afterwards (even immediately afterwards) as a second entrant, *i.e.* as an *imitator*. We assume that in addition to the various standard setup costs associated with bringing a product to market such as dedicated plant and equipment, marketing expenditures, and so forth, the follower incurs a cost of imitation of variable magnitude depending on technological or institutional factors. Introducing the alternative version thus also involves an irrecuperable fixed cost ($K$), though we allow for the extreme case of costless imitation. The imitator's fixed cost may be either higher or lower than the innovator's, depending for instance on the difficulty of reengineering and on the degree of IP protection afforded to the innovator. If the second firm can develop the same product independently, imitation clearly should be no more expensive than innovation ($K \leq I$) in the absence of IP. When the product is complex enough or legal protection is sufficiently strong, imitators must invest in reverse engineering or invent around any patents held by the innovator and the second mover incurs higher entry costs than the leader ($K > I$),[4]

[4] Our focus is the relation between innovation and imitation, but other circumstances can also lead to asymmetric fixed costs for ex-ante identical firms. If developing the new product involves scarce assets, such as prime location in real estate or natural resource extraction, then the imitator may face a higher cost ($K > I$). Also, imperfect competition in input markets may result in asymmetric investment costs. In Billette de Villemeur et al. [2], investment cost is determined endogenously by a strategic input supplier, resulting in a discounted input price for the first firm that invests ($I < K$).

and we allow for the possibility of an arbitrarily high imitation cost $K^* = \infty$.

### 2.2 Firm strategies and leader and follower payoffs

In order to focus on the economics of firm entry decisions, we model the entry process as a game with Markovian strategies along the lines proposed by Thijssen [33] that captures the relevant features of a more general game in stopping times. Firms are thus assumed to choose investment thresholds which determine stochastic investment times. Thus the strategy of a firm $i$, $i = 1, 2$, consists of an initial entry threshold $Y_i \in R_+$ that, once reached for the first time from below and absent prior rival entry, triggers investment, and which is associated with a stopping time $\tau_i := \inf\{t \geq 0 \,|Y_t \geq Y_i\}$. The choice of entry thresholds endogenously determines the role of each firm as innovator or imitator. In the case of identical thresholds ($Y_1 = Y_2$) when it would only be optimal for one of the firms to invest, a standard coordination problem arises and a tie-breaking rule determines firm roles.

Industry dynamics may thus be viewed as resulting from a two stage interaction which unfolds over time, where in a first stage (which determines the onset of the monopoly phase of the industry) the choices of initial entry thresholds $(Y_1, Y_2)$ determine the roles of the firms, and in a second stage (the onset of the duopoly phase), the remaining firm enters at a threshold of its choice. The latter continuation game is a single-firm decision problem, and the remaining firm's follower threshold is denoted by $Y_F^*$, $Y_F^* \geq Y_i$, and associated with a stopping time $\tau_F^* := \inf\{t \geq 0 \,|Y_t \geq Y_F^*\}$, which is specified further below.

Given an initial level of the multiplicative shock $Y_0 = y$ the expected payoffs for innovation and imitation at a threshold $Y_i$ are:

$$L_y\left(Y_i, Y_F^*\right) = \mathbb{E}_y\left[\int_{\tau_i}^{\tau_F^*} e^{-rs}\pi_M Y_s ds - e^{-r\tau_i} I + \int_{\tau_F^*}^{\infty} e^{-rs}\pi_D Y_s ds\right] \tag{1a}$$

$$= \begin{cases} \frac{\pi_M}{r-\alpha}y - I - \frac{\pi_M - \pi_D}{r-\alpha}Y_F^*\left(\frac{y}{Y_F^*}\right)^\beta & , \quad Y_i \leq y \\ \left(\frac{\pi_M}{r-\alpha}Y_i - I\right)\left(\frac{y}{Y_i}\right)^\beta - \frac{\pi_M - \pi_D}{r-\alpha}Y_F^*\left(\frac{y}{Y_F^*}\right)^\beta & , \quad Y_i \geq y \end{cases} \quad \textit{(innovator payoff)} \tag{1b}$$

and

$$F_y(Y_i;K) = \mathbb{E}_y\left[\int_{\tau_i}^{\infty} e^{-rs}\pi_D Y_s ds - Ke^{-r\tau_i}\right] \tag{2a}$$

$$= \begin{cases} \frac{\pi_D}{r-\alpha}y - K & , \quad Y_i \le y \\ \left(\frac{\pi_D}{r-\alpha}Y_i - K\right)\left(\frac{y}{Y_i}\right)^{\beta} & , \quad Y_i \ge y \end{cases} \quad \textit{(imitator payoff)}, \tag{2b}$$

where in both (1b) and (2b), $\beta$ is shorthand for the function of parameters

$$\beta(\alpha,\sigma,r) := \frac{1}{2} - \frac{\alpha}{\sigma^2} + \sqrt{\left(\frac{1}{2} - \frac{\alpha}{\sigma^2}\right)^2 + \frac{2r}{\sigma^2}} \tag{3}$$

with $\beta(\alpha,0,r) = \lim_{\sigma\to 0}\beta = r/\alpha$. The function $\beta$ given in (3) is a standard expression in real option models, satisfying $\beta > 1$. A lower value of $\beta$ is associated with a greater incentive to wait (it is straightforward to check that $\partial\beta/\partial\alpha < 0$, $\partial\beta/\partial\sigma < 0$, and $\partial\beta/\partial r > 0$), so $\beta$ may be interpreted as a measure of "impatience". The $(y/\cdot)^{\beta}$ terms in which $\beta$ occurs reflect the expected discounting of the monetary units that are received when the stochastic process reaches the relevant thresholds for the first time.[5] Here and throughout the rest of the paper $L$ and $F$ subscripts refer to "leader" and "follower".

Both firms may introduce their respective products independently at the same moment, in which case they both incur the fixed cost of innovation, $I$. The resulting payoff is $M_y(Y_i) := L_y(Y_i, Y_i)$ $(= F_y(Y_i; I))$.

The leader (innovator) payoff is composed of two terms, which correspond to the monopoly profit flow of the innovating firm and the possible negative impact of the second firm's entry.

Assuming the current level of the demand shock $y$ is sufficiently small so that it is optimal for firms to delay investment (*e.g.* $y \le (r-\alpha)I/\pi_M$) the payoff functions $L_y$, $F_y$ and $M_y$ are quasiconcave over the relevant parts of their domains and attain non-negative maxima at $Y_L := (\beta(r-\alpha)I)/((\beta-1)\pi_M)$ and $Y_F := (\beta(r-\alpha)K)/((\beta-1)\pi_D)$ and $Y_S := (\beta(r-\alpha)I)/((\beta-1)\pi_D)$ respectively. We refer to the thresholds $Y_L$ and $Y_F$ as the optimal standalone leader and follower thresholds, and the pair of strategies $\{Y_L, Y_F\}$ corresponds to the open loop equilibrium

[5] If $\sigma = 0$ the stopping time $\tau_i$ is deterministic and $(Y_0/Y_i)^{\beta} = e^{-r\tau_i}$ is the standard continuous time discounting term under certainty.

identified by Reinganum [29].[6] A key property of the payoff functions which is used throughout our analysis is that the leader payoff is nondecreasing in the imitation cost provided the follower invests at the optimal follower threshold ($\partial L_y(Y_i, Y_F^*)/\partial K \geq 0$), whereas the follower payoff is decreasing in the imitation cost ($\partial F_y(Y_i; K)/\partial K < 0$).

Note that there are two kinds of simultaneous investment outcomes that can arise in the model. If one firm invests first and thereby takes the role of innovator, but does so at a sufficiently high threshold ($Y_i \geq Y_F$), the remaining firm then chooses to invest immediately after, although it has the follower role and its payoff is $F_y(Y_i; K)$. On the other hand, if both firms happen to invest simultaneously without coordinating their investments, they receive the same payoff $M_y(Y_i)$.

Lastly then, it may occur that at a given threshold $Y_i$ both firms seek to invest whereas it would only be optimal for one of them to do so. This happens if, letting $Y_F^* := \max\{Y_i, Y_F\}$ denote the optimal follower threshold conditional on a first investment at $Y_i$, $F_y(Y_F^*; K) > M_y(Y_i)$. In this case we assume that either firm is equally likely to invest as a leader or as a follower with probability

$$p_y(Y_i; K) = \begin{cases} \frac{F_y(Y_F^*;K) - M_y(Y_i)}{L_y\left(Y_i, Y_F^*\right) + F_y\left(Y_F^*;K\right) - 2M_y(Y_i)} & \text{if} \quad L_y(Y_i, Y_F^*) \geq F_y(Y_F^*; K) \\ 0 & \text{if} \quad L_y(Y_i, Y_F^*) < F_y(Y_F^*; K) \end{cases} \tag{4}$$

so accordingly $1 - 2p_y(Y_i; K)$ is the probability that "mistaken" simultaneous investment occurs.[7]

## 2.3 Equilibrium

We study the symmetric Markov perfect equilibrium of the entry game. One reason for this is that as firms are taken to be symmetric ex-ante, it seems natural to suppose that they hold symmetric beliefs about each other's play at the beginning of the investment game. In so doing, the equilibrium described in Proposition 1 below is

[6] For sufficiently low values of $K$ ($K \in [0, K_l)$ where $K_l := (\pi_D/\pi_M)I$), $Y_F < Y_L$. In this range, if roles were exogenously assigned, a follower would be willing to pay its rival to induce it to invest earlier. By definition the effective follower investment threshold must be $\max\{Y_L, Y_F\}$. We mention this possibility for completeness but the value $K_l$ does not play a significant role in the rest of the analysis.

[7] The tie-breaking rule (4) satisfies the rent-equalization property (Thijssen [33]). One contrast between our model and a standard real option game is that the values of the leader and follower payoffs generally differ at $Y_F$ because of the asymmetry in investment costs.

consistent with the earlier approaches of Fudenberg and Tirole [12] and Hendricks et al. [19].[8] Another reason is that this equilibrium exhibits rent dissipation, a feature that is emphasized in the early timing game literature as surveyed by Fudenberg and Tirole [13], and which leads to a smooth dependence of equilibrium on imitation cost that is of compelling simplicity.

As stated above the game described in Section 2.2 occurs in two effective stages. Firms first compete in initial entry thresholds that endogenously determine their roles as innovators or imitators, and subsequently any remaining firm selects its follower entry threshold:

- Stage 1: both firms select initial entry thresholds $(Y_1, Y_2)$ (or distribution thereof) that determine innovator and imitator roles;
- Stage 2: if a single firm $(i)$ innovates, the remaining firm $(j)$ then selects its imitator entry threshold.

To determine the equilibrium choices, note first that once one of the firms has invested, any firm that remains out of the market holds a standard growth option. It prefers to wait if the first investment has occurred early enough (before $Y_F$ is reached), and otherwise to invest immediately. Thus in the continuation game that follows investment by firm $j$ at a threshold $Y_j$, the optimal policy of firm $i \neq j$ is to invest at a threshold $Y_F^* = \max\{Y_j, Y_F\}$ which results in the optimal follower value $F_y(Y_F^*; K)$.

By backward induction a given firm $i$'s initial payoff is therefore

$$V_y(Y_i, Y_j) = \begin{cases} L_y(Y_i, Y_F^*) & \text{if} \quad Y_i < Y_j \\ p_y(Y_i; K) L_y(Y_i, Y_F^*) + p_y(Y_i; K) F_y(Y_F^*; K) + (1 - 2p_y(Y_i; K)) M_y(Y_i) & \text{if} \quad Y_i = Y_j \\ F_y(Y_F^*; K) & \text{if} \quad Y_i > Y_j \end{cases} . \tag{5}$$

A pair of investment triggers $(Y_1^*, Y_2^*)$ is a pure strategy Markov perfect equilibrium of the duopoly investment game if for all $y$, $V_y\left(Y_i^*, Y_j^*\right) \geq V_y\left(Y_i, Y_j^*\right)$, $i, j \in \{1, 2\}$, $i \neq j$.

[8] In games of attrition models, authors have sometimes proceeded differently. Notably, Hoppe [21] focuses on asymmetric equilibrium in pure strategies. This applies if, for instance, the same entry game is played in several independent markets and pre-play communication enables firm coordination, but we do not allow for this possibility here.

The following proposition establishes the existence a critical imitation cost, $\widehat{K}$,[9] that determines the nature of the duopoly investment game. This imitation cost is defined implicitly by the condition that firms are indifferent in equilibrium between the innovator and imitator payoffs when these are evaluated at the optimal standalone thresholds, that is $L_{Y_L}(Y_L, Y_F) = F_{Y_L}\left(Y_F; \widehat{K}\right)$ (note that $Y_F$ is a function of $K$). It is also useful for the proposition to define another critical imitation cost level, $\widetilde{K} \in \left(0, \widehat{K}\right)$, which is the solution to the condition $L_{Y_L}(Y_L, Y_F) = M_{Y_L}(Y_S)$. Finally for $K \geq \widehat{K}$ so that this is well-defined, let $Y_P$ denote the lower root of the condition $L(Y_P, Y_F) = F(Y_F; K)$. This threshold is usually referred to as the preemption threshold, and here describes the point of indifference between innovating and imitating.[10]

**Proposition 1** *The duopoly investment game has a unique symmetric Markov perfect equilibrium characterized by a critical imitation cost $\widehat{K} \leq I$ such that:*
*(i) If the imitation cost is low ($K < \widehat{K}$), firms play a game of attrition. There is an equilibrium in mixed strategies with innovation triggers distributed continuously over $[Y_S, \infty)$ (if $K \leq \widetilde{K}$) or over a disconnected support of the form $[Y_L, Y_{S'}] \cup [Y_S, \infty)$ (if $\widetilde{K} < K < \widehat{K}$). In the former case imitation occurs immediately after innovation, otherwise it can occur later with positive probability at the imitation threshold $Y_F$ .*
*(ii) If the imitation cost is intermediate ($K = \widehat{K}$), the equilibrium innovation triggers are $(Y_L, Y_L)$ and the innovation and imitation thresholds are $Y_L$ and $Y_F$ respectively.*
*(iii) If the imitation cost is high ($K > \widehat{K}$), firms play a game of preemption. The equilibrium innovation triggers are $(Y_P, Y_P)$ and the innovation and imitation thresholds are $Y_P$ and $Y_F$ respectively.*

In order to illustrate the cases described in Proposition 1, Figures $1-5$ depict

[9] See Section A.1 for a characterization of the critical values:
$\widehat{K} := \left((1 + \beta((\pi_M/\pi_D) - 1)) / (\pi_M/\pi_D)^\beta\right)^{1/(\beta-1)} I$,
and $\widetilde{K} := \left(\beta((\pi_M/\pi_D) - 1) / \left((\pi_M/\pi_D)^\beta - 1\right)\right)^{1/(\beta-1)} I$.

[10] Note that simultaneous innovation almost never occurs in this model. In the attrition range, setting identical stage 1 thresholds ($\widetilde{Y}_1 = \widetilde{Y}_2$) is a zero probability event, although if it ever were to happen firms would invest simultaneously according to the tie-breaking rule. In the preemption range, firms choose identical thresholds ($Y_1^* = Y_2^* = Y_P$) but coordinate so that either firm invests as a leader with equiprobability.

leader and follower payoffs in the first stage of the game, for different values of the imitation cost and assuming $y$ is not too large. Throughout these figures, as the imitation cost increases, the optimal standalone follower threshold $Y_F$ increases and the follower payoff shifts down and towards the right. Because of the longer monopoly phase, the first stage leader payoff $L_y(Y_i, Y_F^*)$ accordingly shifts upward over the range of thresholds $(Y_0, Y_F)$ over which investments are sequential. Note that the optimal standalone leader threshold $Y_L$ is independent of $K$, and the leader payoff function has a kink at $Y_F$ which constitutes the lower bound of the range of thresholds over which innovator and imitator entry are simultaneous. In Figures 1 and 2, there is a *second-mover advantage* (in the sense that $L_y(Y_L, Y_F) < F_y(Y_F; K)$) and the game is one of attrition. Figure 3 represents the intermediate case in which the imitation cost attains its critical value, $K = \widehat{K}$, and there is neither a first-mover advantage nor a second-mover advantage. In Figures 4 and 5, there is a *first-mover advantage* (in the sense that $L_y(Y_L, Y_F) > F_y(Y_F; K)$), and the game is one of preemption, with the first investment occurring at $Y_P$.

In the symmetric equilibrium described in Proposition 1, because the roles of firms are endogenous, positive rent dissipation occurs whenever the firms play a game of attrition ($K < \widehat{K}$) or of preemption ($K > \widehat{K}$). The expected value of firms in equilibrium can therefore be characterized. To state the following corollary, some further notation is necessary. Since $L_y(Y_i, Y_F^*)$ can have two local maxima, in cases of attrition in which the imitation cost is sufficiently low such as that illustrated in Figure 1, its global maximum may be attained at $Y_S := \arg\max M(Y_i)$, which then corresponds to the lower bound of the support of innovator entry thresholds. Thus, define $Y_L^* := \arg\max L_y(Y_i, Y_F^*)$ with $Y_L^* \in \{Y_L, Y_S\}$ to refer to the lower bound of the appropriate attrition equilibrium threshold distribution.

**Corollary 2** *In a symmetric equilibrium the expected payoffs of firms* $\mathbb{E}V_y\left(\widetilde{Y}_1^*, \widetilde{Y}_2^*\right)$ *are identical and equal to* $\min\{L_y(Y_L^*, Y_F^*), F_y(Y_F; K)\}$*, that is (for* $Y_0$ *low enough) to the lowest of the diffusion equilibrium payoffs.*

The dependence of the critical threshold imitation cost $\widehat{K}$ on model parameters is straightforward. The next corollary gives sensitivity results with respect to the intensity of competition in the product market $(\pi_M/\pi_D)$ and discounting $(\beta)$.

**Corollary 3** *The more intense product market competition is* $(\pi_M/\pi_D)$ *and the more firms discount the future* $(\beta)$*, the more likely it is that preemption occurs, and con-*

*versely for attrition:*

$$\frac{\partial \widehat{K}}{\partial \left(\pi_M/\pi_D\right)} < 0 \text{ and } \frac{\partial \widehat{K}}{\partial \beta} < 0. \tag{6}$$

To provide intuition for this corollary, recall that the process $Y_t$ is stochastic and that there is an option value for firms to wait before investing that is positively related to volatility. Provided that there is an inherent advantage to imitation ($K < I$), for some parameter values and in particular for large enough volatility (in which case $K < \widehat{K}$), this option value outweighs any preemptive motive to secure monopoly rents. A similar reasoning holds if the drift in demand is sufficiently high. That is to say, as $\partial\beta/\partial\sigma < 0$ and $\partial\beta/\partial\alpha < 0$ by Corollary 3 *an attrition regime is more likely in industries with greater trend growth and demand volatility.* This particular comparative static is important because it provides a countervailing force to several of the mechanisms that are discussed in the rest of the paper. As the next sections show, institutional conditions such as IP protection and firm choices regarding both technology and licensing generally serve to make market entry regimes more preemptive and attrition relatively rare. One would therefore expect attrition to be relatively rare, except in industries significant enough degree of demand growth or demand uncertainty.

## 3 Imitation cost, industry profit, and welfare

The previous section shows the central role played by the relative fixed cost of imitation in determining the nature of strategic competition and the equilibrium pattern of entry in an industry, and ultimately the industry's levels of consumer surplus and welfare. This imitation cost is likely to be driven by several different factors including technological conditions and the level of IP protection. It thus varies from industry to industry and can be influenced *ex-ante* by regulators, typically through a choice of patent breadth. These considerations raise the question of determining what may be desirable levels of imitation cost. At first glance this decision appears to involve a simple trade-off since a higher imitation cost is socially wasteful but also hastens innovator entry. However different effects arise with regard to imitator entry in the preemption and attrition regimes that need to be examined more carefully.

For simplicity, throughout the discussion that follows we assume that regulators act at a sufficiently early stage of industry development to influence both innovation and imitation decisions, *i.e.* that the initial demand state satisfies $Y_0 \leq (r-\alpha)\, I/\pi_M$, and drop the $y$ subscripts on payoff functions. In addition, let $Y_I = \min\{Y_1^*, Y_2^*\}$ denote the threshold at which innovation occurs in equilibrium, which takes the random value $\min\left\{\widetilde{Y}_1^*, \widetilde{Y}_2^*\right\}$, or $Y_L$ or $Y_P$ depending on the level of $K$ as per Proposition 1.

### 3.1 Industry performance

A useful preliminary step to conducting a welfare analysis is to first consider industry performance only, which allows us to derive an intermediate result regarding industry value. We thus begin by studying the relationship between imitation cost, first- and second-mover advantage, and industry profitability. A first and seemingly obvious consideration that emerges from our framework is that lower imitation cost is a necessary, but not a sufficient condition for second mover advantage. Too see why, note that if firms in an industry have identical fixed costs there is an inherent first-mover advantage that results from the monopoly phase of the entry game $(L(Y_L, Y_F^*) \geq F(Y_F^*; I))$. The degree of first-mover advantage in this case is determined by the relative importance of monopoly profit in the product market $(\pi_M/\pi_D)$. A second-mover advantage, on the other hand, arises through the input market when the relative cost of imitation $(K/I)$ is sufficiently low to compensate for foregoing the period of monopoly profit. Thus the empirical presence of lower costs for imitators, as has been observed by different authors (Mansfield et al. [27], Samuelson and Scotchmer [30]), does not by itself ensure that firms will find it desirable to pursue so-called imitation strategies in a dynamic setting.

Next, in the symmetric equilibrium of our model, there is a monotone relationship between imitation cost and innovation thresholds, as well between imitation cost and imitation lags, as follows. First, the higher is the imitation cost, the higher is the standalone threshold for the follower firm $(Y_F)$, although actual follower entry may occur either at this threshold or possibly later if the investment game is one of attrition. The effect of the resulting delay in follower entry on the innovation threshold $Y_I$ is similar throughout the range of imitation costs. As imitation cost increases, in the attrition regime it is the distribution of innovator entry thresholds $(\widetilde{Y}_I)$ that is shifted leftward whereas in the preemption regime rent equalization directly results in a lower preemption threshold $(Y_P)$. However, the effect of higher imitation cost on the distribution of follower investment (imitation) thresholds $(Y_F^*)$ is not itself monotone in $K$ in every regime. Under attrition, imitator entry occurs at a higher threshold as a result of an increase in imitation cost conditionally upon the innovation threshold realization being low enough (if $\widetilde{Y}_I \leq Y_F$), but the imitator entry threshold is random otherwise and its distribution is shifted leftward if the innovator enters late (if $\widetilde{Y}_I > Y_F$). Nevertheless, the *gap* (and therefore the expected time lag) between innovation and imitation thresholds, $Y_F^*\left(\widetilde{Y}_I\right) - \widetilde{Y}_I$, can be shown to increase stochastically with imitation cost. To summarize, higher imitation cost may properly be said to accelerate innovative investment and to delay the arrival of imitative investment conditional upon innovation having occurred.

Lastly, the equilibrium characterized in Proposition 1 leads to a particularly simple result regarding industry performance. Because in the different regimes of attrition and preemption, competition between firms to secure either second- or first-mover advantages results in the dissipation of any positional rents and since leader value increases in imitation cost whereas follower value decreases, it is only when the level of the imitation cost is such that neither of these regimes occurs (case $(ii)$, $K = \widehat{K}$) that investment thresholds are set optimally from the standpoint of industry profit. Thus all else equal, it is in those industries in which imitation cost reaches the level where firms do not have an incentive to seek positional advantages of either sort that industry value is maximized.

**Proposition 4** *Viewed as a function of imitation cost, expected industry value is single-peaked, constant over* $\left(0, \widetilde{K}\right)$*, and attains its maximum when neither attrition nor preemption occur (at* $K = \widehat{K}$*).*

According to Proposition 4 there exists a range over which expected firm value $\mathbb{E}V\left(\widetilde{Y}_1^*, \widetilde{Y}_2^*\right) = M(Y_S)$ is unaffected by imitation cost. But there is also a range of imitation cost levels $\left(\widetilde{K}, \widehat{K}\right)$ over which greater resource costs are strictly beneficial to the industry. That is to say, if fixed costs are sufficiently high to shield an innovator from instantaneous imitation with positive probability, product introduction is more timely and both firms benefit *ex-ante*. In addition, Proposition 4 is also instrumental in establishing our main welfare results, Proposition 5 and Proposition 7 below.

## 3.2 Optimal protection of innovation

We take the view that regulators can influence the relative cost of imitation (at least upward) through a choice of IP protection level that we interpret as patent breadth. With this single instrument and provided that the natural imitation cost is not so high as to bind the regulator, the imitation cost $K$ may be considered to be a decision variable. We consider a second-best welfare benchmark in which firms are free to select their entry thresholds and product market output or prices.

To provide some intuition for the analysis that follows, expected welfare in this model can be broken down into three parts: expected industry value, consumer surplus from innovator entry, and consumer surplus from imitator entry. The first of these is maximized at the critical imitation cost $\widehat{K}$ (Proposition 4) whereas the other two parts both depend on $K$ directly as well as indirectly through the equilibrium innovation and imitation thresholds. A higher imitation cost unambiguously accelerates innovator entry which raises consumer surplus, so the second of these welfare components is clearly increasing in $K$. But in the case of an attrition regime, the impact of imitation cost on the last of the three components of welfare is more complex, since an increase in $K$ may either delay imitator entry (through its effect on the standalone threshold $Y_F$ at which imitation occurs with positive probability) or hasten it (if the innovation threshold realization is greater than $Y_F$ so that imitation is immediate). Conceivably then, even though raising imitation cost from a sufficiently low initial level $K < \widehat{K}$ increases industry profit and the consumer surplus from innovation, an attrition outcome may still be socially desirable if the loss of consumer surplus resulting from delayed imitation is large enough.

To formalize these insights, suppose that consumer surplus is scaled by the market size parameter $Y_t$, as is the case for firm profits. Let $\mathrm{CS}_M$ and $\mathrm{CS}_D$ then denote the unit flows of consumer surplus under monopoly and under duopoly respectively. The social discount rate is assumed to be identical to that of firms. Recall from Proposition 3 that equilibrium innovative investment threshold $Y_I$ is stochastic in an attrition regime, and that the distributions of the threshold of both innovative investment and follower investment ($Y_F^*$) are functions of $K$. Expected social welfare is

$$W(K) = \mathbb{E}\left[\underbrace{2V\left(\widetilde{Y}_1^*, \widetilde{Y}_2^*\right)}_{\text{industry value}} + \underbrace{\frac{\mathrm{CS}_M}{r-\alpha}\left[\widetilde{Y}_I\right]^{-(\beta-1)} Y_0^{\beta}}_{\text{consumer surplus from innovation}} + \underbrace{\frac{(\mathrm{CS}_D - \mathrm{CS}_M)}{r-\alpha}\left[Y_F^*\right]^{-(\beta-1)} Y_0^{\beta}}_{\text{consumer surplus from imitation}}\right]. \tag{7}$$

The first summand in (7) is the industry's expected value. By Proposition 4 it is equal to $2\min\{L(Y_L^*, Y_F^*), F(Y_F^*; K)\}$, which is single-peaked with respect to $K$ with a maximum at $\widehat{K}$. The second term is the consumer surplus that results from innovative investment. The expected value of this term increases with $K$, since a higher imitation cost shifts the distribution of innovator entry thresholds (which may be degenerate, *e.g.* under preemption) leftward. The third term is the consumer surplus that results from the imitator's entry into the market. The effect of increasing $K$ on this term is ambiguous, as it encompasses the two opposing effects discussed above. Nevertheless, the effect of raising imitation cost on welfare can be partially characterized as follows (see Section A.4 for a proof, the main steps of which are outlined below).

First, within the range of preemption regimes ($K > \widehat{K}$) the innovator and imitator entry thresholds are respectively $Y_P$ and $Y_F$ and the local optimum of (7), $K_P$, has an explicit form. For a range of parameter values $\beta \in [\beta_0, \infty)$, $\beta_0 > 1$, this optimum is a corner solution ($K_P = \infty$) signifying that the social planner's imitation cost instrument is of too limited a reach to attain its welfare objective. Put another way, for sufficiently large $\beta$ the optimal form of preemption is a winner-take-all contest. Since the greatest amount of preemption that the social planner can induce does not generate enough competition to induce firms to enter sufficiently early in such cases, a single firm is active *ex-post* whose investment threshold is determined by the threat of potential entry. On the other hand, if discounting is not too strong so $\beta \in (1, \beta_0)$ as occurs for instance if volatility is large, $K_P$ is finite and strictly greater than $\widehat{K}$ so long as consumer surplus under monopoly is positive ($\mathrm{CS}_M > 0$).

Second, within the range of attrition regimes, there can be another local maximum of welfare. To establish its existence, because (7) is continuous, it is sufficient to show that social welfare is decreasing to the left of the critical value $\widehat{K}$. To see why this may occur set $\mathrm{CS}_M = 0$ for simplicity so that the middle term in (7) drops out. Also, note that the expected industry value term $\mathbb{E}2V\left(\widetilde{Y}_1, \widetilde{Y}_2\right)$ reaches a maximum at $\widehat{K}$ so $\partial \mathbb{E}V\left(\widetilde{Y}_1, \widetilde{Y}_2\right)/\partial K\Big|_{\widehat{K}} = 0$. Then the behavior of social welfare to the left of $\widehat{K}$ is determined by the remaining consumer surplus from imitation ($\mathrm{CS}_D - \mathrm{CS}_M$) term. In an attrition regime and for values of $K$ near $\widehat{K}$ (for $K \in \left(\widetilde{K}, \widehat{K}\right)$) this term has two distinct parts depending on whether the innovator investment threshold realization is below $Y'_S$ (in which case imitator investment occurs at $Y_F$, see Figure 2) or above (in which case imitation immediately follows innovation). Accounting for the equilibrium distribution of $Y_F^*$ therefore gives this term as

$$\frac{(\mathrm{CS}_D - \mathrm{CS}_M)}{r - \alpha} [Y_F]^{-(\beta-1)} Y_0^{\beta} \left( G_\wedge (Y_S; K) + \int_{Y_S}^{\infty} (Y_F/s)^{\beta-1} dG_\wedge(s; K) \right) \quad (8)$$

where $G_\wedge(\cdot; K)$ denotes the distribution of $\widetilde{Y}_I$. However, we have $G_\wedge\left(Y_S; \widehat{K}\right) = 1$ and $\partial G_\wedge\left(Y_S; \widehat{K}\right)/\partial K = 0$. To the left of the critical value $\widehat{K}$ therefore, changes in $K$ have a second-order effect on the distribution of entry thresholds compared with their effect on $Y_F$. Thus an envelope argument on the welfare expression (7) establishes that $\lim_{\widehat{K}_-} \partial W(K)/\partial K < 0$.

Finally, either type of local maximum (under attrition or preemption) can be a global maximum, depending on the relative magnitude of the consumer surplus resulting from innovation and from imitation.

**Proposition 5** *In a constrained social optimum*

*($i$) either attrition or preemption may be optimal;*

*($ii$) if the optimum involves attrition, the imitator incurs a positive cost $K^* > \widetilde{K}$ ("winner-pay-some");*

*($iii$) if the optimum involves preemption, innovation occurs at the threshold $Y_P^* = \psi Y_L$, $\psi \in [(\beta-1)/\beta, 1]$; for $\beta$ large enough, a perpetual monopoly ("winner-take-all") is optimal.*[11]

The upshot of Proposition 5 is that there is no "one size fits all" prescription with respect to balancing the incentives of innovating and imitating firms, suggesting that policy is best determined on a case by case basis according to a number of industry conditions. Nevertheless, the proposition is informative in a number of ways.

To begin with, part ($i$) is of particular importance insofar as some prominent researchers have argued for the abolition of patents altogether (Boldrin and Levine [3]). Our model points to the fact that such an abolition may be desirable only to the extent that the natural cost of imitation is sufficiently high, *i.e.* $K \geq \widetilde{K}$. Moreover, this lower bound on imitation cost has an intuitive characterization, in that industry conditions must be such that an innovator has some positive *ex-ante* probability of earning a monopoly profit, rather than the certainty of facing immediate imitation (even if imitation results in a positive duopoly rent).

In the following corollary, two polar cases provide further economic intuition for part ($i$) of the proposition.

**Corollary 6** *In a constrained social optimum*

[11] See Section A.4 for a characterization of $\psi := \max\left\{\frac{\beta-1}{\beta}, \left(\frac{\mathrm{CS}_D - \mathrm{CS}_M}{\pi_D} + \frac{2}{\beta}\right) / \left(\frac{\mathrm{CS}_D}{\pi_D} - \frac{\beta-1}{\beta}\frac{\mathrm{CS}_M}{\pi_M} + \frac{2}{\beta}\right)\right\}$. It bears mention that in the literature, preemption thresholds generally do not have analytic expressions.

(*i*) *if there is perfect price discrimination under monopoly* ($\mathrm{CS}_M = 0$)*, attrition is socially optimal;*

(*ii*) *if there is either a unit demand or collusion in the product market* ($\mathrm{CS}_D + 2\pi_D = \mathrm{CS}_M + \pi_M$)*, preemption is socially optimal.*

While the optimal value of welfare in the preemption range has a closed form expression, the characterization of the optimal value of welfare in the attrition range is more complex. We therefore make a further restriction and suppose that the static entry incentive is socially excessive ($\pi_D \geq (\mathrm{CS}_D + 2\pi_D) - (\mathrm{CS}_M + \pi_M)$) in order to obtain our next welfare proposition. To provide a rationale for this restriction, recall that in a static setting with symmetric firms and homogeneous goods, Mankiw and Whinston [25] show that there is excess entry in an industry if total output increases whereas individual outputs decrease in the number of firms (the *business-stealing* effect) and argue that these assumptions characterize a broad range of models of oligopoly. In our dynamic setting, this assumption allows us to bound the welfare associated with the imitator's entry, $\mathbb{E}(\,(\mathrm{CS}_D - \mathrm{CS}_M)\,/(r-\alpha))Y_F^*\,(Y_0/Y_F^*)^{\beta}$, by the expected value of a duopoly firm, and hence by $\mathbb{E}\,V\,(Y_L, Y_L)|_{K=\widehat{K}}$ (according to Proposition 4), so as to establish the following.

**Proposition 7** *Suppose that the static private entry incentive is socially excessive. Then, in a constrained social optimum preemption is optimal if*[12]

$$\frac{\mathrm{CS}_M}{\pi_M} \geq \Omega\,(\beta)\,. \tag{9}$$

The right-hand term is decreasing in $\beta$ with $\lim_{\infty} \Omega\,(\beta) = 0$, the condition (9) is satisfied for a given demand specification if there is sufficient discounting, as occurs if industry growth and volatility are sufficiently low.

We conclude this section by illustrating with the case of a common oligopoly specification.

**Example 8** *Suppose that the product market is characterized by a constant elasticity inverse demand* $P = AQ^{-\eta}$*,* $A > 0$ *and* $\eta \in (0,1)$*, and that firms have constant unit variable cost c. Then straightfoward calculations establish*

$$\frac{\mathrm{CS}_M}{\pi_M} = \frac{1}{1-\eta} \text{ and } \frac{\mathrm{CS}_D}{\mathrm{CS}_M} = 2\left(1 + \frac{1}{1-\eta}\right)^{\frac{1}{\eta}-1}. \tag{10}$$

[12]See Section A.5 for a derivation of $\Omega\,(\beta) := 2/\left(\left(\beta^{\beta}/\,(\beta-1)^{\beta-1}\right) - \beta\right)$.

*Therefore,* $\lim_{\eta\to 0}(\mathrm{CS}_D/\mathrm{CS}_M)=\infty$ *and* $\lim_{\eta\to 1}(\mathrm{CS}_M/\pi_M)=\infty$ *so that by Corollary 6 and Proposition 7, the second-best social optimum involves attrition if demand is sufficiently elastic and preemption if demand elasticity approaches unity.*

## 4 Endogenous entry barrier, buyout, and licensing

In this section, we discuss how further real-world aspects of innovation and imitation may be incorporated into the framework of the previous sections. One is the ability of an innovating firm to raise the entry barrier of the imitator, either through technological choices in product development that render reverse engineering more costly or by strengthening the patentability of its product. Another aspect is contracting between the innovator and the imitator, which typically takes the form of technology transfer that reduces the follower's imitation cost in a context similar to a licensing agreement, but can also involve a "pay for delay" agreement or a buyout. From a formal standpoint these extensions both add an intermediate stage to the investment game, once the innovator's entry has occurred and before the imitator invests. Moreover, by raising the standalone value of the innovating firm, they tend to favor first-mover advantage and the emergence of preemption regimes although the implications for imitation timing and welfare generally differ.

### 4.1 Endogenous entry barrier

Suppose that the innovating firm may rely on a varying degree of either legal or technical protection in order to influence the imitation cost of a subsequent entrant. In case of legal protection, the imitation cost level reflects the breadth of patents, with wider patents implying higher costs for inventing around so as to develop a non-infringing imitation. Moreover, firms may decide to pursue patent protection more or less aggressively, as is the case for pharmaceutical firms as discussed in Section 1.3. In case of technical protection, the imitation costs are imparted by reverse engineering, and increase with the complexity of the copied product. For instance, an innovating firm can expend effort to render its product more difficult to disassemble, or even add misleading complexity (Samuelson and Scotchmer [30]).

Such choices may be incorporated into our model by introducing a decision by the innovating firm at the time of its investment to expend an additional irrecuperable cost, which we denote by $\rho$, that raises the imitating firm's fixed cost by an amount $f(\rho)$, where $f$ is taken to be an increasing and weakly concave function, with $f(0)=0$

for simplicity. The cost $\rho$ is deducted from the innovator payoff $L(Y_i, Y_F^*)$ defined in (1a). The investment costs of the innovator and imitator are then redefined as $I(\rho) := I_0 + \rho$ and $K(\rho) := K_0 + f(\rho)$, where $I_0$ and $K_0$ represent baseline values where no effort is exerted on raising rival cost. With respect to the sequence of decisions, the choice of $\rho$ arises once the roles of firms are determined, at the moment the innovator enters and before the second firm's entry so that we have:

- Stage 1': both firms select initial entry thresholds $(Y_1, Y_2)$ that determine innovator and imitator roles;
- Stage 2': if a single firm $(i)$ innovates, it selects a degree of patenting effort and product complexity $(\rho)$;
- Stage 3': the remaining firm $(j)$ then selects its imitator entry threshold.

Proceeding by backward induction, in stage 3' the imitator payoff is a nonincreasing function of $K(\rho)$ and therefore of the innovator's effort $\rho$ whereas its entry threshold $Y_F^*(\rho) = \max\{Y_i, Y_F(\rho)\}$ is nondecreasing. In stage 2', with an endogenous barrier to imitation an innovator that enters at $Y_i$ has an adapted expected payoff $L_{\mathrm{e}}(Y_i, \rho)$ and faces the decision problem $\max_\rho L_{\mathrm{e}}(Y_i, \rho)$, and at an interior optimum the cost-raising effort satisfies

$$\frac{f'(\rho^*)}{(K_0 + f(\rho^*))^\beta} = \frac{\beta^{\beta-1}}{(\beta-1)^\beta} \frac{\pi_D}{\pi_M - \pi_D} \left(\frac{r-\alpha}{\pi_D}\right)^\beta Y_i^{-\beta}. \tag{11}$$

The reasoning for stage 1' proceeds as in the model of Section 2, save that the innovation and imitation payoffs take the respective forms $L_{\mathrm{e}}(Y_i, \rho)$ and $F(Y_F^*(\rho^*); K_0 + f(\rho^*))$. Whenever it is interior (positive) the optimal choice $\rho^*$ results in a higher innovator payoff, whereas the imitator payoff is lower: $(L_{\mathrm{e}}(Y_i, \rho^*) > L(Y_i, Y_F^*(0))$ and $F(Y_i; K_0 + f(\rho^*)) < F(Y_i; K_0))$. The equilibrium is as characterized in Proposition 1, the main differences being that the endogenization of $K$ results in more preemptive strategic investment with a lower critical threshold $\widehat{K}_{\mathrm{e}} < \widehat{K}$ separating the preemption and attrition regimes.

The endogeneity of entry barriers has some noteworthy economic consequences. To begin with, in those industries in which the cost of imitation is large enough so that entry competition is in the preemption range, as equilibrium payoffs are decreasing in imitation cost (Proposition 4), firms have a lower expected value than when the imitation cost is exogenous. To avoid this penalizing outcome firms would prefer to both commit *ex ante* not to exert any cost-raising effort in case they happen to lead the investment process as an outcome of stage 1' (since *ex post*, raising the imitation cost is a dominant strategy for the firm that happens to enter as an innovator in stage 2'). One way to achieve such a commitment is by agreeing to a common and open technological standard.

Moreover, the first-order condition of the innovator is informative as to the role of the baseline cost of imitation $K_0$. Since the left-hand side of (11) is a decreasing function that shifts downward as $K_0$ increases, a straightforward comparative static establishes that the effort to raise the level of entry barriers decreases with the baseline imitation cost, which it supplements ($\partial\rho^*/\partial K_0 < 0$). The latter property is in line with the biopharmaceutical industry case discussed in Section 1.3 where firms typically place greater reliance on patenting in the medications segment, in which natural entry barriers are low, than in the vaccines segment.

Thus,

**Proposition 9** *In an extended framework for strategic investment, with endogenous barriers to imitator entry, in a preemption regime firms benefit from agreeing ex-ante to a common standard; the lower the baseline cost of imitation, the higher the entry barrier set by the innovator.*

### 4.2 Buyout and licensing

The autonomous investment incentives of innovators and imitators having been described, it is then natural to allow for some common forms of contracting between firms. In the context of innovation and imitation, licensing is a particularly important possibility whenever some of the knowledge developed by the innovator can be transferred to the second firm. Other types of contracts that can be observed include a pay-for-delay agreement or a buyout, if these are allowed and provided that an imitator can commit not to enter the market over a certain period. Such agreements are typically concluded by pharmaceutical firms and generic manufacturers. In this context, a buyout in which the acquiring firm shuts down its rival may be thought of as a limiting case of pay-for-delay.

In order to focus broadly on the effects of contracting on entry timing, we make the simplifying assumption that firms have the ability to make a single spot transaction, which may involve a transfer either of technology or asset ownership in exchange for a lump sum payment. This simple form of contract suffices to illustrate a diversity of outcomes. We also assume that the contract is written by the innovator, who holds all the bargaining power.

As a result, the entry game has an intermediate stage, which consists of a dynamic agency problem in which the innovator incentivizes the imitating firm's investment behavior. Let $K_0$ denote an incompressible level of imitation cost reflecting such items as distribution and marketing expenses, and $K_I$ denote that part of the imitator's product development cost that can be eliminated by a technology transfer from the innovator, so the fixed cost of the imitator is $K := K_0 + K_I$. The sequence of moves is:

- Stage 1": both firms select initial entry thresholds $(Y_1, Y_2)$ that determine innovator and imitator roles;
- Stage 2": if a single firm $(i)$ innovates, it proposes a contract involving a transfer $(\varphi)$ from the innovator to the imitator ($\varphi > 0$ for a pay for delay or buyout, $\varphi < 0$ for a technology transfer);
- Stage 3": the remaining firm $(j)$ decides whether or not to accept the contract and selects its entry threshold.

The reservation value of the follower if it rejects any contract with the innovator is the value which results from the equilibrium described in the model of Section 2, $F(Y_F^*; K_0 + K_I)$. Because this reservation value is time-dependent until its realization at Stage 3", it is useful to denote its Stage 2" value as $F_0(Y_t)$ and we assume without loss of generality that the contract is proposed at the time either the innovator or the imitator enters, *i.e.* $Y_t = Y_i$ or $Y_t = Y_F^*$. There are then two cases to consider that depend on the comparative industry profits in monopoly and duopoly.
$(i)$ When the efficiency effect is present ($\pi_M/\pi_D \geq 2$) as occurs in many standard industrial organization settings, if it can do so effectively an innovator prefers to pay the imitator its reservation value at the time of its entry ($\varphi^* = F_0(Y_i)$) in order to delay imitation indefinitely (a buyout). Such an arrangement raises the expected payoff function of the leader and leaves the expected payoff of the follower unchanged, rendering a preemption regime more likely. All else equal, the magnitude of the impact on leader payoff depends on the strength of the efficiency effect, and if it is sufficiently strong or volatility is high enough (if $\pi_M/\pi_D \geq \beta+1$) attrition does not occur for any level of $K$. If $K \geq \widehat{K}$ so that the industry is in a preemption regime, then industry profits are pegged at $F_0(Y_F)$ and unaffected by the possibility of buyout, whereas they are weakly higher otherwise. The effect on consumer surplus is ambiguous, as innovation occurs earlier than it otherwise would but this must be balanced against

the absence of imitator entry into the product market. Taking two extreme examples, with perfect price discrimination under monopoly ($\mathrm{CS}_M = 0$) a takeover may or may not be socially efficient depending on the relative importance of additional innovator value and lost surplus from imitation, whereas if the product market would function as a cartel ($\pi_M/\pi_D = 2$) a buyout increases welfare only to the extent that it economizes on the fixed cost of imitation, $K$.

If a takeover is not allowed the best option for the innovator is to allow follower entry at the standard threshold $Y_F^*$, but set its maximum license fee at this moment $\varphi^* = K_I$ so as to recoup revenue from a part of the imitator's investment cost, thus reducing the duplication of R&D efforts. At the time of innovator entry, the discounted expected value of this fee reduces the innovator's irreversible cost of investment by the expected licensing revenue $K_I(Y_i/Y_F^*)^\beta$ and the leader payoff in stage 1" shifts up for an unchanged payoff function to the imitator, as in the case of a takeover. As with a buyout, a consequence of licensing is a lower critical imitation cost that separates the preemption and attrition regimes and a weakly increasing industry value. With licensing, the effect on consumer surplus is simpler. Licensing accelerates innovation under both preemption and attrition, leaving the arrival of imitation unchanged at $Y_F$, and is therefore unambiguously welfare improving.

*(ii)* If there is sufficient product market complementarity between firms ($\pi_M/\pi_D < 2$), imitation is a positive externality for the industry. The optimal imitator entry threshold for the industry is then $Y_F^{**} := \beta\,(r-\alpha)\,K_0/\,(\beta-1)\,(2\pi_D - \pi_M)$. It is greater than the standalone imitator threshold if the value of transferable technology is relatively small or if product complementarity is not too strong (if $\pi_M/\pi_D > 2 - (K_0/\,(K_0+K_I)))$ and smaller otherwise, in which case an innovator seeks to accelerate imitator entry. If it enters early enough to have leeway and imitation would occur too late otherwise ($Y_i, Y_F^{**} < Y_F$), the innovator induces the industry optimum by setting a license fee

$$\varphi^* = \frac{\pi_D}{r-\alpha} Y_F^{**} - K_0 - F_0\left(Y_F^{**}\right) \tag{12}$$

and we find $\varphi^* < K_I$ in this case. This result is noteworthy because the innovator then subsidizes the licensee to induce imitation at $Y_F^{**}$.[13] Returning to the biopharmaceutical example discussed throughout the paper, this result offers a rationalization

[13] The use of this simple licensing instrument increases welfare since innovation and imitation occur earlier while industry profit does not decrease, but this result is not robust to other forms of licensing.

for observed cooperation in the vaccine industry, when a research-intensive manufacturer transfers knowledge to a local competitor in a developing economy for a lesser payment than the investment that the technology recipient would have made in the absence of agreement (see WHO [38]).

Thus,

**Proposition 10** *In an extended framework for strategic investment, with contracting between the innovator and the imitator,* (*i*) *if the efficiency effect is sufficiently strong industry profits increase with buyouts and only preemption occurs, whereas when buyouts are ruled out licensing is profitable and increases welfare;* (*ii*) *if there are significant product market complementarities and imitation occurs late, the innovator may choose to subsidize its rival's entry in a licensing agreement.*

## 5 Conclusion

We have sought to develop an integrative framework to study some general questions regarding the allocation of resources to innovation and to imitation under imperfect competition. The analysis of the trade-off between static and dynamic inefficiency under imperfect competition highlights a novel channel through which the relative cost of imitation influences welfare, insofar as it impacts the timing of innovation and may even alter the nature of strategic competition (attrition *vs.* preemption). The broad message that emerges from our analysis remains a familiar one if demand growth and volatility are low enough, as typically occurs in mature industries. In this case, sufficient barriers to imitation should exist so that dynamic competition between firms is preemptive in nature, and if discounting is sufficient, it should take the form of a winner-take-all contest. But in those industries in which growth and volatility are sufficiently high, and which are those which are most typically associated with innovation, a form of limited attrition may be optimal in order for the benefits of imitation not to arrive too late. In that case the winner of the attrition game must "pay some", in the form of a positive imitation cost, which may not be so high as to render dynamic competition preemptive.

The theoretical model thus points towards a policy prescription that consists in tailoring IP protection to such general industry or market segment characteristics, as

If $Y_F^{**} > Y_F$ and the innovator can sign a forcing contract that is contingent on the imitator's entry threshold, imitation is optimally delayed by licensing and the consequences for welfare are ambiguous.

may suggest whether the greater risk lies in the direction of insufficient natural innovation incentives and attrition on the one hand, or an excessive preemption and delay in perceiving the benefits of imitation on the other. As described in our introductory example, biopharmaceutical firms must adapt their strategies to the contrasting technological conditions that they face in their medication and vaccine segments. In this particular industry moreover, significant steps have already been taken in order to adjust IP protection in response to identifiable categories of market conditions, in accordance with our policy conclusion. For orphan drugs and rare disease development, the U.S. Food and Drug Administration has enacted an enhanced form of IP protection (Orphan Drug Exclusivity) together with a tax credit that lowers the costs of clinical trials (Grabowski et al. [16]). Our analysis offers theoretical support to regulatory measures of this kind that adapt the relative cost of innovation and imitation to a market's specific characteristics, and which could also involve patent narrowing in case of high demand growth and volatility.

Among the extensions of our model that might be pursued, another step in the analysis would be to study incremental innovation (or "versioning") among existing firms in a market. In this setting, it is possible that simultaneous investment equilibrium solutions arise, suggesting that firms might coordinate on investment timing, and it is not much further to go to examine the possibility of cooperation in product development with these tools as well.

## References


[1] Azevedo A, Paxson D (2014) Developing real option game models, European Journal of Operational Research 237(3):909-920.

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

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

[4] Boyer M, Lasserre P, Moreaux M (2012) A dynamic duopoly investment game without commitment under uncertain market expansion, International Journal of Industrial Organization 30(6):663-681.

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

[6] Cohen W. M., Nelson, R.R., Walsh J. P. (2000) Protecting their intellectual assets: appropriability conditions and why U.S. manufacturing firms patent (or not), NBER Working Paper No. 7552.

[7] Décamps J-P, Mariotti T (2004) Investment timing and learning externalities, Journal of Economic Theory 118:80-102.

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

[9] Denicolò V (2000) Two-stage patent races and patent policy, RAND Journal of Economics 31(3):488-501.

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

[11] 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.

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

[13] Fudenberg D, Tirole J (1987) Understanding rent dissipation: On the use of game theory in industrial organization, American Economic Review 77(2):176-183.

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

[15] Gilbert R, Shapiro C (1990) Optimal patent length and breadth, RAND Journal of Economics 21(1):106-112.

[16] 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.

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

[18] Guey Chuen P, Huan-Yao L, Yee-Shin L, Kulkanya C (2011) Dengue vaccines: challenge and confrontation, World Journal of Vaccines 1:109-130.

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

[20] Henry E, Ruiz-Aliseda, F (2015) Keeping secrets: the economics of access deterrence, mimeo.

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

[22] Huisman K (2001) *Technology Investment: A Game Theoretic Real Options Approach* (Boston: Kluwer Academic Publishers).

[23] Katz M, Shapiro C (1987) R and D rivalry with licensing or imitation, American Economic Review 77(3):402-420.

[24] Lanjouw J (1998) The introduction of pharmaceutical product patents in India: 'Heartless exploitation of the poor and suffering'?, NBER Working Paper No. 6366.

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

[26] Mansfield E (1985) How rapidly does new industrial technology leakout?, Journal of Industrial Economics 34(2):217-223.

[27] Mansfield E, Schwartz M, Wagner S (1981) Imitation costs and patents: An empirical study, The Economic Journal 91(364):907-918.

[28] Mukherjee A, Pennings E (2004) Imitation, patent protection, and welfare, Oxford Economic Papers 56(4):715-733.

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

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

[31] 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.

[32] Steg J-H (2015) Symmetric equilibria in stochastic timing games, working paper 543, Center for Mathematical Economics, University of Bielefield.

[33] Thijssen J (2013) Game theoretic real options and competition risk, in: Bensoussan A, Peng S, Sung J (eds.), *Real Options, Ambiguity, Risk and Insurance* (Amsterdam: IOS Press).

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

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

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

[37] Weeds H (2002) Strategic delay in a real options model of R&D competition, Review of Economic Studies 69(3):729-747.

[38] World Health Organization (2011) Increasing access to vaccines through technology transfer and local production, 34 pp. (http://www.who.int/phi/publications/local_production_vaccines/en/, last accessed 15/7/ 2015).

# A Proofs

## A.1 Proof of Proposition 1

In this section we first identify and characterize the critical threshold $\widehat{K}$. We then study the innovator value function $L_y(Y_i, Y_F^*)$. Finally, we derive the equilibrium strategies in the attrition ($K < \widehat{K}$) and preemption regimes ($K \geq \widehat{K}$).
*Characterization of* $\widehat{K}$

**Proposition 11** *There exists a unique threshold that separates the attrition and preemption regimes of the investment timing game,*

$$\widehat{K} = \left(\frac{1 + \beta\left(\left(\pi_M/\pi_D\right) - 1\right)}{\left(\pi_M/\pi_D\right)^{\beta}}\right)^{1/(\beta-1)} I. \tag{13}$$

We first verify that $\widehat{K}$ is well defined. If $K = 0$, then $Y_F = 0$ so the follower's investment in stage 2 occurs immediately after innovation, $Y_F^* = Y_i$. In that case

$$L_y(Y_i, Y_F^*) = \left(\frac{\pi_D}{r-\alpha}Y_i - I\right)\left(\frac{y}{Y_i}\right)^{\beta} < \frac{\pi_D}{r-\alpha}Y_i\left(\frac{y}{Y_i}\right)^{\beta} = F_y(Y_F^*; 0) \tag{14}$$

for all $Y_i \geq y$. For any $K$, any increase in imitation cost shifts $L_y(Y_i, Y_F^*)$ upward since $Y_F^*$ is nondecreasing in $K$ and $\partial L_y(Y_i, Y_F^*)/\partial Y_F^* \geq 0$. Moreover any increase in imitation cost shifts $F_y(Y_F^*; K)$ downward since $Y_F^*$ maximizes $F_y(Y_i; K)$ and $\partial F_y(Y_i; K)/\partial K < 0$. At $Y_L$ and $Y_F$ therefore, $\partial L_y(Y_L, Y_F^*)/\partial K \geq 0$ and $\partial F_y(Y_F^*; K)/\partial K < 0$, with $\lim_{K\to\infty} F_y(Y_F^*; K) = 0$. Therefore, there exists a unique level of the imitation cost $\widehat{K}$ such that $L_y(Y_L^*, Y_F^*) = F_y\left(Y_F^*; \widehat{K}\right)$, $y \leq Y_L$. As $L_y(Y_S, Y_S) = F_y(Y_F^*; K) \Leftrightarrow K = I$ in which case $L_y(Y_L, Y_F^*) > F_y(Y_F^*; I)$, this threshold is given by the solution in $K$ to $L_{Y_L}(Y_L, Y_F) = F_{Y_L}(Y_F; K)$, and it is direct to verify the expression (13), as well as the property discussed Section 3.1 of the text, $\widehat{K} \leq I$ (see Section A.8.1 in the supplementary section for derivations).
*Characterization of* $L_y(Y_i, Y_F^*)$

We next study the function $L_y(Y_i, Y_F^*)$ for $y$ sufficiently low that firms initially delay investment. There are at most two local maxima, at $Y_L = \arg\max L_y(Y_i, Y_F)$ and $Y_S = \arg\max M_y(Y_i)$, with $Y_L \leq Y_S$. For $K = 0$, $Y_F < Y_L$ so $Y_F^* = \min\{Y_i, Y_F\}$ and $L_y(Y_i, Y_F^*) < M_y(Y_S)$. As argued above, any increase in imitation cost shifts $L_y(Y_i, Y_F^*)$ upward, whereas $M_y(Y_S)$ is unchanged. Therefore, there exists a unique level of the imitation cost $\widetilde{K}$ such that $L_y(Y_L, Y_F^*) = M_y(Y_S)$. This threshold is given by the solution in $K$ to $L_y(Y_L, Y_F) = M_y(Y_S)$, and it is direct to verify that

$$\widetilde{K} = \left( \frac{\beta\left((\pi_M/\pi_D) - 1\right)}{(\pi_M/\pi_D)^{\beta} - 1} \right)^{1/(\beta-1)} I. \tag{15}$$

Then $Y_L$ (resp. $Y_S$) is a unique global maximum of $L_y(Y_i, Y_F^*)$ if $K > \widetilde{K}$ (resp. $K < \widetilde{K}$).

Recall that $K_l := (\pi_D/\pi_M) I$ denotes the imitation cost such that $Y_L = Y_F$. Then the critical imitation cost levels that determine different equilibrium properties in the attrition range are ranked as follows:

**Proposition 12** *The imitation cost levels* $\left\{K_l, \widetilde{K}, \widehat{K}\right\}$ *satisfy* $K_l \leq \widetilde{K} \leq \widehat{K}$ *with strict inequalities if* $\pi_M > \pi_D$.

(see Section A.8.1 for derivation).

*Attrition equilibrium*

For $K < \widehat{K}$ we have $L_y(Y_i, Y_F^*) < F_y(Y_F^*; K)$ so firms engage in a war of attrition. Under the assumption of Markov strategies,any randomization that occurs is over investment triggers and it is with respect to these strategies that we derive the mixed strategy equilibrium.[14] There are two subcases to consider, $i)$ $K < \widetilde{K}$ and $ii)$ $\widetilde{K} \leq K < \widehat{K}$.

*i)* $K < \widetilde{K}$ *subcase*

If $K < \widetilde{K}$, we know from the characterization of $L_y$ above that $L_y(Y_i, Y_F^*)$ has a unique global maximum at $Y_S$ and decreases over $(Y_S, \infty)$. Any investment trigger in $[Y_L, Y_S)$ is thus dominated by investing at $Y_S$ or later as a follower *(see Figure 1)*. The choice of investment triggers over $[Y_S, \infty)$ constitutes a standard war of attrition (see Hendricks et al. [19]), hence there is a unique symmetric equilibrium in which firms randomize investment triggers continuously over this latter interval. To derive the unconditional (*i.e.* low initial $y$) equilibrium investment trigger distribution $G_0$, suppose that $y \leq Y_S$ and assume that firm $j \neq i$ randomizes her investment trigger.

[14] Steg [32] derives a more general equilibrium in which firms choose stopping times. A key difference is that the mixed strategy equilibrium over investment triggers that we derive here does not account explicitly for the fact that the process $Y_t$ exits the region over which attrition occurs with positive probability within any positive time increment. An equilibrium distribution over stopping times does, and results in the same distribution of investment outcomes due to the relatively simple payoff structure in our model, in which the second-mover advantage is global, *i.e.* $L_y\left(Y; Y_F^*\right) \leq F_y\left(Y_F^*; K\right)$ when $K \leq \widehat{K}$.

Then firm $i$'s expected payoff from investing at $Y_i$ is

$$\mathbb{E}_y V_y \left(Y_i, \widetilde{Y}_j\right) = \int_{Y_S}^{Y_i} F_y(s, K) g_j(s; K) ds + M_y\left(Y_i\right)\left(1 - G_j(Y_i; K)\right) \tag{16}$$

where $g_j$ and $G_j$ denote firm $j$'s density and distribution. Firm $i$ mixes over investment thresholds as well if $\partial \mathbb{E} V_y \left(Y_i, \widetilde{Y}_j\right) / \partial Y_i = 0$ over $(Y_S, \infty)$, so that the in equilibrium, the hazard rate of investment triggers is

$$h_0\left(Y_i; K\right) := \frac{g_0\left(Y_i; K\right)}{1 - G_0\left(Y_i; K\right)} = \frac{-M_y'(Y_i)}{F_y\left(Y_i; K\right) - M_y(Y_i)}, \tag{17}$$

the cumulative distribution being given by

$$G_0\left(Y_i; K\right) = 1 - \exp \int_{Y_S}^{Y_i} \frac{M_y'(s)}{F_y\left(s; K\right) - M_y(s)} ds, \tag{18}$$

and resulting in an expected payoff of $M_y\left(Y_S\right)$. Substituting for $F_y$ and $M_y$ and integrating gives the explicit form

$$G_0\left(Y_i; K\right) = 1 - \left(\frac{Y_i}{Y_S}\right)^{\beta \frac{I}{I-K}} \exp\left\{-\beta \frac{I}{I-K}\left(\frac{Y_i}{Y_S} - 1\right)\right\}. \tag{19}$$

*ii)* $\widetilde{K} \leq K < \widehat{K}$ *subcase*

If $\widetilde{K} \leq K < \widehat{K}$, we know from the characterization of $L_y$ above that $L_y\left(Y_i, Y_F^*\right)$ has a global maximum at $Y_L$ and a local maximum at $Y_S$. Because the leader payoff $L_y\left(Y_i, Y_F^*\right)$ is not monotonic over $[Y_L, Y_S]$ the attrition game is nonstandard. Let $Y_{S'}$ denote the unique solution in $[Y_L, Y_F]$ to the condition $L_y\left(Y_{S'}, Y_F\right) = M_y\left(Y_S\right)$. To verify that this threshold is well-defined, note that $Y_L \leq Y_F \leq Y_S$ since $K_l \leq \widetilde{K} \leq K < \widehat{K} \leq I$ and that $L_y\left(Y_i, Y_F^*\right)$ is continuous and weakly decreasing on $[Y_L, Y_F]$ *(see Figure 2)*. In this case, the support of mixed strategies is $(Y_L, Y_{S'}) \cup (Y_S, \infty)$.

To derive the unconditional equilibrium distribution $G\left(Y_i; K\right)$ note first that for $Y_i \geq Y_S$, the expected payoff of firm $i$ has the same form as (16) above, so that the hazard rate over $[Y_S, \infty)$ is $h_0\left(Y_i; K\right)$. For $Y_L \leq Y_i \leq Y_{S'}$ however, the expected payoff of firm $i$ is

$$\mathbb{E}_y V_y \left(Y_i, \widetilde{Y}_j\right) = F_y\left(Y_F; K\right) G\left(Y_i\right) + L_y\left(Y_i, Y_F\right)\left(1 - G(Y_i)\right). \tag{20}$$

Differentiating and rearranging gives the hazard rate over $[Y_L, Y_{S'}]$,

$$h\left(Y_i; K\right) = \frac{-\partial L_y(Y_i, Y_F)/\partial Y_i}{F_y\left(Y_F; K\right) - L_y(Y_i, Y_F)}. \tag{21}$$

The cumulative distribution over $[Y_L, Y_{S'}]$ is therefore given by

$$\overline{G}(Y_i;K) = 1-\exp\int_{Y_L}^{Y_i}\frac{\partial L_y(s,Y_F)/\partial s}{F_y(Y_F;K)-L_y(s,Y_F)}ds \tag{22}$$

$$= \frac{L_y(Y_L,Y_F)-L_y(Y_i,Y_F)}{F_y(Y_F;K)-L_y(Y_i,Y_F)}. \tag{23}$$

Substituting $Y_{S'}$ directly establishes the following intermediate result:

**Proposition 13** *For sufficiently low y the probability of a positive lag between innovation and imitation is*[15]

$$\overline{G}(Y_{S'};K) = \frac{\left((\pi_M/\pi_D)^\beta - 1\right) - \beta\left((\pi_M/\pi_D)-1\right)(I/K)^{\beta-1}}{(I/K)^{\beta-1}-1}. \tag{24}$$

Together the cumulative distribution (22) over $[Y_L, Y_{S'}]$ and the hazard rate $h_0(Y_i;K)$ over $[Y_S,\infty)$ define the equilibrium distribution as

$$G(Y_i;K) = \begin{cases} \overline{G}(Y_i;K) & \text{if } \; Y_L \le Y_i \le Y_{S'} \\ \overline{G}(Y_{S'};K) & \text{if } \; Y_{S'} < Y_i < Y_S \\ \overline{G}(Y_{S'};K) + \left(1-\overline{G}(Y_{S'};K)\right)G_0(Y_i;K) & \text{if } \; Y_S \le Y_i \end{cases} \tag{25}$$

which results in an expected payoff $L_y(Y_L, Y_F^*)$.

*Preemption equilibrium*

For $K >$ (resp. $=$)$\widehat{K}$, $L_{Y_L}(Y_L,Y_F^*) >$ (resp. $=$)$F_{Y_L}(Y_F^*;K)$ so there exists a unique $Y_P < Y_L$ (resp. $Y_P = Y_L$) such that $L_{Y_P}(Y_P,Y_F^*) = F_{Y_P}(Y_F^*;K)$. We refer to *preemption* when the inequalities are strict so $Y_P < Y_L$. Both firms seek to invest at $Y_P$, with equal probability of being an innovator or of effectively entering as an imitator at $Y_F$. The structure of the game and the arguments establishing equilibrium are those of a standard preemption game, although two additional points warrant mention.

If $K < I$, the equilibrium condition $L_y(Y_i,Y_F^*) = F_y(Y_F^*;K)$ has a root $Y_{P'} \in (Y_L, Y_F)$. In this case, in contrast with standard preemption games. If the market entry game were to start at $Y_t > Y_{P'}$, firms would play a war of attrition resulting in an expected payoff $L_y(Y_t,Y_F^*)$. As $\partial L_y(Y_{P'},Y_F^*)/\partial Y_i < 0$, if the game starts at a low enough threshold ($y \le Y_{P'}$) firms prefer to enter before $Y_{P'}$ and this subgame is never reached on the equilibrium path.

[15]Note that for $K = \widetilde{K}$, $Y_L = Y_{S'}$ and $\overline{G}\left(Y_{S'};\widetilde{K}\right) = 0$, whereas at the other extreme $\overline{G}\left(Y_{S'};\widehat{K}\right) = 1$.

Second, although simultaneous investment is generally not an equilibrium in the standard new market model of strategic investment, the suboptimality of simultaneous investment needs to be verified here because of the difference between leader and follower investment costs. Investment at the optimal simultaneous investment threshold $Y_S$ results in a payoff $M(Y_S)$ and evaluating,

$$\frac{L_y(Y_L, Y_F^*)}{M_y(Y_S)} = \left(\frac{\pi_M}{\pi_D}\right)^{\beta} - \beta\left(\frac{\pi_M}{\pi_D} - 1\right)\left(\frac{I}{K}\right)^{\beta-1}. \tag{26}$$

This ratio is increasing in $K$ and therefore over the preemption range for which simultaneous equilibrium might arise, it is minimized at $\widehat{K}$. Substituting $\widehat{K}$ for $K$ and simplifying gives $L_y(Y_L, Y_F^*)/M_y(Y_S) = \left(I/\widehat{K}\right)^{(\beta-1)} \geq 1$, with strict inequality if $\pi_M > \pi_D$. The best response to $Y_{-i} = Y_S$ is thus $Y_L$ for all $K \geq \widehat{K}$. Therefore firms seek to preempt one another before the simultaneous investment threshold is reached. □

## A.2 Proof of Corollary 3

To establish the corollary we characterize the effect of $\beta$ and $\pi_M/\pi_D$ on $\widehat{K}$. Evaluating the relevant partial derivatives and rearranging yields

$$\frac{\partial \widehat{K}}{\partial(\pi_M/\pi_D)} = -\beta\left(\frac{\pi_M}{\pi_D} - 1\right)\left(1 + \beta\left(\frac{\pi_M}{\pi_D} - 1\right)\right)^{\frac{2-\beta}{\beta-1}}\left(\frac{\pi_M}{\pi_D}\right)^{\frac{1-2\beta}{\beta-1}} I \tag{27}$$

so $\partial\widehat{K}/\partial(\pi_M/\pi_D) < 0$ directly,[16] whereas

$$\frac{\partial \widehat{K}}{\partial \beta} = \frac{-1}{(\beta-1)^2}\left(\ln\frac{1+\beta((\pi_M/\pi_D)-1)}{\pi_M/\pi_D} - \frac{(\beta-1)((\pi_M/\pi_D)-1)}{1+\beta((\pi_M/\pi_D)-1)}\right)\widehat{K}. \tag{28}$$

The sign of $\partial\widehat{K}/\partial\beta < 0$ is the opposite of that of the middle (bracketed) 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

$$\ln\frac{1+\beta((\pi_M/\pi_D)-1)}{\pi_M/\pi_D} > \frac{(\beta-1)((\pi_M/\pi_D)-1)}{1+\beta((\pi_M/\pi_D)-1)} \tag{29}$$

which is sufficient to conclude. □

[16] Note that since $\widehat{K}\Big|_{(\pi_M/\pi_D)=1} = I$ this establishes that $\widehat{K} \leq I$.

## A.3 Section 3.1 arguments and industry optimum (Proposition 4)

*Sensitivity analysis of investment thresholds*

Consider first the innovation threshold $Y_I$ (which takes values $Y_L$, or $Y_P$ under preemption, and $\min\left\{\widetilde{Y}_1, \widetilde{Y}_2\right\}$ under attrition). If $K < \widetilde{K}$ (or $K = \widetilde{K}$), the hazard rate over first entry thresholds implied by (17) is

$$h_0(Y_i; K) = \frac{\beta I}{I - K}\left(\frac{1}{Y_S} - \frac{1}{Y_i}\right), \tag{30}$$

so $\partial h / \partial K \geq 0$. For $\widetilde{K} < K < \widehat{K}$, the hazard rate corresponding to (25) is defined by parts. Over $[Y_L, Y_{S'})$ the hazard rate is

$$\overline{h}(Y_i; K) = \frac{-\partial L(Y_i, Y_F^*)/\partial Y_i}{F(Y_F^*; K) - L(Y_i, Y_F^*)} \tag{31}$$

where the numerator is independent of $K$, so $\partial \overline{h}/\partial K = -\left(\partial\left(F - L\right)/\partial K\right)\left(\partial L/\partial Y_i\right)/\left(F - L\right)^2 \geq 0$. Over $[Y_{S'}, \infty)$ we have $\partial h/\partial K = \partial h_0/\partial K \geq 0$. The hazard rate is discontinuous at $Y_{S'}$ and $Y_S$, but as $\partial Y_{S'}/\partial K \geq 0$ and $\partial Y_S/\partial K = 0$, it is increasing in $K$ over the entire range $[Y_S, \infty)$. Finally, for $K > \widehat{K}$, $Y_P$ decreases with $K$. Since the first entry threshold of each firm decreases with $K$, the minimum of these decreases as well. We have therefore established:

**Proposition 14** *In an attrition regime, the hazard rate over innovator entry thresholds increases with $K$ for all $K \leq \widehat{K}$.*

With respect to imitator investment, in the attrition regime the second entry threshold $Y_F^*$ decreases stochastically with respect to $K$ over $(Y_F, \infty)$ where follower entry is immediate, but increases deterministically otherwise. However, the expected difference between the first and second entry thresholds is monotone in $K$. For $K_l \leq K < \widehat{K}$, $Y_F^* - \widetilde{Y}_I = \max\left\{0, Y_F - \widetilde{Y}_I\right\}$ is distributed over $\{0\} \cup [Y_F - Y_{S'}, Y_F - Y_L]$ as $\Pr\{Y_F - Y_i = 0\} = G_\wedge(Y_S; K)$ and $\left(1 - G_\wedge(Y_S; K)\right)/\left(1 - G_\wedge(Y_F - Y_i; K)\right)$ otherwise. So by Proposition 14 the difference between the second and the first entry threshold increases with $K$ (stochastically in the attrition range and deterministically in the preemption range).

*Industry optimum*

The proposition follows directly from the equilibrium values with rent equalization, that is $\mathbb{E}V\left(\widetilde{Y}_1, \widetilde{Y}_2\right) = \min\left\{L\left(Y_L^*, Y_F^*\right), F\left(Y_F; K\right)\right\}$, and the sensitivity of $L$and $F$to

$K$. Note that for $K \leq \widetilde{K}\min\left\{L\left(Y_L^*, Y_F^*\right), F\left(Y_F; K\right)\right\} = M(Y_S)$is independent of $K$, and that at $K = \widehat{K}$, $M(Y_S) \leq L\left(Y_L, Y_F^*\right) = F\left(Y_F; \widehat{K}\right)$. Therefore, $\mathbb{E}V\left(\widetilde{Y}_1, \widetilde{Y}_2\right)$is constant over $\left[0, \widetilde{K}\right)$, increasing over $\left(\widetilde{K}, \widehat{K}\right)$, and decreasing over $\left(\widehat{K}, \infty\right)$. □

## A.4 Imitation cost, consumer surplus, and welfare (Proposition 5)

The argument is divided into four parts. We first characterize the optimal imitation cost level $K_P$in the closure of the preemption regime ($K \geq \widehat{K}$). Second, we establish that $\widetilde{K}$constitutes a lower bound for any optimal imitation cost in an attrition regime ($K_A \geq \widetilde{K}$). Third, we establish the existence of a local optimum of welfare under attrition ($\widetilde{K} \leq K_A < \widehat{K}$). Finally we compare the optimum under preemption with the optimal welfare that is attained in the attrition regime.

*Socially optimal imitation cost in preemption regime*

Suppose that $K \geq \widehat{K}$, so entry thresholds are $Y_P$and $Y_F$. The social welfare function (7) then has the form

$$W\left(K\right) = \left(\frac{\pi_M + \mathrm{CS}_M}{r - \alpha} Y_P - I\right)\left(\frac{Y_0}{Y_P}\right)^{\beta} + \left(\frac{\left(2\pi_D + \mathrm{CS}_D\right) - \left(\pi_M + \mathrm{CS}_M\right)}{r - \alpha} Y_F - K\right)\left(\frac{Y_0}{Y_F}\right)^{\beta}. \tag{32}$$

Noting that $Y_P$ and $Y_F$ are functions of $K$ with $Y_P \leq Y_L$ and $\lim_{K \to \widehat{K}} Y_P = Y_L$, and using the preemption equilibrium condition $L\left(Y_P, Y_F\right) = F\left(Y_F; K\right)$ which implicitly defines the ratio $\left(Y_F/Y_P\right)^{\beta}$, the derivative of (32) can be expressed as

$$\frac{dW}{dK} = \left(\frac{\mathrm{CS}_M}{\pi_M}\left(\beta\frac{\pi_M}{\pi_D}\frac{Y_L}{Y_L - Y_P} - (\beta - 1)\frac{Y_P}{Y_L - Y_P}\right) - \beta\frac{\mathrm{CS}_D}{\pi_D} - 2\right)\left(\frac{Y_0}{Y_F}\right)^{\beta}. \tag{33}$$

If $\mathrm{CS}_M = 0$ the $Y_L$ and $Y_P$ terms in (33) vanish and it is straightforward to see that $dW/dK < 0$, so that $\widehat{K}$ is a maximum. For $\mathrm{CS}_M > 0$, since $\lim_{K \to \widehat{K}} Y_P = Y_L$ (33) satisfies $\lim_{K \to \widehat{K}} dW/dK = +\infty$, and is strictly decreasing in $K$ over its range. So long as $\lim_{K \to \infty} dW/dK < 0$, there is a unique root $K_P > \widehat{K}$ that constitutes an interior optimum which occurs if

$$\left(\beta^2\frac{\pi_M}{\pi_D} - (\beta - 1)^2\right)\frac{\mathrm{CS}_M}{\pi_M} - \beta\frac{\mathrm{CS}_D}{\pi_D} - 2 < 0. \tag{34}$$

For notational simplicity, in what follows we let $K_P = \infty$ if (34) does not hold. Taken as a function of $\beta$ the left-hand side of (34) is a quadratic function, $\Delta\left(\beta\right)$, with $\Delta\left(1\right) = \left(\mathrm{CS}_M - \mathrm{CS}_D - 2\pi_D\right)/\pi_D < 0$ and $\lim_{\infty} \Delta\left(\beta\right) = \infty$. Therefore there exists a unique $\beta_0 > 1$ such that $\Delta\left(\beta_0\right) = 0$. Thus,

**Proposition 15** *The constrained optimization problem* $\max_{K \in [\widehat{K}, \infty]} W(K)$ *has a unique optimum* $K_P$*, and there exists a unique* $\beta_0 > 1$ *such that* $K_P$ *is finite if and only if* $\beta < \beta_0$*.*

For the proof of Proposition 7 in the next section it is also useful to derive the optimal value of welfare that is realized in the preemption range. Several steps (see Section A.8) establish that an optimum preemption threshold has the form $Y_P^* = \psi Y_L$ where

$$\psi = \begin{cases} \frac{\frac{\mathrm{CS}_D - \mathrm{CS}_M}{\pi_D} + \frac{2}{\beta}}{\frac{\mathrm{CS}_D}{\pi_D} - \frac{\beta-1}{\beta}\frac{\mathrm{CS}_M}{\pi_M} + \frac{2}{\beta}}, \beta < \beta_0 \\ \frac{\beta}{\beta-1}, \beta \geq \beta_0 \end{cases} \tag{35}$$

We have $\psi \in \left[\frac{\beta-1}{\beta}, 1\right]$, and from (34) $\psi = (\beta - 1)/\beta$ if $\beta \geq \beta_0$. Moreover, $\psi = 1$ if $\mathrm{CS}_M = 0$. The optimal preemption threshold is $Y_P^* = \psi Y_L$, so $Y_P^* \in [Y_{\mathrm{NPV}}, Y_L]$ where $Y_{\mathrm{NPV}} := (r - \alpha) I / \pi_M$ is the myopic Marshallian investment trigger. The optimal level of welfare under preemption can then be shown to be

$$W_P(K_P) = \frac{\mathrm{CS}_M}{\pi_M} \frac{\psi^{1-\beta}}{1-\psi} \frac{I}{\beta - 1} \left(\frac{Y_0}{Y_L}\right)^{\beta}. \tag{36}$$

*Lower bound on socially optimal imitation cost*

If $K < \widetilde{K}$ (first attrition subcase in Section A.1 above) so firms randomize investment triggers over $[Y_S, \infty)$ according to the distribution $G_0(Y_i; K)$ and imitator entry is immediate, then $W(K) < W\left(\widehat{K}\right)$. To see this, note first that by Proposition 4, industry value is lower at $K$ than at $\widehat{K}$, so it suffices to show that expected consumer surplus is lower also. But at $\widehat{K}$, innovator and imitator entry occur at the standalone thresholds $Y_L$ and $\widehat{Y}_F := \left(\beta(r - \alpha)\widehat{K}\right) / ((\beta - 1)\pi_D)$, whereas the lower bound of the entry threshold distribution under attrition is $Y_S = (\beta(r - \alpha) I) / ((\beta - 1)\pi_D) \geq \widehat{Y}_F$. Therefore, both investments occur later if $K < \widetilde{K}$ than they do at the critical imitation cost $\widehat{K}$ resulting in lower consumer surplus and hence in lower welfare.

*Existence of local maximum in attrition regime*

Consider the value of $W(K)$ just to the left of $\widehat{K}$. Since $V\left(\widetilde{Y}_1, \widetilde{Y}_2\right)$ is maximized at $\widehat{K}$, at this critical value the sign of $\lim_{K \to \widehat{K}_-} dW(K)/dK$ depends only on the behavior of the consumer surplus terms. For simplicity consider the third term, consumer surplus from imitation (the argument for the other term is similar). As noted in the text the consumer surplus from imitation is given by

$$\frac{\text{CS}_D - \text{CS}_M}{r-\alpha}\left[Y_F\right]^{-(\beta-1)} Y_0^{\beta}\left(\underbrace{G_{\wedge}\left(Y_{S'};K\right)}_{\text{lagged imitator entry}} + \underbrace{\int_{Y_S}^{\infty}\left(Y_F/s\right)^{\beta-1} dG_{\wedge}(s;K)}_{\text{immediate imitator entry}}\right). \quad (37)$$

To determine the value of the left derivative at $\widehat{K}$ of (37) recall that the distribution of entry thresholds is given by $G_{\wedge}\left(Y_i;K\right) = 1-\left(1-G\left(Y_i;K\right)\right)^2$. Consider the first summand in (37). Since $\overline{G}\left(Y_{S'};\widehat{K}\right) = 1$, $G_{\wedge}\left(Y_{S'};\widehat{K}\right) = 1$. Moreover $\partial G_{\wedge}/\partial s = 2\left(1-G\right)\left(\partial G/\partial s\right)$ so $\partial G_{\wedge}\left(Y_{S'};K\right)/\partial K|_{\widehat{K}} = 0$. Therefore in (37) only the direct effect of $K$ on $Y_F$ matters for welfare at $\widehat{K}$. A similar argument applies to the consumer surplus from innovation term in (7), except that there is no direct effect since $Y_L$ is independent of $K$.

Therefore,

$$\lim_{K\to\widehat{K}_-}\frac{d\mathbb{E}W\left(K\right)}{dK} = -\left(\beta-1\right)\frac{\text{CS}_D-\text{CS}_M}{r-\alpha}Y_F^{-\beta}Y_0^{\beta}\frac{\partial Y_F}{\partial K} \leq 0. \quad (38)$$

Since $W(K)$ is continuous, we conclude that if $\text{CS}_D > \text{CS}_M$, there exists a local optimum imitation cost level $K_A$ in $\left(\widetilde{K},\widehat{K}\right)$.

*Global welfare optimum*

We therefore know that for $\text{CS}_D > \text{CS}_M$, $\lim_{K\to\widehat{K}-} dW\left(K\right)/dK < 0$ and that for $\text{CS}_M > 0$, $\lim_{K\to\widehat{K}+} dW\left(K\right)/dK > 0$, so that for $\left(\text{CS}_D - \text{CS}_M\right)\text{CS}_M > 0$, welfare has local maxima in both the (upper) attrition and preemption ranges, whereas the local maximum under preemption is $K_P = \widehat{K}$ if $\text{CS}_M = 0$ and $K_A = \widehat{K}$ under attrition if $\text{CS}_D = \text{CS}_M$. Either type of local maximum can be a global maximum depending on the relative magnitude of the consumer surplus resulting from innovation and imitation. □

## A.5 Imitation cost, consumer surplus, and welfare con't (Proposition 7)

To establish the result, an upper bound is first derived for the level of welfare realized in the attrition regime and then compared with a lower bound of the welfare obtained under preemption. These bounds are tight only in the limit ($\beta = 1$), but have the advantage of resulting in a tractable analytic condition (see (48) below).

*Upper bound for welfare under attrition*

The optimal value of expected welfare under attrition can be bounded above as follows. Given the innovation threshold $\min\left\{\widetilde{Y}_1,\widetilde{Y}_2\right\}$ let $\widetilde{Y}_F = Y_F^*\left(\min\left\{\widetilde{Y}_1,\widetilde{Y}_2\right\};K\right)$

denote the (stochastic) imitation threshold for a given imitation cost $K$. The expected social welfare under attrition (7) is

$$\begin{aligned} W_A(K) = \mathbb{E}\left(\frac{\mathrm{CS}_M+\pi_M}{r-\alpha}\min\left\{\widetilde{Y}_1,\widetilde{Y}_2\right\}-I\right)\left(\frac{Y_0}{\min\left\{\widetilde{Y}_1,\widetilde{Y}_2\right\}}\right)^{\beta} \\ +\mathbb{E}\left(\frac{(\mathrm{CS}_D+2\pi_D)-(\mathrm{CS}_M+\pi_M)}{r-\alpha}\widetilde{Y}_F-K\right)\left(\frac{Y_0}{\widetilde{Y}_F}\right)^{\beta}. \quad (39) \end{aligned}$$

To bound the first term, note that its integrand is quasiconcave in the initial investment threshold and $\min\left\{\widetilde{Y}_1,\widetilde{Y}_2\right\} \geq Y_L \geq (\beta(r-\alpha)I)/((\beta-1)(\mathrm{CS}_M+\pi_M))$ where the rightmost term is the global maximizer. The first integrand in (39) is thus decreasing in investment threshold over the relevant range so

$$\begin{aligned} &\mathbb{E}\left(\frac{\mathrm{CS}_M+\pi_M}{r-\alpha}\min\left\{\widetilde{Y}_1,\widetilde{Y}_2\right\}-I\right)\left(\frac{Y_0}{\min\left\{\widetilde{Y}_1,\widetilde{Y}_2\right\}}\right)^{\beta} \\ &\leq \left(\frac{\mathrm{CS}_M+\pi_M}{r-\alpha}Y_L-I\right)\left(\frac{Y_0}{Y_L}\right)^{\beta} = \left(\beta\frac{\mathrm{CS}_M}{\pi_M}+1\right)\frac{I}{\beta-1}\left(\frac{Y_0}{Y_L}\right)^{\beta}. \quad (40) \end{aligned}$$

For the second term in (39), using the assumption that the static entry incentive is excessive,

$$\mathbb{E}\left(\frac{(\mathrm{CS}_D+2\pi_D)-(\mathrm{CS}_M+\pi_M)}{r-\alpha}\widetilde{Y}_F-K\right)\left(\frac{Y_0}{\widetilde{Y}_F}\right)^{\beta} \leq \mathbb{E}\left(\frac{\pi_D}{r-\alpha}\widetilde{Y}_F-K\right)\left(\frac{Y_0}{\widetilde{Y}_F}\right)^{\beta}. \quad (41)$$

The term on the right-hand side is simply the expected follower payoff in equilibrium, that is $\mathbb{E}F\left(Y_F^*\left(\widetilde{Y}_{-i};K\right);K\right) = \mathbb{E}V\left(\widetilde{Y}_1,\widetilde{Y}_2\right)$. Moreover, by Proposition 4, $\mathbb{E}V\left(\widetilde{Y}_1,\widetilde{Y}_2\right)$ is maximized for $K=\widehat{K}$. Therefore (41) holds if

$$\mathbb{E}\left(\frac{(\mathrm{CS}_D+2\pi_D)-(\mathrm{CS}_M+\pi_M)}{r-\alpha}\widetilde{Y}_F-K\right)\left(\frac{Y_0}{\widetilde{Y}_F}\right)^{\beta} \leq \frac{\widehat{K}}{\beta-1}\left(\frac{Y_0}{\widehat{Y}_F}\right)^{\beta}. \quad (42)$$

Then note that $\widehat{Y}_F = \left(\widehat{K}/I\right)(\pi_M/\pi_D)\,Y_L$ and substitute for $\left(\widehat{K}/I\right)^{1-\beta}$ (using (13)) to obtain the equivalent condition

$$\mathbb{E}\left(\frac{(\mathrm{CS}_D+2\pi_D)-(\mathrm{CS}_M+\pi_M)}{r-\alpha}\widetilde{Y}_F-K\right)\left(\frac{Y_0}{\widetilde{Y}_F}\right)^{\beta} \leq \frac{1}{1+\beta\left(\frac{\pi_M}{\pi_D}-1\right)}\frac{I}{\beta-1}\left(\frac{Y_0}{Y_L}\right)^{\beta}. \quad (43)$$

Combining (40) and (43) yields the upper bound

$$W_A(K) \leq \left(\beta\frac{\mathrm{CS}_M}{\pi_M} + 1 + \frac{1}{1+\beta\left(\frac{\pi_M}{\pi_D}-1\right)}\right)\frac{I}{\beta-1}\left(\frac{Y_0}{Y_L}\right)^{\beta}. \tag{44}$$

*Sufficient condition for preemption optimum to be global*

The optimal value of expected welfare under preemption is $W_P(K_P)$ (see Section A.8 for derivation and equation (36)):

$$W_P(K_P) = \frac{\mathrm{CS}_M}{\pi_M}\frac{\psi^{1-\beta}}{1-\psi}\frac{I}{\beta-1}\left(\frac{Y_0}{Y_L}\right)^{\beta}. \tag{45}$$

It is straightforward to check that taken as a function of $\psi$ over $(0,1)$, $\psi^{1-\beta}/(1-\psi)$ is strictly convex and minimized at $\psi_0 := (\beta-1)/\beta$. Substituting $\psi_0$ for $\psi$ in (45) and simplifying thus yields

$$W_P(\infty) = \frac{\mathrm{CS}_M}{\pi_M}I\left(\frac{\beta}{\beta-1}\right)^{\beta}\left(\frac{Y_0}{Y_L}\right)^{\beta} \leq W_P(K_P). \tag{46}$$

Therefore, a sufficient condition for the preemption optimum to be a global optimum of welfare is $W_A(K) \leq W_P(\infty)$. Combining (44) and (46) and simplifying the common $(I/(\beta-1))(Y_0/Y_L)^{\beta}$ terms yields the condition

$$\beta\frac{\mathrm{CS}_M}{\pi_M} + 1 + \frac{1}{1+\beta\left(\frac{\pi_M}{\pi_D}-1\right)} \leq \beta\frac{\mathrm{CS}_M}{\pi_M}\left(\frac{\beta}{\beta-1}\right)^{\beta-1}. \tag{47}$$

Rearranging and using $1/(1+\beta((\pi_M/\pi_D)-1)) \leq 1$, a sufficient condition for (47) to hold is

$$\frac{\mathrm{CS}_M}{\pi_M} \geq \frac{2}{\beta}\frac{1}{\left(\frac{\beta}{\beta-1}\right)^{\beta-1}-1} =: \Omega(\beta). \tag{48}$$

To characterize the right-hand side of (48), note first that using l'Hôpital's rule, $\lim_1\left(\frac{\beta}{\beta-1}\right)^{\beta-1} = 1$ and $\lim_\infty\left(\frac{\beta}{\beta-1}\right)^{\beta-1} = \mathrm{e}$, so $\Omega(1) = \infty$ and $\lim_\infty \Omega(\beta) = 0$. Moreover, $2/\beta$ is decreasing and $d\left(\frac{\beta}{\beta-1}\right)^{\beta-1}/d\beta = \left(\frac{\beta}{\beta-1}\right)^{\beta-1}\left(-\frac{1}{\beta} + \ln\frac{\beta}{\beta-1}\right)$ which is positive since $\ln(\beta/(\beta-1)) > 1/\beta$ by the logarithm inequality, so $\Omega(\beta)$ is decreasing over this range. □

## A.6 Endogenous entry barrier

In stage 3', the imitator payoff depends on the cost-raising effort $\rho$:

$$F(Y_i;K) = \left(\frac{\pi_D}{r-\alpha}Y_i - K_0 - f(\rho)\right)\left(\frac{Y_0}{Y_i}\right)^{\beta}. \tag{49}$$

The optimal standalone imitator threshold is $Y_F(\rho) = \beta(r-\alpha)(K_0 + f(\rho)) / ((\beta-1)\pi_D)$, yielding an optimal choice $Y_F^*(\rho) = \max\{Y_i, Y_F(\rho)\}$. In stage 2', an innovator having entered at the threshold $Y_i$ chooses a level of effort that maximizes:

$$L_{\mathrm{e}}(Y_i, \rho) = \left(\frac{\pi_M}{r-\alpha} Y_i - I_0 - \rho\right) + \frac{\pi_D - \pi_M}{r-\alpha} Y_F^*(\rho) \left(\frac{Y_i}{Y_F^*(\rho)}\right)^\beta . \tag{50}$$

Note that the term $Y_F^*(\rho)$ generally introduces a kink in the innovator's stage 2' payoff. For example with $K_0 \leq ((\beta-1)/\beta)(\pi_D/\pi_M) I_0$, $Y_F(0) \leq Y_{\mathrm{NPV}}$ so that $Y_F^*(\rho) = Y_i$ both in attrition and preemption regimes for some range of effort $\rho \in [0, \overline{\rho}]$. In such cases the innovator's stage 2' decision problem may present a corner solution. Moreover in an attrition regime, since the innovator threshold is random, the optimal endogenous entry barrier is itself a random variable in stage 1'. However, to determine the critical imitation cost $\widehat{K}_{\mathrm{e}}$ that separates the two regimes, it is sufficient to consider the case in which innovator entry occurs at the threshold at which there are no positional rents, *i.e.* $Y_{L,\mathrm{e}}(\rho^*) = \beta(r-\alpha)(I_0 + \rho^*) / ((\beta-1)\pi_M)$ where $\rho^*$ solves $\max_\rho L_{\mathrm{e}}(Y_{L,\mathrm{e}}, \rho)$ such that $L_{\mathrm{e}}(Y_{L,\mathrm{e}}, \rho^*) = F\left(Y_F; \widehat{K}_{\mathrm{e}}\right)$. Since at $K = \widehat{K}$, $\left(\frac{Y_t}{Y_{L,\mathrm{e}}(\rho^*)}\right)^\beta L_{\mathrm{e}}(Y_{L,\mathrm{e}}(\rho^*), \rho^*) \geq \left(\frac{Y_t}{Y_L}\right)^\beta L_{\mathrm{e}}(Y_{L,\mathrm{e}}(0), 0) = L(Y_L, Y_F) = F\left(Y_F; \widehat{K}\right)$, it immediately follows that $\widehat{K}_{\mathrm{e}} = K_0 + f(\rho^*) \leq \widehat{K}$. □

## A.7 Buyout and licensing

Depending on the effect of entry on industry profit, there are two cases to consider.

*Case i: efficiency effect* ($\pi_M/\pi_D \geq 2$)

Suppose that the innovator, at the time of investment, can offer a payment of $\varphi$ to buy its rival's option on duopoly profits. The innovator's decision in stage 2" in this case is $\max_{\varphi \geq F_0(Y_i)} L_{\mathrm{b}}(Y_i, \varphi)$ where

$$L_{\mathrm{b}}(Y_i, \varphi) := \left(\frac{\pi_M}{r-\alpha} Y_i - I - \varphi\right)\left(\frac{Y_0}{Y_i}\right)^\beta \tag{51}$$

and $\varphi \geq F_0(Y_i)$ is the rival firm's participation constraint. As imitator entry reduces industry flow profit, a takeover is always efficient for the firms and it is straightforward to verify that at an optimum $L_{\mathrm{b}}(Y_i, F_0(Y_i)) > L(Y_i, Y_F^*)$.

To establish that a buyout can increase welfare, consider the case where imitator entry would leave the consumer surplus unchanged, as occurs if $2\pi_D = \pi_M$ (*i.e.*, a unit demand or a cartel in a homogeneous product market). If $K \geq \widehat{K}$, preemption occurs, and industry value is pegged to $F_0(Y_t)$ regardless of whether takeovers are allowed or not. A buyout is efficient in this case if the first firm enters earlier. This occurs when the innovator can make a purchase offer to its rival, *i.e.* if the lower root of $L_{\mathrm{b}}\left(Y_i, F_0(Y_i)\right) = F\left(Y_F; K\right)$ is lower than $Y_P$, which holds since $L_{\mathrm{b}}\left(Y_i, F_0(Y_i)\right) > L\left(Y_i, Y_F^*\right)$.

To establish that attrition can be eliminated, consider the limiting case $K = 0$. In this case, follower entry is immediate for all $Y_i$, so $F_0(Y_i) = \pi_D Y_i / (r - \alpha)$ and the stage 1" leader payoff is therefore

$$L_{\mathrm{b}}\left(Y_i, F_0(Y_i)\right) := \left(\frac{\pi_M - \pi_D}{r - \alpha} Y_i - I\right)\left(\frac{Y_0}{Y_i}\right)^{\beta}. \tag{52}$$

Let $Y_{\mathrm{b}} := \beta\left(r - \alpha\right) I / \left(\left(\beta - 1\right)\left(\pi_M - \pi_D\right)\right)$ denote the maximizer of the latter function. Solving $L_{\mathrm{b}}(Y_{\mathrm{b}}, F_0(Y_{\mathrm{b}})) \geq F_0(Y_{\mathrm{b}})(Y_t/Y_{\mathrm{b}})^{\beta}$ gives the condition under which preemption arises even with a maximal second mover advantage ($K = 0$) as $\pi_M/\pi_D \geq \beta + 1$.

If a buyout is not possible then the innovator may license its technology to the imitator when it enters. The innovator's decision in stage 2" takes the form $\max_{\varphi \leq K_I} V_{\mathrm{l}}\left(\varphi\right)$ where

$$V_{\mathrm{l}}\left(\varphi\right) = \left(\frac{\pi_M - \pi_D}{r - \alpha} Y_F^*(\varphi) - \varphi\right)\left(\frac{Y_0}{Y_F^*(\varphi)}\right)^{\beta} \tag{53}$$

and the rival's participation constraint is $F\left(Y_F^*; K_0 + \varphi\right) \geq F_0(Y_F^*)$. In (53), $Y_F^*(\varphi)$ is the follower's investment threshold is generally a function of the fee $\varphi$ (if $\varphi < K_I$), although at an optimum $\varphi^* = K_I$ and $Y_F^*(\varphi) = Y_F^*$. In stage 1" then, the leader value is

$$L_{\mathrm{l}}\left(Y_i, Y_F^*\right) := \left(\frac{\pi_M}{r - \alpha} Y_i - I\right)\left(\frac{Y_0}{Y_i}\right)^{\beta} + \left(\frac{\pi_D - \pi_M}{r - \alpha} Y_F^* + K_I\right)\left(\frac{Y_0}{Y_F^*}\right)^{\beta} \tag{54}$$

so licensing simply has a level effect on the leader payoff if $Y_i < Y_F$. Setting $L_{\mathrm{l}}\left(Y_L, Y_F\right) = F\left(Y_F; K\right)$ defines the critical threshold $\widehat{K}_{\mathrm{l}} < \widehat{K}$ that separates the attrition and preemption regimes. To establish the effect of licensing on welfare, there are three cases to consider: 1) If $K \geq \widehat{K}$, the industry is preemptive whether licensing occurs or not. Industry value and the timing of imitation are then unaffected by licensing, whereas the preemption threshold decreases since $L_{\mathrm{l}}\left(Y_i, Y_F^*\right) > L\left(Y_i, Y_F^*\right)$ so innovation occurs earlier and welfare increases. 2) Alternatively, if $K \in \left(\widehat{K}_{\mathrm{l}}, \widehat{K}\right)$, then the industry switches from an attrition regime to preemption when licensing is allowed. As compared with the previous case, the increase in welfare is also due to an increase in industry value and earlier imitation. 3) Finally, if $K \leq \widehat{K}_{\mathrm{l}}$, the industry is in an attrition regime whether licensing occurs or not. Industry value is pegged on the optimal leader value, which increases in comparison to the baseline model, and the imitation is either unaffected (if $\widetilde{Y}_I \leq Y_F$) or occurs earlier if innovation occurs earlier. What remains to be verified is that the distribution of innovation thresholds shifts left with licensing. We do this in the case that $K$ is not too small, $\widetilde{K}_l < K \leq \widehat{K}_l$, (the argument for $K \leq \widetilde{K}_l$ is similar).

Note first that the support of the mixed strategy distribution, $[Y_L, Y_{S',l}] \cup [Y_{S,l}, \infty)$, is larger with licensing. $Y_L$ is unaffected by licensing, whereas $Y_{S,l} = \beta (r - \alpha)(I - K_I) / ((\beta - 1)\pi_D) < Y_S$. Finally, for $Y_i < Y_F$ licensing shifts $L(Y_i, Y_F)$ upward by $K_I (Y_t/Y_F)^{\beta}$, which is weakly larger than the upward shifts of $M(Y_i)$ $(K_I (Y_t/Y_i)^{\beta})$, so $Y_{S',l} > Y_{S'}$ (*see the graphic construction of* $Y_{S'}$ *in Figure 2*).

Next, it is necessary to examine the impact of licensing on the hazard rate of $\widetilde{Y}_I$. Over $[Y_S, \infty)$, the hazard rate implied by (18), adapted to the licensing specification, becomes

$$h_{0,l}(Y_i; K_0 + \varphi^*) = \beta \frac{I - K_I}{I - K_I - K} \left( \frac{1}{Y_{S,l}} - \frac{1}{Y_i} \right). \tag{55}$$

Comparing with $h_0$ in (30), we find

$$\frac{h_{0,l}(Y_i; K_0 + \varphi^*)}{h_0(Y_i; K)} = \frac{(I - K_I)(I - K)}{I(I - K_I - K)} \frac{\frac{1}{Y_{S,l}} - \frac{1}{Y_i}}{\frac{1}{Y_S} - \frac{1}{Y_i}} > 1. \tag{56}$$

Over $[Y_L, Y_{S'})$, the hazard rate implied by (22), adapted to licensing, becomes

$$\overline{h}_l(Y_i; K_0) = \frac{-\partial L_l(Y_i)/\partial Y_i}{F(Y_F^*; K) - L_l(Y_i)} \tag{57}$$

and since the slope of $L_l$ is independent of $\varphi$ so that licensing only has a positive level effect, $\overline{h}_l(Y_i; K_0) > \overline{h}(Y_i; K)$.

*Case ii: product market complementarity* ($\pi_M/\pi_D < 2$)

If the second firm's entry increases industry profit, there is an optimal imitator entry threshold for the industry $Y_F^{**} := \beta (r - \alpha) K_0 / (\beta - 1)(2\pi_D - \pi_M)$, which may be either greater or smaller than $Y_F$ as noted in the text. With a simple flat license fee instrument an innovator cannot induce imitation beyond $Y_F$ (it could with a forcing contract or a combination of a flat fee and a royalty payment but we do not pursue this further here) so the most interesting case to consider is if imitation occurs too late from an industry standpoint and the innovator has some leeway regarding imitator entry *i.e.* $Y_i, Y_F^{**} < Y_F$. (Otherwise, the optimal license fee is $K_I$, imitation occurs at $Y_F^*$, and the outcome is comparable to the previous case.) If these inequalities do hold, then in stage 2", the innovator's problem is $\max_{\varphi \leq K_I} V_l(\varphi)$ where

$$V_l(\varphi) = \left( \varphi - \frac{\pi_M - \pi_D}{r - \alpha} Y_F^{**} \right) \left( \frac{Y_0}{Y_F^{**}} \right)^{\beta} \tag{58}$$

and the follower's participation constraint is $F\left(Y_F^{**}; K_0+\varphi\right) \geq F_0(Y_F^{**})$. An optimal license fee satisfies this constraint with equality, i.e. at the optimal imitation threshold $Y_F^{**}$,

$$\frac{\pi_D}{r-\alpha}Y_F^{**} - K_0 - \varphi^* = \left(\frac{\pi_D}{r-\alpha}Y_F - K_0 - K_I\right)\left(\frac{Y_F^{**}}{Y_F}\right)^{\beta} \tag{59}$$

whereas if $\varphi = K_I$, as $Y_F^{**} < Y_F$ and since $Y_F$ is a maximizer,

$$\frac{\pi_D}{r-\alpha}Y_F^{**} - K_0 - \varphi < \left(\frac{\pi_D}{r-\alpha}Y_F - K_0 - K_I\right)\left(\frac{Y_F^{**}}{Y_F}\right)^{\beta}$$

so $\varphi^* < K_I$. $\square$

## A.8 Additional derivations

This section details some of the intermediate steps and lengthier derivations referred to above.

### A.8.1 Derivation of $\widehat{K}$, $\widetilde{K}$, and $K_l \leq \widetilde{K} \leq \widehat{K}$ ranking

To find $\widehat{K}$, set $L_{Y_L}\left(Y_L, Y_F\right) = F_{Y_L}\left(Y_F; \widehat{K}\right)$ *i.e.*

$$\left(\frac{\pi_M}{r-\alpha}Y_L - I\right)\left(\frac{Y_t}{Y_L}\right)^{\beta} + \frac{\pi_D - \pi_M}{r-\alpha}Y_F\left(\frac{Y_t}{Y_F}\right)^{\beta} = \left(\frac{\pi_D}{r-\alpha}Y_F - \widehat{K}\right)\left(\frac{Y_t}{Y_F}\right)^{\beta} \tag{60}$$

or, substituting for $Y_L$ and $\widehat{Y}_F = \beta\left(r-\alpha\right)\widehat{K}/\left(\beta-1\right)\pi_D$ (at $K = \widehat{K}$)

$$\frac{I}{\beta-1}\left(\frac{Y_t}{Y_L}\right)^{\beta} + \frac{\beta}{\beta-1}\left(1-\frac{\pi_M}{\pi_D}\right)\widehat{K}\left(\frac{Y_t}{\widehat{Y}_F}\right)^{\beta} = \frac{\widehat{K}}{\beta-1}\left(\frac{Y_t}{\widehat{Y}_F}\right)^{\beta}. \tag{61}$$

Then multiply by $\left(\widehat{Y}_F/Y_t\right)^{\beta}$ and note that $\widehat{Y}_F/Y_L = \left(\pi_M/\pi_D\right)\left(\widehat{K}/I\right)$ to get

$$\frac{I}{\beta-1}\left(\frac{\pi_M}{\pi_D}\right)^{\beta}\left(\frac{\widehat{K}}{I}\right)^{\beta} + \frac{\beta}{\beta-1}\left(1-\frac{\pi_M}{\pi_D}\right)\widehat{K} = \frac{\widehat{K}}{\beta-1}. \tag{62}$$

Multiplying by $\left(\beta-1\right)/\widehat{K}$ and regrouping terms,

$$\left(\frac{\pi_M}{\pi_D}\right)^{\beta}\left(\frac{\widehat{K}}{I}\right)^{\beta-1} = 1+\beta\left(\frac{\pi_M}{\pi_D}-1\right) \tag{63}$$

so

$$\widehat{K} = \left(\left(1+\beta\left(\left(\pi_M/\pi_D\right)-1\right)\right)/\left(\pi_M/\pi_D\right)^{\beta}\right)^{1/(\beta-1)} I. \tag{64}$$

To find $\widetilde{K}$, set $L_{Y_L}(Y_L, Y_F) = M_{Y_L}(Y_S)$ *i.e.*

$$\left(\frac{\pi_M}{r-\alpha}Y_L - I\right)\left(\frac{Y_t}{Y_L}\right)^\beta + \frac{\pi_D - \pi_M}{r-\alpha}Y_F\left(\frac{Y_t}{Y_F}\right)^\beta = \left(\frac{\pi_D}{r-\alpha}Y_S - I\right)\left(\frac{Y_t}{Y_S}\right)^\beta. \quad (65)$$

Substituting for $Y_L$, $Y_F$, and $Y_S$ gives

$$\frac{I}{\beta-1}\left(\frac{Y_t}{Y_L}\right)^\beta + \frac{\beta}{\beta-1}\left(1-\frac{\pi_M}{\pi_D}\right)\widetilde{K}\left(\frac{Y_t}{Y_F}\right)^\beta = \frac{I}{\beta-1}\left(\frac{Y_t}{Y_S}\right)^\beta. \quad (66)$$

Multiply by $(Y_F/Y_t)^\beta$ and note that $Y_F/Y_L = (\pi_M/\pi_D)\left(\widetilde{K}/I\right)$ and $Y_F/Y_S = \widetilde{K}/I$ to get

$$\frac{I}{\beta-1}\left(\frac{\pi_M}{\pi_D}\right)^\beta\left(\frac{\widetilde{K}}{I}\right)^\beta + \frac{\beta}{\beta-1}\left(1-\frac{\pi_M}{\pi_D}\right)\widetilde{K} = \frac{I}{\beta-1}\left(\frac{\widetilde{K}}{I}\right)^\beta. \quad (67)$$

Regrouping terms on either side,

$$\frac{I}{\beta-1}\left(\left(\frac{\pi_M}{\pi_D}\right)^\beta - 1\right)\left(\frac{\widetilde{K}}{I}\right)^\beta = \frac{\beta}{\beta-1}\left(\frac{\pi_M}{\pi_D} - 1\right)\widetilde{K} \quad (68)$$

and multiplying by $(\beta-1)/\widetilde{K}$,

$$\left(\left(\frac{\pi_M}{\pi_D}\right)^\beta - 1\right)\left(\frac{\widetilde{K}}{I}\right)^{\beta-1} = \beta\left(\frac{\pi_M}{\pi_D} - 1\right) \quad (69)$$

so

$$\widetilde{K} = \left(\beta\left((\pi_M/\pi_D) - 1\right)/\left((\pi_M/\pi_D)^\beta - 1\right)\right)^{1/(\beta-1)} I. \quad (70)$$

The different critical imitation cost levels are ranked as $K_l \leq \widetilde{K} \leq \widehat{K}$, with strict inequalities if $\pi_M > \pi_D$. Indeed, straightforward calculations show that $\widetilde{K} \geq K_l$ if and only if

$$(\beta-1)(\pi_M/\pi_D)^\beta - \beta(\pi_M/\pi_D)^{\beta-1} + 1 \geq 0, \quad (71)$$

and that $\widehat{K} \geq \widetilde{K}$ if and only if

$$(\pi_M/\pi_D)^\beta - \beta\left((\pi_M/\pi_D) - 1\right) - 1 \geq 0. \quad (72)$$

Both of these conditions hold for all $\beta, \pi_M/\pi_D \geq 1$ (it suffices to evaluate them at $\pi_M/\pi_D = 1$ and to observe that the derivative with respect to $\pi_M/\pi_D$ is non-negative).

### A.8.2 Welfare under preemption

*Characterization of* $Y_P(K)$

Over $\left(\widehat{K},\infty\right)$ the condition $L(Y_P, Y_F) = F(Y_F; K)$ implicitly defines the preemption threshold $Y_P$ as a $\mathcal{C}^1$ function of $K$ (see Section A.1):

$$\left(\frac{\pi_M}{r-\alpha}Y_P - I\right)\left(\frac{Y_0}{Y_P}\right)^\beta + \frac{\pi_D - \pi_M}{r-\alpha}Y_F\left(\frac{Y_0}{Y_F}\right)^\beta = \left(\frac{\pi_D}{r-\alpha}Y_F - K\right)\left(\frac{Y_0}{Y_F}\right)^\beta. \quad (73)$$

Dividing by $Y_0$ and moving $Y_F$ terms to the right-hand side gives

$$\left(\frac{\pi_M}{r-\alpha}Y_P - I\right)Y_P^{-\beta} = \left(\frac{\pi_M}{r-\alpha}Y_F - K\right)Y_F^{-\beta} \quad (74)$$

or, substituting $(\beta(r-\alpha)K)/((\beta-1)\pi_D)$ for $Y_F$ and factoring $K^{1-\beta}$,

$$\left(\frac{\pi_M}{r-\alpha}Y_P - I\right)Y_P^{-\beta} = \left(\frac{\beta}{\beta-1}\frac{\pi_M}{\pi_D} - 1\right)\left(\frac{\beta}{\beta-1}\frac{r-\alpha}{\pi_D}\right)^{-\beta}K^{1-\beta}. \quad (75)$$

The condition (75) has the form $f(Y_P) = g(K)$ and thus defines $dY_P/dK = g'(K)/f'(Y_P)$ where

$$f'(Y_P) = \left(-(\beta-1)\frac{\pi_M}{r-\alpha}Y_P + \beta I\right)Y_P^{-\beta-1} \quad (76)$$

and

$$g'(K) = -(\beta-1)\left(\frac{\beta}{\beta-1}\frac{\pi_M}{\pi_D} - 1\right)\left(\frac{\beta}{\beta-1}\frac{r-\alpha}{\pi_D}\right)^{-\beta}K^{-\beta} = -\frac{\beta-1}{K}g(K). \quad (77)$$

The sign $g'(K) < 0$ is direct whereas for any preemption threshold $Y_P$, $Y_P < Y_L$, and therefore $f'(Y_P) > 0$. Finally note that from (76) and (77), using the identity $f(Y_P) = g(K)$ and simplifying the numerator and the denominator by $Y_P^{-\beta}$,

$$\frac{dY_P}{dK} = -\frac{\beta-1}{K}\frac{\frac{\pi_M}{r-\alpha}Y_P - I}{-(\beta-1)\frac{\pi_M}{r-\alpha} + \beta(I/Y_P)}. \quad (78)$$

*Interior preemption optimum* $K_P$

Suppose that condition (34) holds so that the preemption optimum is interior. In a preemption equilibrium innovator and imitator entry occur at $Y_P$ and $Y_F$ so social welfare is

$$W(K) = \left(\frac{\mathrm{CS}_M + \pi_M}{r-\alpha}Y_P - I\right)\left(\frac{Y_0}{Y_P}\right)^\beta + \left(\frac{(\mathrm{CS}_D + 2\pi_D) - (\mathrm{CS}_M + \pi_M)}{r-\alpha}Y_F - K\right)\left(\frac{Y_0}{Y_F}\right)^\beta. \quad (79)$$

Substituting for $Y_F$ in the second term and factoring $K$,

$$W(K) = \left( \frac{\mathrm{CS}_M + \pi_M}{r-\alpha} Y_P - I \right) \left( \frac{Y_0}{Y_P} \right)^{\beta}$$
$$+ \left( \frac{\beta}{\beta-1} \frac{(\mathrm{CS}_D + 2\pi_D) - (\mathrm{CS}_M + \pi_M)}{\pi_D} - 1 \right) \left( \frac{\beta}{\beta-1} \frac{r-\alpha}{\pi_D} \right)^{-\beta} Y_0^{\beta} K^{1-\beta}. \quad (80)$$

In a constrained social optimum the planner's problem over the preemption range is $\max_{K \geq \widehat{K}} W(K)$. The derivative of (80) is

$$W'(K) = \left( -(\beta-1) \frac{\mathrm{CS}_M + \pi_M}{r-\alpha} Y_P + \beta I \right) \left( \frac{Y_0}{Y_P} \right)^{\beta} \frac{1}{Y_P} \frac{dY_P}{dK}$$
$$-(\beta-1) \left( \frac{\beta}{\beta-1} \frac{(\mathrm{CS}_D + 2\pi_D) - (\mathrm{CS}_M + \pi_M)}{\pi_D} - 1 \right) \left( \frac{\beta}{\beta-1} \frac{r-\alpha}{\pi_D} \right)^{-\beta} Y_0^{\beta} K^{-\beta}.$$

At an interior optimum the socially optimal imitation cost $K_P$ satisfies the first-order condition $W_P'(K_P) = 0$, but it is more convenient to obtain an expression for the corresponding socially optimal preemption threshold $Y_P^*$ from the first-order condition. Substituting for $dY_P/dK$ (expression (78)) in the first-order condition and multiplying by $K/((\beta-1)Y_0)$ gives

$$-\frac{\left( -(\beta-1) \frac{\mathrm{CS}_M + \pi_M}{r-\alpha} Y_P^* + \beta I \right) \left( \frac{\pi_M}{r-\alpha} Y_P^* - I \right)}{-(\beta-1) \frac{\pi_M}{r-\alpha} Y_P^* + \beta I} Y_P^{*-\beta}$$
$$-\left( \frac{\beta}{\beta-1} \frac{(\mathrm{CS}_D + 2\pi_D) - (\mathrm{CS}_M + \pi_M)}{\pi_D} - 1 \right) \left( \frac{\beta}{\beta-1} \frac{r-\alpha}{\pi_D} \right)^{-\beta} K^{*1-\beta} = 0. \quad (82)$$

From the preemption condition (75),

$$\left( \frac{\beta}{\beta-1} \frac{r-\alpha}{\pi_D} \right)^{-\beta} K^{1-\beta} = \frac{\left( \frac{\pi_M}{r-\alpha} Y_P^* - I \right) Y_P^{*-\beta}}{\frac{\beta}{\beta-1} \frac{\pi_M}{\pi_D} - 1}. \quad (83)$$

Substituting into the second term in (82), cancelling $\left( \frac{\pi_M}{r-\alpha} Y_P^* - I \right) Y_P^{*-\beta}$ terms which appear in both parts, and rearranging yields an equivalent condition in terms of $Y_P^*$ only,

$$\frac{(\beta-1) \frac{\mathrm{CS}_M + \pi_M}{r-\alpha} Y_P^* - \beta I}{-(\beta-1) \frac{\pi_M}{r-\alpha} Y_P^* + \beta I} = \frac{\frac{\beta}{\beta-1} \frac{(\mathrm{CS}_D + 2\pi_D) - (\mathrm{CS}_M + \pi_M)}{\pi_D} - 1}{\frac{\beta}{\beta-1} \frac{\pi_M}{\pi_D} - 1}. \quad (84)$$

There is a unique solution to (84) which can be expressed as $Y_P^* = \psi Y_L = \psi (\beta (r-\alpha) I) / ((\beta-1)\pi_M)$, in which case the numerator and denominator of the left hand side simplify yielding, after rearrangement of the right-hand side also,

$$\frac{\frac{\mathrm{CS}_M + \pi_M}{\pi_M} \psi - 1}{1-\psi} = \frac{\frac{(\mathrm{CS}_D + 2\pi_D) - (\mathrm{CS}_M + \pi_M)}{\pi_D} - \frac{\beta-1}{\beta}}{\frac{\pi_M}{\pi_D} - \frac{\beta-1}{\beta}} \quad (85)$$

and it is straightforward to check that the unique solution is

$$\psi = \frac{\frac{\mathrm{CS}_D - \mathrm{CS}_M}{\pi_D} + \frac{2}{\beta}}{\frac{\mathrm{CS}_D}{\pi_D} - \frac{\beta-1}{\beta}\frac{\mathrm{CS}_M}{\pi_M} + \frac{2}{\beta}}. \tag{86}$$

Note that setting $Y_P^* > Y_{\mathrm{NPV}}$ is equivalent to setting $\psi > (\beta - 1)/\beta$ and yields condition (34) in the text.

It is now possible to return to the social welfare expression (80) and obtain an explicit form for the value of social welfare at the optimum. First, the identity (83) can be used to substitute terms in the second summand of $W(K_P)$ so as to obtain an expression in terms of $Y_P^*$ only,

$$W_P(Y_P^*) = \left(\frac{\mathrm{CS}_M + \pi_M}{r-\alpha} Y_P^* - I\right)\left(\frac{Y_0}{Y_P^*}\right)^\beta + \left(\frac{\beta}{\beta-1}\frac{(\mathrm{CS}_D + 2\pi_D) - (\mathrm{CS}_M + \pi_M)}{\pi_D} - 1\right)\frac{\frac{\pi_M}{r-\alpha}Y_P^* - I}{\frac{\beta}{\beta-1}\frac{\pi_M}{\pi_D} - 1}\left(\frac{Y_0}{Y_P^*}\right)^\beta. \tag{87}$$

Regrouping terms

$$W_P(Y_P^*) = \frac{\left(\frac{\mathrm{CS}_M+\pi_M}{r-\alpha}Y_P^* - I\right)\left(\frac{\beta}{\beta-1}\frac{\pi_M}{\pi_D} - 1\right)}{\frac{\beta}{\beta-1}\frac{\pi_M}{\pi_D} - 1}\left(\frac{Y_0}{Y_P^*}\right)^\beta \tag{88}$$

$$+\frac{\left(\frac{\beta}{\beta-1}\frac{(\mathrm{CS}_D+2\pi_D)-(\mathrm{CS}_M+\pi_M)}{\pi_D} - 1\right)\left(\frac{\pi_M}{r-\alpha}Y_P^* - I\right)}{\frac{\beta}{\beta-1}\frac{\pi_M}{\pi_D} - 1}\left(\frac{Y_0}{Y_P^*}\right)^\beta. \tag{89}$$

Substituting for $Y_P^* = (\beta(r-\alpha)\psi I)/((\beta-1)\pi_M)$ $(= \psi Y_L)$ and factoring $I$,

$$W_P(Y_P^*) = \frac{\left(\frac{\beta}{\beta-1}\frac{\mathrm{CS}_M+\pi_M}{\pi_M}\psi - 1\right)\left(\frac{\beta}{\beta-1}\frac{\pi_M}{\pi_D} - 1\right)}{\frac{\beta}{\beta-1}\frac{\pi_M}{\pi_D} - 1} + \frac{\left(\frac{\beta}{\beta-1}\frac{(\mathrm{CS}_D+2\pi_D)-(\mathrm{CS}_M+\pi_M)}{\pi_D} - 1\right)\left(\frac{\beta}{\beta-1}\psi - 1\right)}{\frac{\beta}{\beta-1}\frac{\pi_M}{\pi_D} - 1}\psi^{-\beta} I\left(\frac{Y_0}{Y_L}\right)^\beta. \tag{90}$$

It is straightforward to check that after substituting the expression for $\psi$ given by (86) and some algebra,

$$W_P(Y_P^*) = \frac{\mathrm{CS}_M}{\pi_M}\frac{\psi^{1-\beta}}{1-\psi}\frac{I}{\beta-1}\left(\frac{Y_0}{Y_L}\right)^\beta. \tag{91}$$

□