The Effect of Mergers on Innovation*

Kaustav Das[†] Tatiana Mayskaya[‡] Arina Nikandrova[§]

5 February 2024

Abstract

We study the effect of a merger on R&D activity in a dynamic model in which there is uncertainty about the feasibility of innovation. The merger has three effects on innovation. Once an innovation has already taken place, the merger may reduce the number of follow-up innovations (cannibalization effect). At the same time, the merger may increase the probability of the first game-changer innovation (appropriability effect) and bring this innovation forward in time (informational effect). A surprising policy implication of our model is that the benefit of the merger may be higher if the first and subsequent innovations are closer substitutes.

Keywords: R&D, mergers, Poisson process, strategic experimentation, learning

JEL Codes: C73, D83

1 Introduction

In its recent review of the merger between chemical companies Dow and DuPont, the European Commission (EC) applied a novel innovation-based theory of harm. The concern was that after the merger, the parties would find it profitable to reduce overall R&D

*We sincerely thank Rabah Amir, Ying Chen, Brendan Daley, Andrey Golubov, Sergei Izmalkov, Ali Khan, Vijay Krishna, Igor Letina, Romans Pancs; as well as the participants at the SAET conference 2019 Ischia, Lancaster Game Theory Conference 2019, the 6th World Congress of the Game Theory Society (GAMES) 2021, CMA Durham Workshop 2023, the PET 2023 conference Santiago; and seminar audiences of the University of Birmingham; Birkbeck; City, University of London; NERA Economic Consulting; University of Nottingham; Johns Hopkins University; Pennsylvania State University; HSE University; University of Warwick and University of Leicester for their helpful comments. Remaining errors are ours.

†Department of Economics, Finance and Accounting, University of Leicester, School of Business, UK, e-mail: daskaustav84@gmail.com.

‡International College of Economics and Finance / Faculty of Economic Sciences, HSE University, Moscow, Russia, e-mail: tmayskaya@gmail.com.

§Department of Economics, City, University of London, UK, e-mail: a.nikandrova@gmail.com.

investments, thereby decreasing the availability of yet undiscovered pesticide products in the future. The Commission argued that prior to the merger, each party had incentives to develop new products to capture sales from its competitors. After the merger, however, by introducing new products, instead of capturing sales from its competitors, the merged entity might divert sales from its own products. In other words, an innovation by one merging party now cannibalizes its merger partner's profits. The merger internalizes the reduction of profits that the competition between newly developed products brings. As a result, incentives to innovate are reduced. The described mechanism is commonly referred to as the cannibalization effect[1] and is very much rooted in the traditional horizontal merger analysis.[2]

Although heavily relied on by the Commission, the cannibalization effect does not fully capture all the consequences of a merger for innovation incentives. The existing literature recognizes that innovation incentives also depend on the extent to which an innovator can capture the benefits generated by its innovation efforts. A firm's incentives to innovate are muted if future competition from similar products is expected to rapidly diminish the profit that the innovation generates. A merger may alleviate this concern by giving the merged entity control over the introduction of new products in the future. The increased incentives to innovate due to the improved ability of the merged entity to appropriate the benefits of its innovation efforts are often referred to as the appropriability effect of the merger.[3,4]

Though both the cannibalization and the appropriability effects have been recognized in the existing literature, the trade-off between them is not well understood. We propose a unifying formal framework that encompasses existing verbal arguments (e.g., Shapiro (2012)) and derives some novel insights not easily accessible without a formal model.

We consider two firms engaging in costly R&D activity along a research avenue. The research avenue can be one of two types: good or bad. A good research avenue rewards R&D effort with a product innovation at exponentially distributed time — in the manner

[1]Though not explicitly called so, the cannibalization effect was first introduced by Arrow (1962). Tirole (1988) refers to this effect as the replacement effect because, by introducing new products, the merged entity would replace only itself, while absent the merger, each party would replace its competitor.

[2]The very same idea that a merger internalizes the negative externality that one product imposes on the profits from another product is typically used to explain a merged entity's incentives to increase prices. The Commission's Dow/DuPont decision refers to the EC's Horizontal Merger Guidelines on numerous occasions.

[3]The appropriability effect was first described in Schumpeter (1942).

[4]In the Dow/DuPont decision, the EC equates appropriability to the intellectual property rights protection (paragraph 5, Annex 4 of the decision) and claims that the appropriability is not affected by competition because an innovator can always capture the marginal social value of an innovation (Appendix A to Annex 4 of the decision). This view, however, fails to recognize that the marginal social value of an innovation may decline with the number of similar products in existence.

of exponential bandits. A bad research avenue never generates an innovation. Initially, the firms do not know the type of the research avenue. Over time, each firm learns about the type of the avenue from its own and its rival's observable research activity and observable innovation successes.[5] We assume that there can be, at most, two sequential innovations in the market. The innovations are substitutes, and so, the introduction of the second innovation reduces the profit that the first innovation generates.

The goal is to compare the firms' innovation incentives when they compete with each other and when they are merged to form a single entity, referred to as the merged entity. The main difference between the two settings is that, while the competing firms choose their R&D efforts non-cooperatively, the merged entity implements the R&D effort that is jointly optimal for the two merging parties. We assume that the merger changes neither the cost of R&D nor the profitability of an innovation, keeping the number of innovations in the market fixed. Thus, we implicitly disregard potential synergies as well as the effects of the merger on price competition. Shutting down the synergy and the price channels allows us to focus purely on dynamic incentives to innovate.[6]

In our model, the cannibalization effect arises in relation to the follow-up innovation: once the first innovation has already taken place, the merged entity may block the second innovation because it may divert sales from the first innovation. The cannibalization effect does not arise in relation to the first innovation because we assume that the first innovation is a game-changer (so-called drastic or radical innovation) and does not have any substitutes from which to divert sales. In contrast, the appropriability effect arises in relation to the game-changer innovation: for example, when the merged entity blocks the second innovation, the profitability of the first innovation increases, and, hence, the merger increases the probability of the first innovation.

Our model provides new insights for policy makers on how to resolve the cannibalization/appropriability trade-off. Perhaps surprisingly, the benefit of the merger, brought by the appropriability effect, may be higher if the first and second innovations are close demand substitutes.[7] Intuitively, when innovations are close substitutes, expectations of

[5] We assume that a firm's research activity and whether this activity produced an innovation are observable to its rival. This assumption is justified because in many industries, R&D activities are often required to be publicly reported. For example, in the pharmaceutical industry, patented drug development involves public disclosure of a number of milestone developments, and data set Pharmaprojects tracks all drug projects throughout their whole life-cycle. Moreover, in our model, we conjecture that the qualitative results do not change if we assume that only the end results of R&D activities — successful innovations — are observable. See *"Remark on the observability assumption"* at the end of Section 3 for more details.

[6] We briefly discuss the synergy and the price channels in *"Remark on the additive payoff assumption"* at the end of Section 2, but the full analysis of how these channels interact with the forces present in our model is outside the scope of this paper.

[7] Our conclusion that the benefit of the merger may be higher for closer substitutes may seem surprising

intense competition in the future dampen the competing firms' immediate incentives to innovate, thus reducing the probability of the first innovation. In contrast, the merged entity may avoid competition by not producing the second innovation at all, which implies that an increased substitutability of the innovations does not decrease the merged entity's incentives to produce the first innovation. Thus, within our model, under some conditions, the expectation that the prospective innovations will be close substitutes is an argument for, rather than against, the merger.[8]

In addition, we identify a novel effect of the merger that arises from the dynamic nature of the model. Once a firm launches a newly developed product, its competitor immediately learns that the research avenue is, indeed, promising and will eventually yield an innovation. Thus, launching a new product generates positive informational externalities for the other firm in the market. The expectation that the rival's costly R&D effort may provide information about the feasibility of the innovation gives rise to the free-riding problem, thereby incentivizing firms to slow down their R&D activities prior to the first innovation. The merged entity internalizes the informational externalities and, thus, has no reason to slow down R&D. We refer to this mechanism as the informational effect. The nature of this effect is familiar from the strategic experimentation literature since Keller et al. (2005), in which the individual incentives to free-ride on others' experimentation efforts lead to an aggregate experimentation level that is lower than socially optimal. Within our model, the informational effect manifests itself through the timing of the first innovation: conditional on the arrival of the first innovation, the merged entity may take less time to generate it.[9,10]

Regarding the Dow/DuPont merger, the Commission approved the merger conditional on the divestiture of major parts of DuPont's global pesticide business, including its global R&D organization. To justify the decision, the Commission maintained that after the merger, Dow and DuPont would face reduced innovation incentives that would *"manifest themselves in the form of: (1) immediate reduction of incentives to continue with*

because in traditional merger analysis, the price increase following a merger is expected to be stronger when the merging parties' products are closer substitutes. Hence, mergers between close competitors are typically frowned upon.

[8]Our model lends support to the pro-merger arguments put forward by Dow and DuPont (see paragraph 2115 in the Commission's Dow/DuPont decision).

[9]Katz and Shelanski (2007) note that, often, "proponents of the permissive merger policy contend that [...] market consolidation may in fact help to speed innovation by bringing complementary assets together." Although we also conclude that the merger may speed up innovation arrival, the mechanism is different because we assume away any synergy of resources.

[10]Literature on research joint ventures (RJV) also features the free-riding problem in R&D (see Katz (1986) and others). In contrast to our model, RJV models of R&D are static and so cannot provide insights on the timing of innovations; in that literature, free-riding arises because the competitors' R&D effort directly enters each firm's payoff function.

some existing innovation efforts [...] in the case of overlapping lines of research [...], and (2) reduced incentives to develop in the longer term the same number of new products as the combined targets of the Parties before the Transaction" (paragraph 2014 of the Commission's Dow/DuPont decision). Our analysis indicates that while the second concern of the Commission is justified by the cannibalization effect, the first concern may be misplaced if the firms were developing game-changer innovations. Due to the appropriability and informational effects, the immediate incentives to undertake R&D may well be higher for the merged entity.

Our theoretical findings point to a potential bias in the empirical literature. Arguably, identifying the cannibalization effect of a merger is easier than identifying the appropriability and informational effects. The cannibalization effect shuts down the development of substitute products, while the appropriability and informational effects promote and speed up the development of truly novel products. The absence of novel products is harder to detect than the absence of substitute products. This asymmetry in the ease with which the effects can be identified might bias the empirical literature towards finding the negative effect of a merger. Indeed, recent empirical papers, such as Ornaghi (2009), Haucap et al. (2019) and Cunningham et al. (2021), look at pharmaceutical mergers and document a decline in merged entities' patenting and R&D activity following the merger. Despite the potential bias, there are empirical papers that identify a positive effect of the merger: for example, Guadalupe et al. (2012) find that a target company's innovation increases after an acquisition by a multinational firm.

A number of recent papers analyze the impact of a merger on a one-off innovation in a static setting. For example, Federico et al. (2017, 2018), Denicolò and Polo (2018) and Jullien and Lefouili (2020) work with a simple model of product innovation; Motta and Tarantino (2021) and Mukherjee (2022) consider a model in which firms engage in R&D investments aimed at reducing the cost of production (see Jullien and Lefouili (2018) for comprehensive review). In contrast to this literature, we work in a dynamic setting and our results hinge on the dynamic nature of our model. In particular, the sequential order of stochastic innovations allows us to shed new light on the appropriability/cannibalization trade-off, and the informational effect arises because we model R&D efforts as accruing incrementally over time.

There is a large body of literature on research joint ventures (see, for example, Katz (1986); D'Aspremont and Jacquemin (1988); Kamien et al. (1992); Amir et al. (2003)). That literature typically focuses on process innovation, which reduces the cost of production, and compares two settings. In one setting, firms form a joint venture to conduct R&D in a jointly owned lab run; in the other setting, each firm conducts R&D independently in its own lab. In both settings, firms behave non-cooperatively in the product market. In contrast, we focus on product innovation and assume that the merger changes the nature of product market competition. Prior to the merger, firms decide whether to work towards the second innovation in a non-cooperative fashion; after the merger, the merged entity internalizes payoff externality that the second innovation imposes on the first.

We formally model R&D activity as strategic experimentation with exponential bandits and observable actions, as in Keller et al. (2005). In contrast to canonical models that feature only informational externalities, our model features both payoff and informational externalities. Other papers that introduce payoff externalities into a strategic experimentation setting with observable actions include Besanko and Wu (2013), Boyarchenko and Levendorski (2014), Cripps and Thomas (2019), Das and Klein (Forthcoming), and Thomas (2021).[11]

A growing literature applies variations on the technological ladder model to various R&D and antitrust issues. The technological ladder model was first developed in the context of economic growth (Aghion and Howitt, 1992; Aghion et al., 1997, 2001); the current applications consider the impact of firm asymmetries (Cabral, 2018) or of a reduction in the number of firms (Marshall and Parra, 2019) on R&D incentives. This literature is complementary to our paper. The literature is concerned with infinite sequences of innovations, whereby today's technology leaders may become technology laggards tomorrow. In contrast, we consider a sequence of, at most, two potential innovations and assume that the leader maintains the first-mover advantage after the follow-up innovation arrives. Moreover, in contrast to our approach, the technological ladder models do not involve learning about the feasibility of innovation.

The rest of the paper is organized as follows. Section 2 describes the setup. Section 3 presents our main result with three effects of merger on innovations — cannibalization, appropriability and informational effects. Section 4 provides a detailed analysis of the model. Section 5 concludes.

2 Model

To identify the effects of a merger, we compare outcomes in two settings: one in which there are two competing firms and one in which these firms are merged into a

[11]No information externalities and extreme payoff externalities are featured in dynamic models of R&D races with winner-takes-all payoff structure and irrevocable exit. Within this literature, Malueg and Tsutsui (1997) study the effect of the number of players on R&D incentives.

single entity, referred to as the merged entity. The key metrics for the comparison are the number and the expected arrival time of innovations. In formulating our policy recommendations, we operate under the implicit assumption that consumers value a greater variety of innovations and their expeditious arrival.

Competing firms

Two firms, indexed by $i \in \{1, 2\}$, undertake R&D to produce an innovation. Each firm is restricted to producing, at most, one innovation.[12] Time is continuous and indexed by $t \in [0, \infty)$. The discount rate is $r > 0$.

Firms undertake research along a research avenue. At each moment t prior to producing an innovation, firm i chooses the research intensity $x_i(t) \in [0, 1]$. The flow cost of research is $cx_i(t)$, with $c > 0$.

The research avenue can be one of two types: *good* or *bad*. A bad avenue can never succeed in generating an innovation. A good avenue generates an innovation for firm i according to a Poisson process with intensity equal to firm i's research intensity, $x_i(t)$.

Ex-ante, firms do not know the type of the research avenue; at time $t = 0$, they have a common prior belief $p(0)$ that the avenue is good. Both firms can perfectly observe each other's research intensity allocations and innovation arrivals. Hence, the firms continue to share a common posterior belief $p(t)$ at all t.

Until a firm innovates, it gets the flow payoff of 0. If a firm innovates first, as long as it remains the sole innovator, it receives a flow payoff of $\pi > 0$. Once its competitor also innovates, flow payoffs of both firms change: the leader (which innovated first) gets $\lambda\pi$, and the follower gets $\phi\pi$. Parameters are such that

$$1 \geq \lambda \geq \phi \geq 0. \tag{1}$$

If $\lambda = 1$ and $\phi = 0$, then the payoff structure is winner-takes-all; if $\lambda = \phi = 1$, then the payoff structure is standard for a strategic experimentation model with no payoff externalities. In Appendix B, we use a Stackelberg game with differentiated products to provide microfoundations for the assumptions that competition from the second innovation weakly lowers the leader's payoff ($\lambda\pi \leq \pi$) and that the leader's payoff $\lambda\pi$ is weakly higher than the follower's payoff $\phi\pi$.

[12] In Appendix C, we show that, assuming that the game ends after the second innovation, our results remain qualitatively the same if each firm can continue research after producing the first innovation.

Merged entity

We set up the merged entity's problem so that it is possible to make the equilibrium outcome with two independent firms meaningfully comparable to the merged entity's optimum. In particular, the merged entity can innovate twice, but the merger does not create any synergy of resources. Thus, prior to any innovation, the merged entity has two units of research intensity, $X(t) = x_1(t) + x_2(t) \in [0, 2]$, to devote to the avenue; after the first innovation, the merged entity has only one unit of research intensity, $X(t) \in [0, 1]$.

The flow payoff of the merged entity is 0 prior to the first innovation, π after the first innovation, and $(\lambda + \phi)\pi$ after the second innovation. The merged entity's flow cost of research is $cX(t)$. In other words, the merged entity's payoff is simply the sum of the independent firms' payoffs.

Remark on the additive payoff assumption. The assumption that the merged entity's payoff is simply the sum of the independent firms' payoffs requires some discussion.

In practice, the merged entity's payoff may be higher than the sum of the merging parties' payoffs because the merger may change the competitiveness of the relevant product markets.[13] The merger-induced change in the competitiveness of the product market, studied in Federico et al. (2017) and Federico et al. (2018), has an ambiguous effect on consumer welfare. Higher prices hurt consumers after an innovation has been developed. However, the prospect of less intense price competition strengthens the appropriability effect, thus incentivizing R&D.

Potentially, the merger may also bring synergy of resources, thus allowing the merged firm to reduce production or R&D costs.[14] The cost-reduction effect of the merger has been studied, for example, in Atallah (2016) and Denicolò and Polo (2021).

We abstract away from such considerations and assume that, keeping the dynamics of innovations the same, the merger changes neither the profitability of an innovation nor the cost of R&D. This assumption allows us to focus purely on incentives to innovate, in isolation from the standard unilateral effects of a merger.

[13]In the Dow/DuPont decision, according to the EC (see para 2044 in the decision), "*following a merger, the merging parties coordinate the pricing of their products and thus a merger may increase the prices and profits of the merged entity. Less intense competition in the product market can increase the net revenues earned by a product line both when the firms innovate to improve the products in that line and when they do not.*"

[14]As noted in the Dow/DuPont decision (see para 3296), "efficiencies brought about by a merger may counteract [...] the potential harm to consumers that it might otherwise have."

3 Main result

Our main result demonstrates that the merger has three effects. First, the merged entity in some instances blocks the second innovation that the competing firms would have delivered (*cannibalization effect*). Second, the merger may increase the likelihood of the first innovation (*appropriability effect*), and, third, the merger may bring the first innovation forward in time (*informational effect*).

Cannibalization effect. The merged entity faces reduced incentives to produce the second innovation because it internalizes negative payoff externalities that the competing firms impose on each other. In particular, in contrast to the competing firms, the merged entity takes into account that the competition from the second innovation reduces, or *cannibalizes*, the profit from the first innovation, π, to $\lambda\pi$. We refer to this as the cannibalization effect.

Appropriability effect. The flip side of the cannibalization effect is that, for the merged entity, the marginal benefit of undertaking R&D to produce the first innovation is higher. Indeed, due to reduced competition from the second innovation, the first innovation generates higher expected profit, which incentivizes the merged entity to undertake more research to produce the first innovation. Moreover, even if both the competing firms and the merged entity produce the second innovation, the merged entity still has higher incentives to undertake research on the first innovation because, in contrast to the leader of the competing firms, it reaps the benefit not only from the first innovation, but also from the second innovation. As a result, the merger increases the likelihood that the first innovation will be produced. The positive effect of the merger on this likelihood occurs because the merged entity fully *appropriates* the benefits of its own R&D efforts — that is why we refer to this effect as the appropriability effect.

Informational effect. The merged entity also internalizes positive informational externalities that the competing firms impose on each other. In particular, an innovation by the leader immediately informs the competitor that the research avenue is good and, thus, is capable of generating another innovation. Hence, before the arrival of the first innovation, the competing firms have incentives to reduce their own research intensity in the hope of free-riding on the opponent's R&D efforts. In contrast, the merged entity cannot fall back on somebody else's efforts to discover whether an innovation is feasible and so undertakes research at higher intensity, which brings the first innovation forward

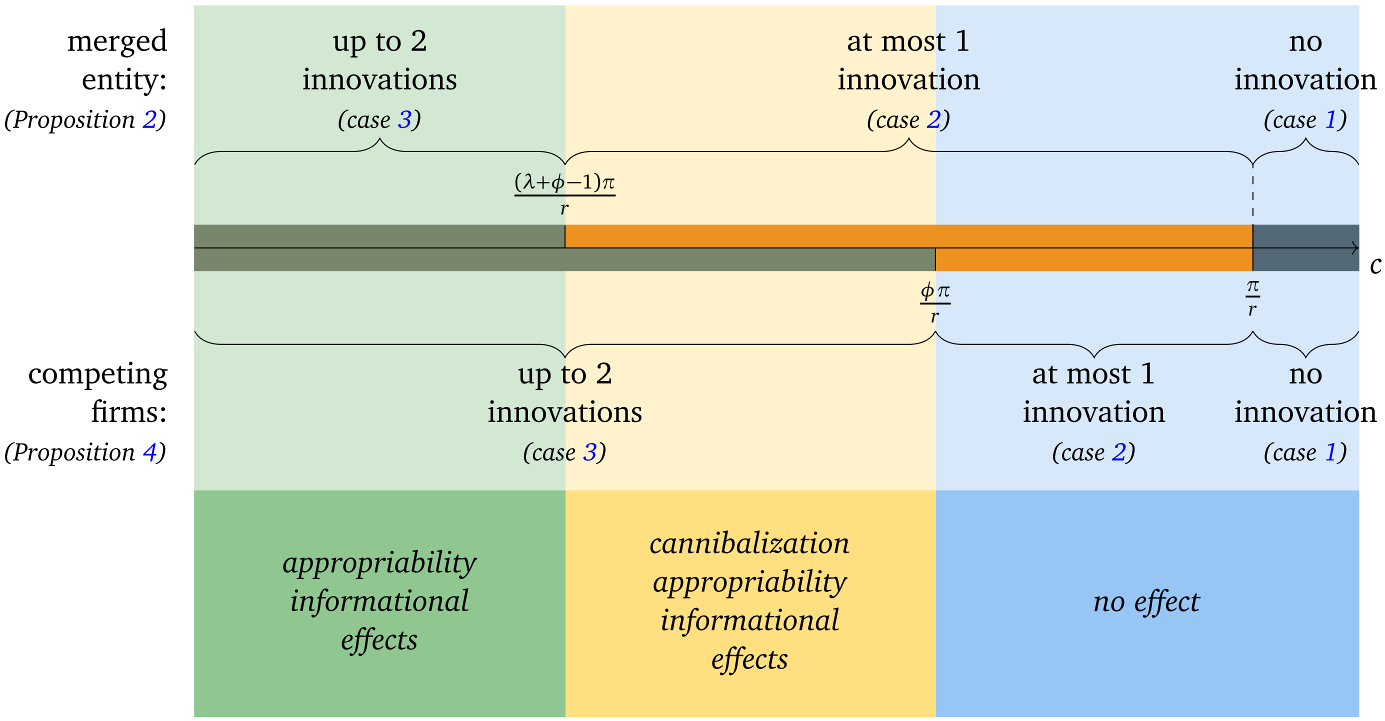

Figure 1: The number of innovations in each setting and the effects of the merger. The case numbers in the brackets correspond to cases in Proposition 2 for the merged entity setting and to cases in Proposition 4 for the competing firms setting.

in time. We refer to the positive effect of the merger on the timing of the first innovation as the informational effect because it is connected to *informational* externalities.

Theorem 1 derives the parameter regions in which the cannibalization, appropriability and informational effects are present. Figure 1 visualizes these regions. All three effects are present in the yellow region, where the costs are sufficiently low for the competing firms to produce the second innovation but sufficiently high for the merged entity to block the second innovation. If the second innovation is not profitable for the merged entity even when research is costless — that is, if $\lambda + \phi < 1$ — the yellow region extends to the low cost as well, making the green region, which features only the appropriability and informational effects, disappear.

Theorem 1.

1. *If*

$$\frac{\phi\pi}{r} < c, \tag{2}$$

then the merger has no effect on the number and the probability of innovations.

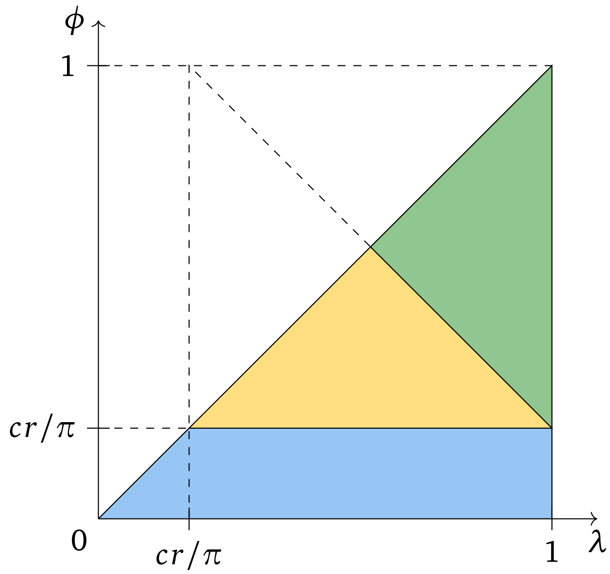

Figure 2: The effect of the merger on innovations. In the blue region, the merger has no effect; in the green region, the merger has an unambiguously positive effect; in the yellow region, the merger blocks the second innovation but increases the probability and the speed of arrival of the first innovation.

2. *If*
$$\frac{(\lambda+\phi-1)\pi}{r} < c \leq \frac{\phi\pi}{r}, \tag{3}$$
then the merger blocks the second innovation but increases the probability that the first innovation arrives and, moreover, brings it forward in time. The increase in the probability of the first innovation does not depend on ϕ and decreases in λ. The decrease in the expected arrival time of the first innovation (conditional on the first innovation arriving) increases in ϕ and decreases in λ.

3. *If*
$$c \leq \frac{(\lambda+\phi-1)\pi}{r}, \tag{4}$$
then the merger has an unambiguously positive effect: while it does not block the second innovation, it increases the probability that the first innovation arrives and brings it forward in time.

Figure 2 visualizes the regions from Theorem 1 on the λ/ϕ plane. Since $\lambda > \phi$, the relevant part of the plane is the lower triangle.

Case 2 corresponds to the yellow region in Figure 2, where the follower's payoff is sufficiently large (that is, ϕ is relatively high), but together with the leader's payoff it is relatively low (that is, $\phi + \lambda$ is sufficiently small). Sufficiently high ϕ ensures that the competing firms produce the second innovation, which triggers free-riding prior to the arrival of the first innovation, thus leading to the informational effect of the merger. Sufficiently low $\phi + \lambda$ reflects substantial competition between innovations, which induces the merged entity not to produce the second innovation (the cannibalization effect). Since the competing firms produce the second innovation while the merged entity does not, the profitability of the first innovation is higher for the merged entity, and, thus, the merged entity produces the first innovation with a higher probability than the competing firms (the appropriability effect).

According to case 2 in Theorem 1, the increase in the probability of the first innovation, which the merger induces — that is, the strength of the appropriability effect — does not depend on ϕ and decreases in λ. Intuitively, a higher λ lowers the negative payoff externality that competition from the second innovation imposes on the first innovation. As a result, the expected payoff from innovating first increases, which, in turn, strengthens the preemption motive for research. Hence, a higher λ boosts the incentives of the competing firms to conduct research on the first innovation. The incentives of the merged entity are unaffected, as the merged entity does not produce the second innovation. Hence, the increase in the probability of the first innovation, which the merger induces, decreases in λ.

In addition, according to case 2 in Theorem 1, the decrease in the expected arrival time of the first innovation — that is, the strength of the informational effect of the merger — increases in ϕ and decreases in λ. The intuition for the impact of λ on the arrival time is the same as that provided above in relation to the probability of the first innovation. The impact of ϕ is positive because higher ϕ increases the positive informational externality that one competing firm imposes on the other when undertaking research on the first innovation. Therefore, higher ϕ aggravates the free-riding problem of the competing firms and, thus, increases the positive effect of the merger, which eliminates free-riding altogether.

In sum, in the yellow region of Figure 2, the regulatory authorities need to trade off the loss of the second innovation, due to cannibalization effect, against a higher probability and speed of arrival of the first innovation due to appropriability and information effects, respectively. A surprising policy implication is that in this region, the benefit of the merger is higher when the first and the second innovations are closer substitutes. This conclusion follows because the appropriability effect becomes stronger as λ decreases. Intuitively, low λ reflects intense competition between the innovations, which occurs when the innovations are close substitutes. Hence, as innovations become closer sub-

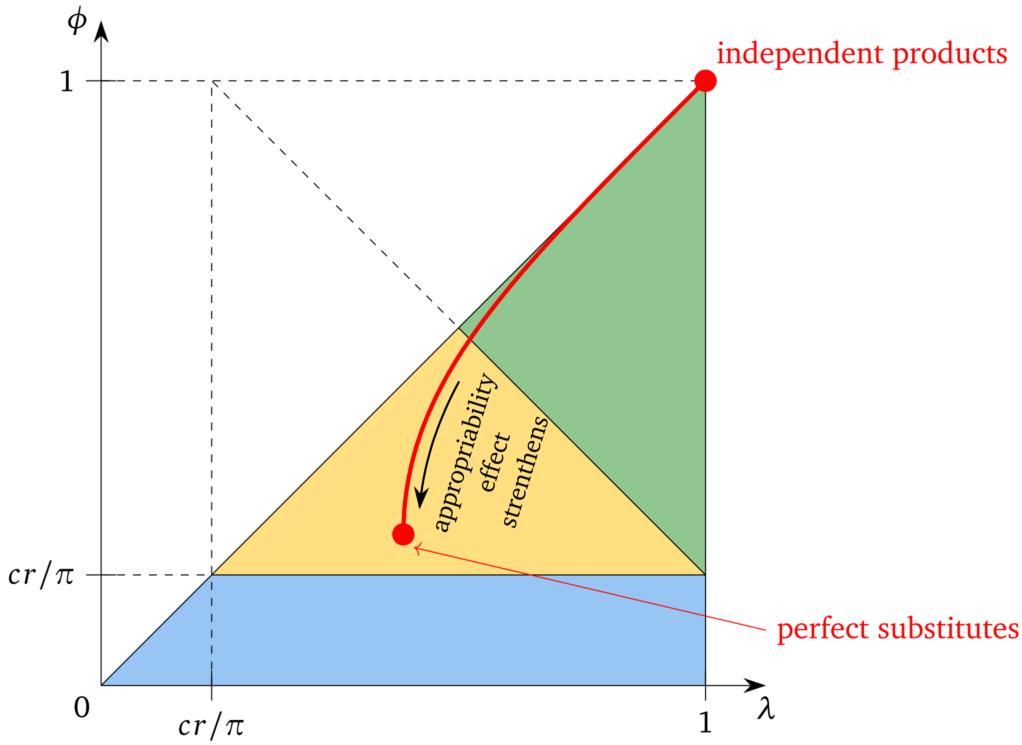

Figure 3: Parametrization of payoffs from Appendix B: The red curve traces the feasible payoff pairs $(\lambda(\theta), \phi(\theta))$ as the degree of product substitutability θ varies from $\theta = 0$ (independent products) to $\theta = 1$ (perfect substitutes).

stitutes (that is, as λ decreases), the merger causes a more significant increase in the probability of the first innovation (that is, the appropriability effect becomes stronger).

The intuition that low λ occurs when the innovations are close substitutes is supported by microfoundations for λ and ϕ from Appendix B, which is based on the Stackelberg competition model with differentiated products. The red curve in Figure 3 illustrates the parametrization of λ and ϕ through the degree of substitutability between the products — see (B.10) in Appendix B. The shape of this curve is determined by two forces: the strength of competition and the first-mover advantage of the leader. As innovations become closer substitutes, both ϕ and λ decrease due to more intense competition between innovations, with λ decreasing less because of the first-mover advantage. Starting from the point $\lambda = \phi = 1$, at which consumers view the innovations as independent products, as the degree of demand substitutability increases, the feasible payoffs move down along the red curve, first passing through the green parameter region where the merger is desirable. As the substitutability of innovations increases further, the payoffs move into the yellow region. At the boundary between the green and the yellow regions, there is a discontinuous drop in the desirability of the merger because in the yellow region, the merger blocks the second innovation. As the substitutability increases further, the payoffs move deeper into the yellow region, where the merger, while continuing to block the second innovation, brings about an ever higher increase in the probability of the first innovation. Hence, the decision of the competition authority to block the merger may be non-monotonic in the degree of substitutability between innovations: it may be beneficial to allow the merger when innovations are either very close or very distant substitutes, but to block the merger when innovations are moderate substitutes — that is, when the firms' payoffs lie in the yellow region (where the merger blocks the second innovation) but are close to the boundary with the green region (where the merger does not have negative effects on innovations).

The above policy implication relies on the assumption that the regulatory authority deems whether the first innovation occurs (the appropriability effect) more important than when it occurs (the informational effect). Putting significant weight on the informational effect complicates matters. The strength of the informational effect increases in ϕ and decreases in λ. Hence, the change in the informational effect is ambiguous as the degree of substitutability increases. In particular, the strength of the informational effect may decrease, counteracting the increase in the appropriability effect, in which case the assessment of whether the merger is more beneficial when innovations are closer substitutes depends on the priorities of the regulatory authority.

Remark on the incentives to merge. In our model, the firms always (weakly) jointly benefit from the merger because, by the additive payoff assumption, the merged entity maximizes the sum of individual firms' payoffs. Since the merger is always profitable, a mere notification of the intention to merge does not provide competition authorities with any information that is relevant for its merger assessment.

Remark on the observability assumption. In the competing firms setting, our solution relies heavily on the assumption that each firm observes its rival's research intensities. Bonatti and Hörner (2011) considers a model of strategic experimentation in teams in which individual team members' learning intensity is hidden. Marlats and Ménager (2021) allow players to choose whether, at a cost, to observe the other player's experimentation effort and outcomes. Based on the findings in these papers, we conjecture that relaxing the assumption of observable research intensities mitigates the free-riding problem but does not completely eliminate it. Thus, the informational effect of the merger would still be present, albeit weaker.

Remark on the no dead-end discoveries assumption. We also assume that the R&D activity never reveals that the research avenue is a dead end, which can never generate a successful innovation. Akcigit and Liu (2016) show that if R&D activity could result in both successful innovations, which are observable by rivals, and dead-end discoveries, which are not observable by rivals, then competition between firms would result in wasteful duplication of dead-end research. In this case, the merger may have an additional positive effect — elimination of wasteful duplicative dead-end research.

Remark on the two-to-one merger assumption. We assume that prior to the merger, there are only two firms, and no other firms undertake R&D along the research avenue. Introducing competitors would complicate the analysis without qualitatively affecting conclusions. With more firms, the effects that we identify in our reduced-form model would be present, though possibly weaker.

4 Analysis

In this section, we prove Theorem 1. We start by analyzing the merged entity's optimization problem (Proposition 2) and then proceed to solve the competing firms' game (Proposition 4). We conclude with Section 4.3, in which Theorem 1 is proven as a direct consequence of Propositions 2 and 4.

There are two distinct stages in the settings with the competing firms and with the merged entity: stage 1, when no innovation has arrived yet; and stage 2, which begins after the first innovation has arrived. In each setting, we begin our analysis with stage 2 and then proceed to stage 1.

4.1 Merged entity

Stage 2: After the first innovation

After the first innovation, the merged entity knows for sure that the research avenue is good and, thus, can generate the second innovation. Whether to undertake research to get the second innovation depends on whether the second innovation is profitable. The flow payoff with a single innovation is π, while after the second innovation, it is $(\lambda+\phi)\pi$. Thus, the second innovation is profitable if $\lambda+\phi>1$.

Proposition 1 characterizes the optimal research decision of the merged entity and its payoff at the optimum. It states that, conditional on the second innovation being profitable, the merged entity undertakes research if and only if the flow cost c is sufficiently low.[15]

Proposition 1. *If*

$$\frac{(\lambda+\phi-1)\pi}{r} \geq c, \tag{5}$$

then the merged entity continues research at full intensity until the second innovation arrives; otherwise, it aborts research and the second innovation never arrives. The merged entity's expected payoff from the second stage is

$$V_M = \frac{\pi}{r} + \frac{1}{1+r}\max\left\{\frac{(\lambda+\phi-1)\pi}{r} - c, 0\right\}. \tag{6}$$

Proof. See Appendix A.1. □

Stage 1: Before the first innovation

Prior to any innovation, the merged entity is uncertain about the type of research avenue. In the absence of innovation, the belief of the merged entity that the avenue is good, $p(t)$, evolves according to the law of motion derived from Bayes' rule:

$$p'(t) = -X(t)(1-p(t))p(t). \tag{7}$$

Formula (7) implies that if the merged entity undertakes research with positive intensity, it becomes progressively more pessimistic that the avenue is good — that is, $p(t)$ decreases.

At the optimum, if the belief drops below a certain threshold — which we refer to as a stopping threshold — the merged entity optimally aborts research and the game ends with no innovation. If the merged entity succeeds before the belief reaches the stopping threshold, $p(t)$ immediately jumps to 1 and the game proceeds to the second stage.

Proposition 2 derives the optimal stopping threshold and uses Proposition 1 to fully characterize the merged entity's uniquely optimal research strategy.

Proposition 2.

1. *If*

$$\frac{\pi}{r} \leq c, \tag{8}$$

then the merged entity does not undertake research.

[15] If the merged entity is indifferent between producing the second innovation and aborting research after the first innovation, we assume it chooses to continue research.

2. *If*
$$\frac{(\lambda+\phi-1)\pi}{r} < c < \frac{\pi}{r}, \tag{9}$$
then the merged entity undertakes research at full intensity as long as its current belief $p(t)$ *is above threshold*
$$\check{p} = \frac{cr}{\pi}. \tag{10}$$
Once an innovation arrives or the belief falls to $\check{p}$*, the merged entity completely aborts research efforts.*

3. *If*
$$c \leq \frac{(\lambda+\phi-1)\pi}{r}, \tag{11}$$
then the merged entity undertakes research at full intensity as long as its current belief $p(t)$ *is above threshold*
$$\check{p} = \frac{cr}{\pi}\frac{1+r}{\lambda+\phi+r-rc/\pi}. \tag{12}$$
Once the belief falls to $\check{p}$*, the merged entity aborts research efforts. If an innovation arrives before that, the merged entity undertakes research at full intensity until the second innovation arrives.*

Proof. See Appendix A.2. □

Given expression (6) for V_M, the stopping threshold, defined in (10) and in (12), is equal to c/V_M. This expression is intuitive because it equalizes the flow cost of research, c, with the expected benefit from the first innovation, pV_M. The difference in definitions (10) and (12) comes from the difference in the expression for V_M, which depends on whether there is research in stage 2.

4.2 Competing firms

Stage 2: After the first innovation

Once the first innovation arrives, all uncertainty about the type of the research avenue is resolved: it is common knowledge that the avenue is good and, thus, can generate the second innovation.

Since each firm is restricted to producing, at most, one innovation, the leader — that is, the firm that produced the first innovation — is not active anymore. The firm that did not innovate — the potential follower — has a choice between continuing research or giving up and leaving the competitor to be the sole innovator. The optimal decision depends on the cumulative payoff that the second innovation brings to the follower, $\phi\pi/r$. If it is large relative to the flow cost of research c, then it is optimal to undertake research until the second innovation arrives. Proposition 3 summarizes the optimal research decision of the potential follower and the firms' equilibrium payoffs.[16]

Proposition 3. *If*

$$\frac{\phi\pi}{r} \geq c, \tag{13}$$

then the potential follower undertakes research at full intensity until the second innovation arrives; otherwise, it aborts research and the second innovation never arrives. The potential follower's expected payoff from the second stage is

$$V_F = \frac{1}{1+r} \max\left\{\frac{\phi\pi}{r} - c, 0\right\}, \tag{14}$$

while the leader's expected payoff is

$$V_L = \frac{\lambda + r}{1+r}\frac{\pi}{r} \quad \text{if (13) holds,} \quad V_L = \frac{\pi}{r} \quad \text{otherwise.} \tag{15}$$

Proof. See Appendix A.3. □

Comparing Proposition 3 and Proposition 1, we conclude that, conditional on the potential follower and the merged entity making the same research decision in the second stage, we always have $V_L + V_F = V_M$. This is a direct consequence of the additive payoff assumption. However, the conditions under which undertaking research is optimal in the second stage, (5) and (13), are different. In particular, condition (13) is weaker than condition (5), which means that the merger may block the second innovation. We will come back to this point in Section 4.3.

Stage 1: Before the first innovation

Prior to any innovation, the firms are uncertain about the type of the research avenue. Since each firm's research intensity and research outcomes are publicly observable, the firms share a common belief that the avenue is good, $p(t)$. In the absence of any innovation, this belief evolves according to (7), where $X = x_1 + x_2$. Formula (7) implies that when the firms undertake research with positive intensity, in the absence of innovation, they become progressively more pessimistic that the avenue is good.

[16] As for the merged entity, we assume that the potential follower chooses to undertake research whenever indifferent.

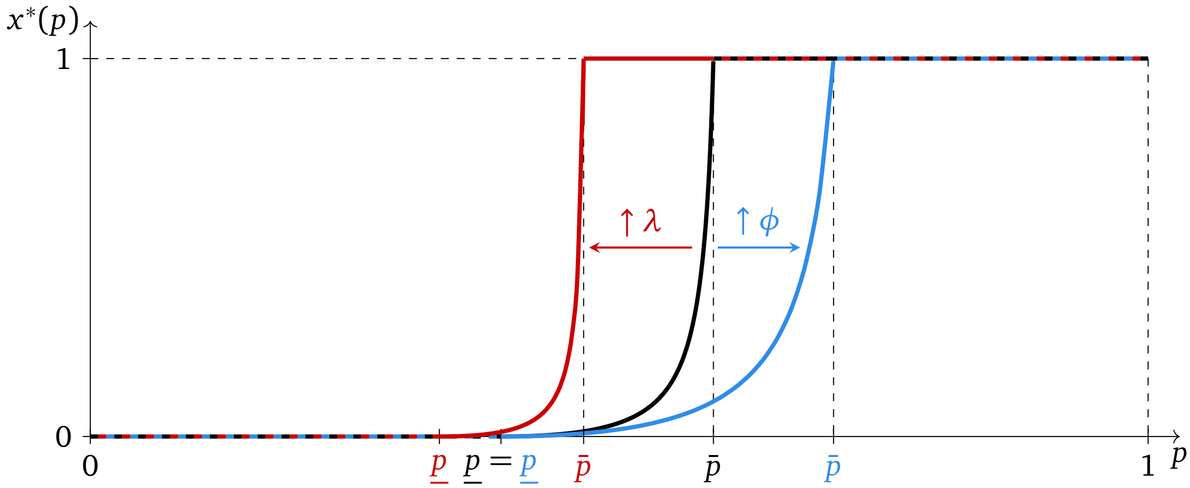

Figure 4: Equilibrium research intensity of each firm as a function of common belief p that the avenue is good. The black curve corresponds to $c/\pi = 3$, $r = 0.1$, $\lambda = 0.75$, $\phi = 0.6$. The red and blue curves illustrate how the intensity changes with λ and ϕ: the red curve corresponds to $\lambda = 0.9$ and $\phi = 0.6$; the blue curve corresponds to $\lambda = 0.75$ and $\phi = 0.7$.

We are looking for a symmetric Markov perfect equilibrium of the game, in which the firms use identical stationary Markovian strategies with belief p as the state variable. Proposition 4 proves that such an equilibrium is unique and provides the equilibrium characterization. The form of the equilibrium closely resembles that in Keller et al. (2005), in which there are no payoff externalities.

In the equilibrium, as in the merged entity setting, if the belief drops below a certain threshold — which we again refer to as a stopping threshold — the firms abort research and the game ends with no innovation.

Before the belief reaches the stopping threshold, the equilibrium research intensity depends on whether there is innovation in the second stage. If there is no research in the second stage, then, in the first stage, the firms undertake research at full intensity. If the potential follower continues research in the second stage, then, in the first stage, after reaching a certain belief threshold $\bar{p}$ above the stopping threshold $\underline{p}$, the firms gradually reduce their equilibrium research intensity from 1 to 0, as illustrated in Figure 4. As a result, their common belief $p(t)$ approaches but never reaches the stopping threshold $\underline{p}$.

Proposition 4.

1. If

$$\frac{\pi}{r} \leq c, \tag{16}$$

then neither firm undertakes research.

2. *If*

$$\frac{\phi\pi}{r} < c < \frac{\pi}{r}, \tag{17}$$

then the firms undertake research at full intensity as long as their current belief $p(t)$ *is above threshold*

$$\hat{p} = \frac{cr}{\pi}. \tag{18}$$

Once an innovation arrives or the belief falls to $\hat{p}$*, the firms completely abort research efforts.*

3. *If*

$$c \leq \frac{\phi\pi}{r}, \tag{19}$$

then the firms undertake research at full intensity as long as their current belief $p(t)$ *is above threshold* $\bar{p}$*. Once the belief falls to* $\bar{p}$*, the firms undertake research at intensity* $x_1(t) = x_2(t) = x^*(p(t))$*, which decreases over time, from 1 at* $p(t) = \bar{p}$ *to 0 at* $t \to +\infty$*. In the absence of innovation, the belief approaches threshold* $\underline{p}$ *at* $t \to +\infty$*. Once an innovation arrives, the firm that did not innovate undertakes research at full intensity until the second innovation arrives.*

Threshold $\underline{p}$ *is defined as*

$$\underline{p} = \frac{cr}{\pi}\frac{1+r}{\lambda+r}. \tag{20}$$

Threshold $\bar{p} \in [\underline{p}, 1)$ *is defined as a unique solution to*

$$\frac{1}{\underline{p}} - \frac{1}{\bar{p}} - \frac{1-\bar{p}}{\bar{p}}\frac{r}{1+r}\ln\frac{\bar{p}(1-\underline{p})}{(1-\bar{p})\underline{p}} = \frac{\phi\pi - cr}{cr(1+r)^2}. \tag{21}$$

The equilibrium research intensity $x^*(p)$ *on* $p \in (\underline{p}, \bar{p})$ *is defined as*

$$x^*(p) = \frac{\dfrac{1}{\underline{p}} - \dfrac{1}{p} - \dfrac{1-p}{p}\ln\dfrac{p(1-\underline{p})}{(1-p)\underline{p}}}{\dfrac{1}{r}\left(\dfrac{1}{p} - \dfrac{1}{\bar{p}}\right) + \dfrac{1}{\underline{p}} - \dfrac{1}{\bar{p}} - \dfrac{1-\bar{p}}{\bar{p}}\ln\dfrac{\bar{p}(1-\underline{p})}{(1-\bar{p})\underline{p}}}. \tag{22}$$

Proof. See Appendix A.4. □

As in the merged entity setting, the stopping thresholds $\hat{p}$ and $\underline{p}$, defined in (18) and (20), respectively, equalize the flow cost of research, c, with the expected benefit from the first innovation, $\hat{p}V_L$ and $\underline{p}V_L$. The difference in the definitions of $\hat{p}$ and $\underline{p}$ comes from the difference in the expression for V_L, see (15).

In the region $(\underline{p}, \bar{p})$, where the equilibrium research intensity is strictly between 0 and 1, the firms must be indifferent to undertaking research. Intuitively, the indifference region $(\underline{p}, \bar{p})$ appears because of the free-riding effect: once one firm innovates, the other firm learns that the research avenue is good, thus free-riding on past research efforts of the innovator. Then, during the first stage, in the indifference region, if a competitor undertakes research at full intensity, each firm's best response is to pause its own research efforts and see whether the competitor's research efforts will be successful. Such a free-riding effect was first identified in the Poisson environment by Keller et al. (2005).

Lemma 1 highlights some comparative statics of thresholds $\underline{p}$ and $\bar{p}$ and the equilibrium research intensity $x^*(p)$. These comparative statics are illustrated in Figure 4 and will be used later in Section 4.3.

Lemma 1. *Threshold* $\underline{p}$ *defined in* (20) *decreases in* λ *and is independent of* ϕ*. Threshold* $\bar{p}$ *defined in* (21) *decreases in* λ *and increases in* ϕ*. The equilibrium research intensity* $x^*(p)$ *defined in* (22) *increases in* λ *and decreases in* ϕ*.*

Proof. See Appendix A.5. □

Lemma 1 implies that the length of the indifference region, $(\underline{p}, \bar{p})$, is increasing in ϕ. Intuitively, since higher ϕ increases the follower's payoff V_F but does not affect the leader's payoff V_L, higher ϕ increases the benefit from free-riding on the information that the leader's innovation generates. The larger the free-riding benefit, the larger the indifference region, in which the firms reduce their research intensities to prevent the competitor from free-riding.

Lemma 1 states that the indifference region $(\underline{p}, \bar{p})$ moves down as λ increases. Intuitively, higher λ increases the payoff from becoming the leader, thereby strengthening the preemption motive in the first stage. Thus, in the first stage, the firms compete more aggressively to become the leader, which implies that the firms undertake research at full intensity for a wider range of beliefs (i.e., interval $(\bar{p}, 1)$ is larger) and continue some research at more pessimistic beliefs (i.e., $\underline{p}$ is lower).

The comparative statics of $\underline{p}$ and $\bar{p}$ determines the comparative statics of $x^*(p)$. The equilibrium research intensity $x^*(p)$ decreases from 1 to 0 as the belief moves from $\bar{p}$ to $\underline{p}$. Hence, it is reasonable to expect that, for a fixed p, $x^*(p)$ increases as the indifference region $(\underline{p}, \bar{p})$ moves down.

In contrast to the competing firms environment in Proposition 4, the optimal research intensity of the merged entity in Proposition 2 is always either 0 or 2 (full intensity). In other words, there is no belief region in which the merged entity is indifferent between undertaking and not undertaking research. This is expected because the merger eliminates the free-riding problem, which may emerge when the firms compete.

4.3 The effect of the merger

Figure 1 compares the results of Propositions 2 and 4 in terms of the number of innovations. Three distinct cases emerge; they correspond to three cases in Theorem 1.

Case 1 in Theorem 1: The merger has no effect on innovations.

This case is defined by restriction (2) and comprises two separate regions:

1) $\pi/r \le c$ (case 1 in Proposition 2 and case 1 in Proposition 4): Neither the competing firms nor the merged entity undertake any research, and so there are no innovations in either setting.

2) $\phi\pi/r < c < \pi/r$ (case 2 in Proposition 2 and case 2 in Proposition 4): In both settings, the cumulative research intensity is 2 until an innovation arrives or until the belief reaches threshold $\hat{p} = \check{p} = cr/\pi$, and there is no research after either of the events.

In sum, if (2) holds, then neither the merged entity nor the competing firms have incentives to work on the second innovation after the first innovation arrives. The merger has no effect because, intuitively, in the absence of the second innovation, there are no payoff externalities, and informational externalities are irrelevant.

Case 3 in Theorem 1: The merger has unambiguously positive effect on innovations.

It is instructive to consider case 3 — in which there are only two effects of the merger, both working in the same direction — before case 2, in which an additional effect appears, working in the opposite direction.

Case 3 in Theorem 1 is defined by restriction (4) and corresponds to case 3 in Proposition 2 and case 3 in Proposition 4. Both the merged entity and the competing firms have incentives to work on the second innovation after the first innovation arrives. However, before the first innovation arrives, the merged entity undertakes research at full intensity until belief threshold $\check{p}$, while the competing firms undertake research at full intensity only until belief threshold $\bar{p} > \check{p}$ and then gradually lower their research intensity, with

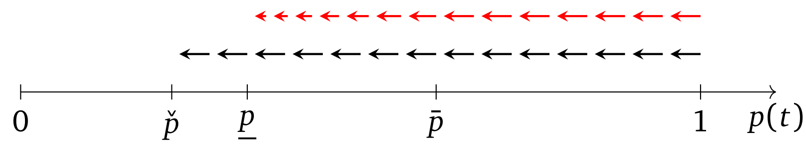

Figure 5: Phase diagram of belief updating (7), assuming $c < \phi\pi/r$. The black arrows indicate the belief trajectory of the merged entity. The red arrows indicate the belief trajectory of the competing firms.

belief approaching threshold $\underline{p} > \check{p}$ (see Figure 5). Thus, the merger positively affects the arrival of the first innovation in two ways.

First, since $\underline{p} > \check{p}$, the competing firms undertake research over a shorter interval of beliefs. Intuitively, the merged entity has higher incentives to innovate in the first stage because, unlike the leader of the competing firms, it reaps the benefit from the second innovation, V_F; formally, V_F captures the difference between $\underline{p}$ and $\check{p}$ because $\underline{p} = c/V_L$, $\check{p} = c/V_M$ and $V_M = V_L + V_F$. By Lemma 2 below, $\underline{p} > \check{p}$ implies that the merger increases the probability of the first innovation. Intuitively, a shorter interval of beliefs at which some research is undertaken means that the cumulative investment in research is lower, and the probability of the first innovation is an increasing function of the cumulative investment.

Lemma 2. *For both the merged entity and the competing firms, given a prior belief p_0 and a stopping threshold p_S where all research activity ceases, the probability of the first innovation is equal to*

$$1 - \frac{1/p_0 - 1}{1/p_S - 1}, \tag{23}$$

which is decreasing in p_S.

Proof. See Appendix A.6. □

Second, the merger increases the research intensity on the belief interval $(\underline{p}, \bar{p})$. Hence, conditional on the first innovation arriving, in expectation, this innovation takes less time to arrive when the firms are merged.

Case 2 in Theorem 1: The merger increases the probability and the arrival speed of the first innovation but blocks the second innovation.

This case is defined by restriction (3) and corresponds to case 2 in Proposition 2 and case 3 in Proposition 4. The competing firms have incentives to work on the second innovation after the first innovation arrives, while the merged entity does not. Thus, the merger blocks the second innovation. Moreover, as in case 3 in Theorem 1, since the stopping threshold of the merged entity is lower than the stopping threshold of the competing firms, $\underline{p} > \check{p}$, and since the competing firms do not always undertake research at full intensity, as illustrated in Figure 5, the merger increases the probability of the first innovation and brings it forward in time.

The increase in the probability of the first innovation that the merger induces does not depend on ϕ and decreases in λ. Indeed, by Lemma 2, for a fixed prior, the probability of the first innovation depends only on the stopping threshold, and this dependence is negative. Since $\check{p}$ defined in (10) depends on neither λ nor ϕ, the probability with which the merged entity produces the first innovation does not depend on λ and ϕ. At the same time, since $\underline{p}$ defined in (20) decreases in λ and is independent of ϕ, the probability with which the competing firms produce the first innovation increases in λ and is independent of ϕ.

The decrease in the expected arrival time of the first innovation that the merger induces increases in ϕ and decreases in λ. Indeed, for the merged entity, the expected arrival time of the first innovation is independent of ϕ and λ because, as previously noted, the belief threshold $\check{p}$ defined in (10) is independent of ϕ and λ. In contrast, for the competing firms, the expected arrival time of the first innovation decreases in λ and increases in ϕ because, as stated in Lemma 1 and illustrated in Figure 4, as λ increases or ϕ decreases, each firm's equilibrium research intensity increases, weakly for all p and strictly for some p.

5 Conclusion

The Dow/DuPont merger gave rise to a flurry of academic papers that identify various positive and negative effects of mergers on innovation. Despite this heightened interest, the literature failed to come to a consensus regarding a presumed effect of mergers on innovation (see, Jullien and Lefouili (2018)). We contribute to this growing literature by modeling the R&D process in continuous time and allowing for multiple sequential innovations. The dynamic nature of our model allows us to gain new insight into the cannibalization/appropriability trade-off and capture the novel informational effect.

References

Aghion, Philippe and Peter Howitt, "A model of growth through creative destruction," *Econometrica*, 3 1992, *60* (2), 323–351. 6

__, Christopher Harris, and John Vickers, "Competition and growth with step-by-step innovation: An example," *European Economic Review*, 4 1997, *41* (3-5), 771–782. 6

__, __, Peter Howitt, and John Vickers, "Competition, imitation and growth with step-by-step innovation," *Review of Economic Studies*, 7 2001, *68* (3), 467–492. 6

Akcigit, Ufuk and Qingmin Liu, "The role of information in innovation and competition," *Journal of the European Economic Association*, 2016, *14* (4), 828–870. 15

Amir, Rabah, Igor Evstigneev, and John Wooders, "Noncooperative versus cooperative R&D with endogenous spillover rates," *Games and Economic Behavior*, 2 2003, *42* (2), 183–207. 5

Arrow, Kenneth J, "Economic welfare and the allocation of resources for invention," in NBER, ed., *The Rate and Direction of Inventive Activity: Economic and Social Factors*, London: Princeton University Press, 1962, pp. 609–626. 2

Atallah, Gamal, "Endogenous efficiency gains from mergers," *Southern Economic Journal*, 7 2016, *83* (1), 202–235. 8

Besanko, David and Jianjun Wu, "The impact of market structure and learning on the tradeoff between R&D competition and cooperation," *Journal of Industrial Economics*, 3 2013, *61* (1), 166–201. 6

Bonatti, Alessandro and Johannes Hörner, "Collaborating," *American Economic Review*, 2011, *101* (2), 632–663. 14

Boyarchenko, Svetlana and Sergei Levendorski, "Preemption games under Lévy uncertainty," *Games and Economic Behavior*, 11 2014, *88*, 354–380. 6

Cabral, Luis M.B., "Standing on the shoulders of dwarfs: dominant firms and innovation incentives," Working paper, 2018. Available at SSRN: https://ssrn.com/abstract=3235598. 6

Cripps, Martin W. and Caroline D. Thomas, "Strategic experimentation in queues," *Theoretical Economics*, 2019, *14* (2), 647–708. 6

Cunningham, Colleen, Florian Ederer, and Song Ma, "Killer acquisitions," *Journal of Political Economy*, 3 2021, *129* (3), 649–702. 5

Das, Kaustav and Nicolas Klein, "Do stronger patents lead to faster innovation? The effect of duplicative search," *International Economic Review*, Forthcoming. Available at https://nicolasklein.com/data/documents/optimal_patent_ier_revision_july_12_resubmission.pdf. 6

D'Aspremont, Claude and Alexis Jacquemin, "Cooperative and noncooperative R&D in duopoly with spillovers," *Americal Economic Review*, 1988, *78* (5), 1133–1137. 5

Denicolò, Vincenzo and Michele Polo, "Duplicative research, mergers and innovation," *Economics Letters*, 2018, *166*, 56–59. 5

_ and _ , "Mergers and innovation sharing," *Economics Letters*, 2021, *202*, 109841. 8

Federico, Giulio, Gregor Langus, and Tommaso Valletti, "A simple model of mergers and innovation," *Economics Letters*, 8 2017, *157*, 136–140. 5, 8

_ , _ , and _ , "Horizontal mergers and product innovation," *International Journal of Industrial Organization*, 7 2018, *59*, 1–23. 5, 8

Guadalupe, Maria, Olga Kuzmina, and Catherine Thomas, "Innovation and foreign ownership," *American Economic Review*, 12 2012, *102* (7), 3594–3627. 5

Haucap, Justus, Alexander Rasch, and Joel Stiebale, "How mergers affect innovation: theory and evidence," *International Journal of Industrial Organization*, 3 2019, *63*, 283–325. 5

Jullien, Bruno and Yassine Lefouili, "Horizontal mergers and innovation," *Journal of Competition Law and Economics*, 9 2018, *14* (3), 364–392. 5, 24

_ and _ , "Mergers and investments in new products," TSE Working Paper No. 949, 2020. Available at http://tse-fr.eu/pub/32923. 5

Kamien, Morton I., Eitan Muller, and Israel Zang, "Research joint ventures and R&D cartels," *Americal Economic Review*, 1992, *82* (5), 1293–1306. 5

Katz, Michael L., "An analysis of cooperative research and development," *RAND Journal of Economics*, 1986, *17* (4), 527–543. 4, 5

_ and Howard A. Shelanski, "Mergers and innovation," *Antitrust Law Journal*, 2007, *74* (1), 1–85. 4

Keller, Godfrey, Sven Rady, and Martin Cripps, "Strategic experimentation with exponential bandits," *Econometrica*, 2005, *73* (1), 39–68. 4, 6, 19, 21

Malueg, David A and Shunichi O Tsutsui, "Dynamic R&D competition with learning," *RAND Journal of Economics*, 1997, *28* (4), 751–772. 6

Marlats, Chantal and Lucie Ménager, "Strategic observation with exponential bandits," *Journal of Economic Theory*, 4 2021, *193*, 105232. 14

Marshall, Guillermo and Álvaro Parra, "Innovation and competition: the role of the product market," *International Journal of Industrial Organization*, 7 2019, *65*, 221–247. 6

Motta, Massimo and Emanuele Tarantino, "The effect of horizontal mergers, when firms compete in prices and investments," *International Journal of Industrial Organization*, 9 2021, *78*, 102774. 5

Mukherjee, Arijit, "Merger and process innovation," *Economics Letters*, 4 2022, *213*, 110366. 5

Ornaghi, Carmine, "Mergers and innovation in big pharma," *International Journal of Industrial Organization*, 1 2009, *27* (1), 70–79. 5

Schumpeter, Joseph A., *Capitalism, socialism and democracy*, Routledge, 1942. 2

Shapiro, Carl, "Competition and innovation: Did Arrow hit the bulls eye?," in Josh Lerner and Scott Stern, eds., *The Rate and Direction of Inventive Activity Revisited*, University of Chicago Press, 2012, pp. 361–404. 2

Thomas, Caroline D., "Strategic experimentation with congestion," *American Economic Journal: Microeconomics*, 2 2021, *13* (1), 1–82. 6

Tirole, Jean, *The theory of industrial organization*, MIT Press, 1988. 2

Appendix A Proofs

A.1 Proof of Proposition 1

Given the research intensity X, the merged entity's instantaneous payoff is

$$\underbrace{\pi\,\mathrm{dt}}_{\substack{\text{flow payoff}\\\text{before innovation}}} - \underbrace{cX\,\mathrm{dt}}_{\text{cost of research}} + \underbrace{X\,\mathrm{dt}}_{\text{prob of innovation}} \underbrace{\times(\lambda+\phi)\pi\int_0^{+\infty} e^{-rs}\,\mathrm{ds}}_{\text{payoff after innovation}} = \left(\pi + \frac{(\lambda+\phi)\pi}{r}X - cX\right)\mathrm{dt}. \tag{A.1}$$

Let V_M be the merged entity's expected payoff from the second stage. Then, the merged entity's expected discounted continuation payoff is

$$\underbrace{(1-X\,\mathrm{dt})}_{\text{prob of no innovation}} \times \underbrace{(1-r\,\mathrm{dt})}_{\text{discount factor}} V_M \overset{\mathrm{dt}\approx 0}{=} (1-X\,\mathrm{dt}-r\,\mathrm{dt})\,V_M. \tag{A.2}$$

The merged entity's overall payoff is the sum of the instantaneous and expected discounted continuation payoffs:

$$V_M = \max_{X\in[0,1]}\left\{\pi\,\mathrm{dt} + \left(\frac{(\lambda+\phi)\pi}{r} - c - V_M\right)X\,\mathrm{dt} + (1-r\,\mathrm{dt})V_M\right\}, \tag{A.3}$$

where the maximization is performed over the research intensity at the current moment. The optimization problem (A.3) implies that continuing research (i.e., $X = 1$) is optimal if and only if

$$\frac{(\lambda+\phi)\pi}{r} - c \geq V_M. \tag{A.4}$$

If $X = 1$ is optimal, then (A.3) implies that

$$0 = \pi + \frac{(\lambda+\phi)\pi}{r} - c - (1+r)V_M. \tag{A.5}$$

If $X = 0$ is optimal, then (A.3) implies that $V_M = \pi/r$. Therefore, condition (5) is necessary and sufficient for $X = 1$ to be optimal, and the merged entity's payoff is defined by (6).

A.2 Proof of Proposition 2

Consider the merged entity at a given moment in time and let $p(t) = p$ be the current belief that the research avenue is good. Given the research intensity X, the merged entity's instantaneous payoff is

$$-\underbrace{cX\,\mathrm{dt}}_{\text{cost of research}} + \underbrace{pX\,\mathrm{dt}}_{\text{prob of innovation}} \times V_M, \tag{A.6}$$

where V_M is the merged entity's payoff from the second stage.

Let $W(p)$ be the merged entity's cumulative payoff, provided that the initial belief is $p(0)=p$. In the absence of innovation, the merged entity updates its belief according to (7). Then, the merged entity's expected discounted continuation payoff is

$$\underbrace{(1-pX\,\mathrm{d}t)}_{\text{prob of no innovation}} \times \underbrace{(1-r\,\mathrm{d}t)}_{\text{discount factor}} W(p+\mathrm{d}p) \overset{(7)}{=} (1-pX\,\mathrm{d}t)(1-r\,\mathrm{d}t)W(p-X(1-p)p\,\mathrm{d}t)$$

$$\overset{\mathrm{d}t\approx 0}{=} (1-pX\,\mathrm{d}t - r\,\mathrm{d}t)\,W(p) - W'(p)X(1-p)p\,\mathrm{d}t\,. \quad \text{(A.7)}$$

The overall payoff is the sum of the instantaneous and expected discounted continuation payoffs, (A.6) and (A.7), respectively:

$$W(p) = \max_{X\in[0,2]} \left\{ (pV_M - c)X\,\mathrm{d}t + (1-pX\,\mathrm{d}t - r\,\mathrm{d}t)\,W(p) - W'(p)X(1-p)p\,\mathrm{d}t \right\}, \quad \text{(A.8)}$$

which yields the Hamilton-Jacobi-Bellman equation

$$0 = \max_{X\in[0,2]} \left\{ \left(pV_M - c - pW(p) - p(1-p)W'(p)\right)X - rW(p) \right\}. \quad \text{(A.9)}$$

The linearity in X of the maximand in (A.9) implies that either $X=0$ or $X=2$ is optimal. Hence, the belief interval $p\in[0,1]$ has two regions: (1) a region where $X=0$ is optimal, (2) a region where $X=2$ is optimal.

Region with no research. If $X=0$, then, by (A.9), $W(p)=0$. According to (A.9), choosing $X=0$ is optimal only if

$$c \geq p\left(V_M - W(p) - (1-p)W'(p)\right). \quad \text{(A.10)}$$

Since $W(p)=0$, condition (A.10) can be simplified as

$$c \geq pV_M. \quad \text{(A.11)}$$

Region with full intensity. For the region of beliefs, in which $X=2$, we can solve the differential equation (A.9):

$$2\left(pV_M - c - pW(p) - p(1-p)W'(p)\right) = rW(p), \quad \text{(A.12)}$$

for $W(p)$ explicitly up to a constant of integration:

$$W(p) = \frac{2\,(pV_M - c)}{2+r} - \frac{4c(1-p)}{r(2+r)} + C_2\left(\frac{1-p}{p}\right)^{r/2}(1-p). \quad \text{(A.13)}$$

According to (A.9), choosing $X=2$ is optimal only if

$$c \leq p\left(V_M - W(p) - (1-p)W'(p)\right). \quad \text{(A.14)}$$

Equation (A.12) allows rewriting condition (A.14) as

$$W(p) \geq 0. \tag{A.15}$$

Optimum. To combine the regions into a unique Markovian optimal strategy, first note that $p = 0$ belongs to the region with no research. Indeed, the limit $p \to 0$ of (A.13) is either $\pm\infty$ or finite and negative, depending on the value of C_2. Negative limit contradicts the optimality condition (A.15). Value $+\infty$ is also non-feasible because the cumulative payoff cannot exceed V_M, the payoff from the discovery net of any research cost.

No-research solution. Suppose that the no-research region covers the whole belief interval, $p \in (0,1)$. Since choosing zero intensity is optimal only if condition (A.11) holds, this is an optimum if and only if

$$c \geq V_M. \tag{A.16}$$

The no-research solution corresponds to case 1 in Proposition 2. Given expression (6) for V_M, condition (A.16) is equivalent to condition (8).

Full-intensity solution. Suppose that there exists $\check{p} \in (0,1)$ such that there is no research for $p \in [0, \check{p})$ and the beliefs just above $\check{p}$ belong to the full-intensity region. Then the value-matching and smooth-pasting properties, $W(\check{p}) = W'(\check{p}) = 0$, together with (A.13), give the expression for $\check{p}$ and the constant of integration:

$$\check{p} = \frac{c}{V_M}, \quad C_2 = \frac{4c}{r(2+r)} \left(\frac{\check{p}}{1-\check{p}} \right)^{r/2}. \tag{A.17}$$

A necessary condition for the full-intensity solution to exist is

$$c < V_M \tag{A.18}$$

because threshold $\check{p} = c/V_M$ must be below 1.

Denote by $\check{\check{p}} \leq 1$ the highest belief of the full-intensity region. Let us show that condition (A.15) holds (as strict inequality) for all $p \in (\check{p}, \check{\check{p}}]$. Differentiating (A.13) twice yields

$$W''(p) = C_2 \frac{r(2+r)}{4(1-p)p^2} \left(\frac{1-p}{p} \right)^{r/2}, \tag{A.19}$$

which is positive because $C_2 > 0$. By the value-matching and smooth-pasting properties, $W(\check{p}) = W'(\check{p}) = 0$. Thus, $W(p) > 0$ for all $p \in (\check{p}, \check{\check{p}}]$.

Since condition (A.15) holds as strict inequality for all $p \in (\check{p}, \check{\check{p}}]$, at belief $\check{\check{p}}$, this region cannot be adjacent to the no-research region where $W(p) = 0$ because $W(p)$ is continuous. Thus, the full intensity region ends at $\check{\check{p}} = 1$.

The full-research solution corresponds to cases 2 and 3 in Proposition 2.

A.3 Proof of Proposition 3

Consider the firm that did not innovate — the potential follower. Given the research intensity x, its instantaneous payoff is

$$-\underbrace{cx\,\mathrm{dt}}_{\text{cost of research}} + \underbrace{x\,\mathrm{dt}}_{\text{prob of innovation}} \times \underbrace{\phi\pi\int_0^{+\infty} e^{-rs}\,\mathrm{ds}}_{\text{payoff from innovation}} = \left(\frac{\phi\pi}{r} - c\right)x\,\mathrm{dt}. \tag{A.20}$$

Let V_F be the potential follower's expected payoff. Then, its expected discounted continuation payoff is

$$\underbrace{(1 - x\,\mathrm{dt})}_{\text{prob of no innovation}} \times \underbrace{(1 - r\,\mathrm{dt})}_{\text{discount factor}} V_F \overset{\mathrm{dt}\approx 0}{=} (1 - x\,\mathrm{dt} - r\,\mathrm{dt})\,V_F. \tag{A.21}$$

The potential follower's overall payoff is the sum of the instantaneous and expected discounted continuation payoffs:

$$V_F = \max_{x\in[0,1]}\left\{\left(\frac{\phi\pi}{r} - c - V_F\right)x\,\mathrm{dt} + (1 - r\,\mathrm{dt})V_F\right\}, \tag{A.22}$$

where the maximization is performed over the research intensity at the current moment. The optimization problem (A.22) implies that continuing research (i.e., $x = 1$) is optimal if and only if

$$\frac{\phi\pi}{r} - c \geq V_F. \tag{A.23}$$

If $x = 1$ is optimal, then (A.22) implies that

$$0 = \frac{\phi\pi}{r} - c - (1 + r)V_F. \tag{A.24}$$

If $x = 0$ is optimal, then (A.22) implies that $V_F = 0$. Therefore, condition (13) is necessary and sufficient for $x = 1$ to be optimal, and the potential follower's payoff is defined by (14).

Consider the firm that innovated on the first stage — the potential leader. Given the competitor's research decision x, the potential leader's payoff is

$$V_L = \int_0^{+\infty}\left(\underbrace{\pi\int_0^t e^{-rs}\,\mathrm{ds}}_{\text{payoff before 2nd innovation}} + \underbrace{\lambda\pi\int_t^{+\infty} e^{-rs}\,\mathrm{ds}}_{\text{payoff after 2nd innovation}}\right) \times \underbrace{e^{-t}\,\mathrm{dt}}_{\text{prob of 2nd innovation at time } t} = \frac{\lambda + r}{1 + r}\frac{\pi}{r} \tag{A.25}$$

if $x = 1$ and

$$V_L = \pi\int_0^{+\infty} e^{-rt}\,\mathrm{dt} = \frac{\pi}{r} \tag{A.26}$$

if $x = 0$.

A.4 Proof of Proposition 4

Consider firm i at a given moment in time. Let $p(t) = p$ be the current belief that the research avenue is good. Let us fix the instantaneous research intensity of the competitor by x_{-i}. Then, by choosing intensity x_i, firm i's instantaneous payoff is

$$- \underbrace{cx_i \, \mathrm{d}t}_{\text{cost of research}} + \underbrace{px_i \, \mathrm{d}t}_{\substack{\text{prob of innovation} \\ \text{from } i}} \times V_L + \underbrace{px_{-i} \, \mathrm{d}t}_{\substack{\text{prob of innovation} \\ \text{from } -i}} \times V_F. \tag{A.27}$$

Expression (A.27) assumes that once an innovation arrives, the game proceeds to the second phase, where the firm that innovated and the firm that did not innovate get V_L and V_F, respectively.

Let $V(p)$ be each firm's cumulative payoff in the game, provided that the initial belief is $p(0) = p$. In the absence of innovation, the firms update their belief according to (7). Thus, firm i's expected discounted continuation payoff is

$$\begin{aligned}
&\underbrace{(1 - p(x_1 + x_2)\,\mathrm{d}t)}_{\text{prob of no innovation}} \times \underbrace{(1 - r\,\mathrm{d}t)}_{\text{discount factor}} V(p + \mathrm{d}p) \\
&\qquad \overset{(7)}{=} (1 - p(x_1 + x_2)\,\mathrm{d}t)(1 - r\,\mathrm{d}t) V\left(p - (x_1 + x_2)(1 - p)p\,\mathrm{d}t\right) \\
&\qquad \overset{\mathrm{d}t \approx 0}{=} (1 - p(x_1 + x_2)\,\mathrm{d}t - r\,\mathrm{d}t) V(p) - V'(p)(x_1 + x_2)(1 - p)p\,\mathrm{d}t.
\end{aligned} \tag{A.28}$$

The overall payoff is the sum of the instantaneous and expected discounted continuation payoffs, (A.27) and (A.28), respectively:

$$\begin{aligned}
V(p) = \max_{x_i \in [0,1]} \Big\{ &(-cx_i + px_i V_L + px_{-i} V_F)\,\mathrm{d}t \\
&+ (1 - p(x_1 + x_2)\,\mathrm{d}t - r\,\mathrm{d}t) V(p) - V'(p)(x_1 + x_2)(1 - p)p\,\mathrm{d}t \Big\},
\end{aligned} \tag{A.29}$$

which yields the Hamilton-Jacobi-Bellman equation

$$0 = \max_{x_i \in [0,1]} \left\{ -cx_i + px_i V_L + px_{-i} V_F - rV(p) - p(x_1 + x_2)\big(V(p) + (1 - p)V'(p)\big) \right\}. \tag{A.30}$$

Maximization in (A.30) gives firm i's best response correspondence:

$$x_i^* \begin{cases} = 0 & \text{if } c > p\left(V_L - V(p) - (1 - p)V'(p)\right), \\ \in [0, 1] & \text{if } c = p\left(V_L - V(p) - (1 - p)V'(p)\right), \\ = 1 & \text{if } c < p\left(V_L - V(p) - (1 - p)V'(p)\right). \end{cases} \tag{A.31}$$

In a symmetric equilibrium, the belief interval $p \in [0,1]$ has three regions: (1) a region where both firms do not undertake research, (2) a region where both firms choose the same intermediate intensity, and (3) a region where both firms choose the maximum intensity of research. The second region corresponds to the middle case in (A.31), in which a firm is indifferent among all research intensities — hence, we refer to it as the indifference region.

Region with no research. If $x_i = x_{-i} = 0$, then, by (A.30), $V(p) = 0$. According to (A.31), choosing $x_i = 0$ is optimal if and only if $c \geq pV_L$.

The indifference region. The condition for the middle case in (A.31) gives the differential equation for $V(p)$, which we can solve explicitly up to a constant of integration:

$$V(p) = V_L - c - c(1-p)\ln\frac{p}{1-p} + C_M(1-p). \tag{A.32}$$

Expressing $V'(p)$ through $V(p)$ from the condition for the middle case in (A.31) yields

$$V'(p) = \frac{p\,(V_L - V(p)) - c}{p(1-p)}. \tag{A.33}$$

Substituting (A.33) into (A.30) and assigning $x_i = x_{-i} = x^*(p)$ give the expression for the equilibrium intensity:

$$x^*(p) = \frac{rV(p)}{c - p(V_L - V_F)}. \tag{A.34}$$

Since the equilibrium intensity $x^*(p)$ cannot exceed 1, (A.34) implies that

$$rV(p) \leq c - p(V_L - V_F). \tag{A.35}$$

Region with full intensity. For the region of beliefs, in which $x_i = x_{-i} = 1$, we can solve the differential equation (A.30):

$$p\,(V_L + V_F) - c - rV(p) = 2p\left(V(p) + (1-p)V'(p)\right), \tag{A.36}$$

for $V(p)$ explicitly up to a constant of integration:

$$V(p) = \frac{p\,(V_L + V_F) - c}{2+r} - \frac{2c(1-p)}{r(2+r)} + C_1\left(\frac{1-p}{p}\right)^{r/2}(1-p). \tag{A.37}$$

According to (A.31), choosing full intensity is optimal if and only if

$$c \leq p\left(V_L - V(p) - (1-p)V'(p)\right). \tag{A.38}$$

Equation (A.36) allows rewriting (A.38) as

$$rV(p) \geq c - p(V_L - V_F). \tag{A.39}$$

Equilibrium. To combine the regions into a symmetric equilibrium, first note that, in any equilibrium, each firm's cumulative payoff in the game must be non-negative (otherwise, the firm can deviate to no research and get a strictly higher payoff of zero) and cannot exceed V_L, the payoff of the sole innovator:

$$0 \leq V(p) \leq V_L. \tag{A.40}$$

Condition (A.40) implies that $p = 0$ belongs to the region with no research. Indeed, the limit $p \to 0$ of (A.32) is $+\infty$ and the limit $p \to 0$ of (A.37) is either $\pm\infty$ or finite and negative, depending on the value of C_1. Thus, at $p = 0$, neither (A.32) nor (A.37) is compatible with restriction (A.40), which leaves only the no-research region with $V(p) = 0$.

The remaining construction of the equilibrium relies on the continuity of $V(p)$ and of $V'(p)$ at the regions' borders (the value matching and smooth-pasting properties).

Depending on what kind of region is adjacent to the no-research region on the right, three cases emerge. As we will see, in each case, there exists at most one equilibrium. Moreover, the conditions for equilibrium existence in different cases are non-overlapping, which proves the uniqueness of equilibrium.

No-research equilibrium. Suppose that the no-research region covers the whole belief interval, $p \in (0,1)$. Since choosing zero intensity is optimal only if $c \geq pV_L$, this is an equilibrium if and only if

$$c \geq V_L. \tag{A.41}$$

The no-research equilibrium corresponds to case 1 in Proposition 4. Given the expression (15) for V_L, condition (A.41) is equivalent to condition (16).

Full-intensity equilibrium. Suppose that there exists $\hat{p} \in (0,1)$ such that there is no research for $p \in [0, \hat{p})$ and the beliefs just above $\hat{p}$ belong to the full-intensity region. Then $V(\hat{p}) = V'(\hat{p}) = 0$, together with (A.37), give the expression for $\hat{p}$ and the constant of integration:

$$\hat{p} = \frac{c}{V_L + V_F}, \quad C_1 = \frac{2c}{r(2+r)} \left(\frac{\hat{p}}{1-\hat{p}} \right)^{r/2}. \tag{A.42}$$

A necessary condition for the full-intensity equilibrium to exist is

$$c < V_L + V_F \tag{A.43}$$

because threshold $\hat{p} = c/(V_L + V_F)$ must be below 1. Moreover, full intensity is optimal for beliefs just above $\hat{p}$ only if condition (A.39) holds at $p = \hat{p}$:

$$0 \overset{V(\hat{p})=0}{=} rV(\hat{p}) \geq c - \hat{p}(V_L - V_F) \overset{\text{(A.42)}}{=} \frac{2cV_F}{V_L + V_F}. \tag{A.44}$$

Hence, another necessary condition for the full-intensity equilibrium to exist is

$$V_F = 0 \tag{A.45}$$

because, by Proposition 3, $V_L > 0$ and $V_F \geq 0$.

Denote by $\hat{\hat{p}} \leq 1$ the highest belief of the full-intensity region. Let us show that condition (A.39) holds (as a strict inequality) for all $p \in (\hat{p}, \hat{\hat{p}}]$. The derivative of $rV(p) - c + p(V_L - V_F)$ with respect to p is positive at $p = \hat{p}$ because $V'(\hat{p}) = 0$. Differentiating (A.37) twice yields

$$V''(p) = C_1 \frac{r(2+r)}{4(1-p)p^2} \left(\frac{1-p}{p}\right)^{r/2}, \tag{A.46}$$

which is positive because $C_1 > 0$. Thus, the derivative of $rV(p) - c + p(V_L - V_F)$ is increasing. Consequently, the derivative of $rV(p)-c+p(V_L-V_F)$ is positive, and, so, $rV(p)-c+p(V_L-V_F)$ is increasing for all $p > \hat{p}$ throughout the full-intensity region. At $p = \hat{p}$, expression $rV(p) - c + p(V_L - V_F)$ is equal to 0 because $V(\hat{p}) = 0$ and because $V_F = 0$ by (A.45). Thus, $rV(p)-c+p(V_L-V_F)$ is positive in the full-intensity region $p \in (\hat{p}, \hat{\hat{p}}]$, which means that condition (A.39) holds as strict inequality for all $p \in (\hat{p}, \hat{\hat{p}}]$.

Since condition (A.39) holds as strict inequality for all $p \in (\hat{p}, \hat{\hat{p}}]$, at belief $\hat{\hat{p}}$, the full-intensity region $p \in (\hat{p}, \hat{\hat{p}}]$ cannot be adjacent to the indifference region where (A.35) holds. Since $V''(p) > 0$ by (A.46) and since $V(\hat{p}) = V'(\hat{p}) = 0$ by the value-matching and smooth-pasting properties, we conclude that $V(p) > 0$ for all $p \in (\hat{p}, \hat{\hat{p}}]$. Hence, at belief $\hat{\hat{p}}$, the full-intensity region cannot be adjacent to the no-research region where $V(p) = 0$. Thus, the full-intensity region ends at $\hat{\hat{p}} = 1$.

The full-research equilibrium corresponds to case 2 and to $c = \phi\pi/r$ in case 3 in Proposition 4. Given the expressions (14) and (15) for V_F and V_L, conditions (A.43) and (A.45) together are equivalent to $\phi\pi/r \leq c < \pi/r$. If $\phi\pi/r < c < \pi/r$, then (13) does not hold and threshold $\hat{p} = c/(V_L + V_F)$ becomes $\hat{p} = cr/\pi$, as in (18). If $c = \phi\pi/r$, then (13) holds and threshold $\hat{p} = c/(V_L + V_F)$ becomes $\hat{p} = cr/\pi \times (1+r)/(\lambda+r) = \underline{p}$, as in (20). Note that if $c = \phi\pi/r$, then threshold $\bar{p}$, defined in (21), is equal to $\underline{p}$, which means that there is no indifference region in case 3; yet, this case is still different from case 2 in that the potential follower undertakes research in the second stage, which lowers equilibrium $\hat{p} = c/V_L$ from cr/π to $cr/\pi \times (1+r)/(\lambda+r)$.

Indifference equilibrium. Suppose there exists $\underline{p} \in (0,1)$ such that there is no research for $p \in [0, \underline{p})$ and the beliefs just above $\underline{p}$ belong to the indifference region. Then $V(\underline{p}) = V'(\underline{p}) = 0$, together with (A.32), give the expression for $\underline{p}$ and the constant of integration:

$$\underline{p} = \frac{c}{V_L}, \quad C_M = c \ln \frac{c}{V_L - c} - V_L. \tag{A.47}$$

Substituting C_M from (A.47) into (A.32) yields

$$V(p) = pV_L - c - c(1-p) \ln \frac{p(V_L - c)}{(1-p)c}. \tag{A.48}$$

A necessary condition for the indifference equilibrium to exist is

$$c < V_L \tag{A.49}$$

because threshold $\underline{p} = c/V_L$ must be below 1. From (A.48) it is clear that $V(p) < pV_L - c$ for all $p > \underline{p} = c/V_L$. Hence, by (A.33), the derivative of $V(p)$ is bounded below by $V_L - c/p$, which is positive for all $p > \underline{p} = c/V_L$. Thus, $V(p)$ is increasing in p throughout the indifference region. Since $V(p)$ is equal to 0 at the lower bound of the indifference region, $V(\underline{p}) = 0$, $V(p)$ is positive for all p in the indifference region.

Since $V(p)$ is positive, the numerator in (A.34) is positive for all p in the indifference region. The denominator in (A.34) is positive for beliefs just above $\underline{p} = c/V_L$ if and only if

$$V_F > 0. \tag{A.50}$$

Given the expression (14) for V_F, condition (A.50) is equivalent to

$$c < \frac{\phi\pi}{r}. \tag{A.51}$$

Since the research intensity $x^*(p)$, defined in (A.34), must be positive throughout the indifference region, condition (A.51) is another necessary condition for the indifference equilibrium to exist.

Note that condition (A.51) implies condition (A.49). Indeed, given the expression (15) for V_L, under restriction (A.51), condition (A.49) becomes

$$c < \frac{\lambda + r}{1 + r}\frac{\pi}{r}, \tag{A.52}$$

which holds if (A.51) holds.

We now argue that $x^*(p)$ is increasing from 0 at $p = \underline{p}$ to 1 at some $p = \bar{p} \in (\underline{p}, 1)$, and so, since the equilibrium research intensity must be less or equal to 1, the highest possible belief in the indifference region is $\bar{p}$. Throughout the indifference region, since $V(p)$ is increasing from $V(\underline{p}) = 0$, the research intensity $x^*(p)$, defined in (A.34), is also increasing from $x^*(\underline{p}) = 0$, as long as the denominator in (A.34) remains positive. If the denominator becomes negative at some belief $\bar{\bar{p}} < 1$, the research intensity $x^*(p)$ becomes unbounded as p approaches $\bar{\bar{p}}$ from below and there is some $\bar{p} \in (\underline{p}, \bar{\bar{p}})$ such that $x^*(\bar{p}) = 1$. If the denominator is still positive at $p = 1$, then $\bar{p} \in (\underline{p}, 1)$ such that $x^*(\bar{p}) = 1$ also exists because $x^*(1) > 1$, which we now show. Substituting $p = 1$ into (A.48) and then into (A.34) yields $x^*(1) = r(V_L - c)/(c - V_L + V_F)$, and so, condition $x^*(1) > 1$ is equivalent to

$$\frac{V_L}{c} - 1 > \frac{V_F}{c(1+r)}. \tag{A.53}$$

Condition (A.53) follows from (A.51), which is a necessary condition for the indifference equilibrium to exist. Indeed, under restriction (A.51), by (14) and (15), (A.53) becomes

$$\frac{\overbrace{\lambda\pi - cr}^{\geq \phi\pi - rc}}{cr(1+r)} + \frac{\overbrace{(\pi/r - c)}^{>0 \text{ if } c<\phi\pi/r} r}{c(1+r)} > \frac{\phi\pi - rc}{cr(1+r)} \overbrace{\frac{1}{1+r}}^{<1}, \tag{A.54}$$

which always holds for $c < \phi\pi/r$. Thus, $x^*(1) > 1$, and so, there always exists $\bar{p} \in (\underline{p}, 1)$ such that $x^*(\bar{p}) = 1$, and $x^*(p)$ is increasing in $p \in (\underline{p}, \bar{p})$ from 0 to 1.

Threshold $\bar{p} \in (\underline{p}, 1)$ is uniquely defined as a solution to $x^*(\bar{p}) = 1$, or equivalently,

$$rV(\bar{p}) = c - \bar{p}(V_L - V_F). \tag{A.55}$$

Substituting $V(\bar{p})$ from (A.48) into equation (A.55) yields

$$\frac{V_L}{c} - \frac{1}{\bar{p}} - \frac{1-\bar{p}}{\bar{p}} \frac{r}{1+r} \ln \frac{\bar{p}(V_L - c)}{(1-\bar{p})c} = \frac{V_F}{c(1+r)}. \tag{A.56}$$

The left-hand side of (A.56) is increasing in $\bar{p}$ for $\bar{p} \in (c/V_L, 1)$, is equal to 0 at $\bar{p} = c/V_L = \underline{p}$, and converges to $V_L/c - 1$ as $\bar{p} \to 1$. Hence, by condition (A.53), equation (A.56) has a unique solution $\bar{p} \in (\underline{p}, 1)$.

Since $V(p)$ is positive for all $p \in (\underline{p}, \bar{p}]$, the indifference region cannot end with the no-research region where $V(p) = 0$; that is, it ends with the full-intensity region. The necessary condition for the full-intensity region, (A.39), precludes a configuration in which the indifference region ends before $\bar{p}$. Hence, the indifference region ends at $\bar{p}$, at which the full-intensity region starts.

The beliefs just below $p = \bar{p}$ belong to the indifference region, and so, for these beliefs, the relevant expression for $V(p)$ is given in (A.48). The beliefs just above $p = \bar{p}$ belong to the region with full intensity, and so, for these beliefs, the relevant expression for $V(p)$ is given in (A.37). The continuity of $V(p)$ and of $V'(p)$ at $p = \bar{p}$ gives the equation for $\bar{p}$, which coincides with (A.56), and the constant of integration:

$$C_1 = \frac{2c}{r}\left(1 + \frac{1+r}{2+r}\frac{\bar{p}}{1-\bar{p}}\left(\frac{V_F}{c(1+r)} - \frac{V_L}{c} + 1\right)\right)\left(\frac{\bar{p}}{1-\bar{p}}\right)^{r/2}. \tag{A.57}$$

Suppose that $V(p)$ is defined by (A.37), with C_1 defined in (A.57), for all $p \in (\bar{p}, 1)$. Let us show that condition (A.39) holds (as a strict inequality) for all $p \in (\bar{p}, 1)$. The second derivative of $V(p)$, expressed in (A.46), has the same sign for all $p \in (\bar{p}, 1)$. Hence, the derivative of $rV(p) - c + p(V_L - V_F)$ is monotone. Differentiating (A.37) once and then substituting (A.57) and $p = \bar{p}$ yields

$$\begin{aligned} V'(\bar{p}) &= -\frac{V_L - V_F}{r} + \frac{c}{\bar{p}}\left(\frac{1}{r} + \frac{1+r}{r}\frac{\bar{p}}{1-\bar{p}}\left(\frac{V_L}{c} - \frac{1}{\bar{p}} - \frac{V_F}{c(1+r)}\right)\right) \\ &\overset{(A.56)}{=} -\frac{V_L - V_F}{r} + \frac{c}{\bar{p}}\left(\frac{1}{r} + \ln\frac{\bar{p}(V_L - c)}{(1-\bar{p})c}\right) \overset{\bar{p}>c/V_L}{>} -\frac{V_L - V_F}{r} + \frac{c}{r\bar{p}} > -\frac{V_L - V_F}{r}, \end{aligned} \tag{A.58}$$

which implies that the derivative of $rV(p)-c+p(V_L-V_F)$ is positive at $p=\bar{p}$. Differentiating (A.37) and then taking the limit $p\to 1$ gives

$$V'(1)=\frac{2c+r(V_L+V_F)}{r(2+r)}>0, \tag{A.59}$$

which implies that the derivative of $rV(p)-c+p(V_L-V_F)$ is positive at the limit $p\to 1$. Hence, the derivative of $rV(p)-c+p(V_L-V_F)$ is positive for all $p\in(\bar{p},1)$, and so, $rV(p)-c+p(V_L-V_F)$ is increasing for all $p\in(\bar{p},1)$. At $p=\bar{p}$, the value of $rV(p)-c+p(V_L-V_F)$ is 0 by (A.55). Thus, $rV(p)-c+p(V_L-V_F)$ is positive for all $p\in(\bar{p},1)$, which means that condition (A.39) holds as a strict inequality for all $p\in(\bar{p},1)$.

Denote by $\hat{\bar{p}}\leq 1$ the highest belief of the full-intensity region. Suppose that $\hat{\bar{p}}<1$. Since condition (A.39) holds as a strict inequality for all $p\in(\bar{p},1)$, it holds as a strict inequality for $p=\hat{\bar{p}}$. Hence, the full-intensity region cannot be adjacent to the indifference region where (A.35) holds. Moreover, for all $p\in(\bar{p},1)$, including $p=\hat{\bar{p}}$, $V(p)$ is positive because

$$rV(p)\overset{(A.39)}{>}c-V_L+V_F\overset{(A.53)}{>}r(V_L-c)\overset{(A.49)}{>}0. \tag{A.60}$$

Hence, at belief $\hat{\bar{p}}$, the full intensity region cannot be adjacent to the no-research region where $V(p)=0$. Thus, the full-intensity region ends at $\hat{\bar{p}}=1$.

The indifference equilibrium corresponds to $c<\phi\pi/r$ in case 3 in Proposition 4. Given the expressions (14) and (15) for V_F and V_L, the expression (A.47) for $\underline{p}$ becomes (20) and the equation (A.56) for $\bar{p}$ becomes (21). The expression (A.34) for the equilibrium intensity becomes (22) after substituting $V(p)$ from (A.48), V_F from (A.56) and $V_L=c/\underline{p}$.

A.5 Proof of Lemma 1

The comparative statics of $\underline{p}$ trivially follows from (20).

Threshold $\bar{p}$ defined in (21) decreases in λ because the right-hand side of (21) does not depend on λ and $\bar{p}$, while the left-hand side of (21) increases in $\bar{p}$:

$$\frac{\partial}{\partial\bar{p}}\left(\frac{1}{\underline{p}}-\frac{1}{\bar{p}}-\frac{1-\bar{p}}{\bar{p}}\frac{r}{1+r}\ln\frac{\bar{p}(1-\underline{p})}{(1-\bar{p})\underline{p}}\right)=\frac{1}{\bar{p}^2(1+r)}\left(1+r\ln\frac{\bar{p}(1-\underline{p})}{(1-\bar{p})\underline{p}}\right)>0, \tag{A.61}$$

and increases in λ: it depends on λ only through $\underline{p}$, which, by (20), decreases in λ, and

$$\frac{\partial}{\partial\underline{p}}\left(\frac{1}{\underline{p}}-\frac{1}{\bar{p}}-\frac{1-\bar{p}}{\bar{p}}\frac{r}{1+r}\ln\frac{\bar{p}(1-\underline{p})}{(1-\bar{p})\underline{p}}\right)=-\frac{\bar{p}(1-\underline{p})+r(\bar{p}-\underline{p})}{\bar{p}(1+r)(1-\underline{p})\underline{p}^2}<0. \tag{A.62}$$

Threshold $\bar{p}$ defined in (21) increases in ϕ because the right-hand side of (21) increases in ϕ and is independent of $\bar{p}$, while the left-hand side of (21) increases in $\bar{p}$ by (A.61) and is independent of ϕ.

The equilibrium research intensity $x^*(p)$ defined in (22) decreases in $\bar{p}$:

$$\frac{\partial}{\partial \bar{p}}\left(\frac{\frac{1}{\underline{p}}-\frac{1}{p}-\frac{1-p}{p}\ln\frac{p(1-\underline{p})}{(1-p)\underline{p}}}{\frac{1}{r}\left(\frac{1}{p}-\frac{1}{\bar{p}}\right)+\frac{1}{\underline{p}}-\frac{1}{\bar{p}}-\frac{1-\bar{p}}{\bar{p}}\ln\frac{\bar{p}(1-\underline{p})}{(1-\bar{p})\underline{p}}}\right)=$$

$$-\frac{\frac{1}{\underline{p}}-\frac{1}{p}-\frac{1-p}{p}\ln\frac{p(1-\underline{p})}{(1-p)\underline{p}}}{\left(\frac{1}{r}\left(\frac{1}{p}-\frac{1}{\bar{p}}\right)+\frac{1}{\underline{p}}-\frac{1}{\bar{p}}-\frac{1-\bar{p}}{\bar{p}}\ln\frac{\bar{p}(1-\underline{p})}{(1-\bar{p})\underline{p}}\right)^2}\frac{1}{\bar{p}^2 r}\left(1+r\ln\frac{\bar{p}(1-\underline{p})}{(1-\bar{p})\underline{p}}\right) \quad \text{(A.63)}$$

is negative. Indeed, $\frac{1}{\underline{p}}-\frac{1}{p}-\frac{1-p}{p}\ln\frac{p(1-\underline{p})}{(1-p)\underline{p}}$ is positive because it is increasing in $p > \underline{p}$ and equal to 0 at $p = \underline{p}$.

The equilibrium research intensity $x^*(p)$ defined in (22) decreases in $\underline{p}$:

$$\frac{\partial}{\partial \underline{p}}\left(\frac{\frac{1}{\underline{p}}-\frac{1}{p}-\frac{1-p}{p}\ln\frac{p(1-\underline{p})}{(1-p)\underline{p}}}{\frac{1}{r}\left(\frac{1}{p}-\frac{1}{\bar{p}}\right)+\frac{1}{\underline{p}}-\frac{1}{\bar{p}}-\frac{1-\bar{p}}{\bar{p}}\ln\frac{\bar{p}(1-\underline{p})}{(1-\bar{p})\underline{p}}}\right)=$$

$$-\frac{\frac{p-\underline{p}}{r}\left(\frac{\bar{p}}{p}-1\right)+\left\{\left(\bar{p}-\underline{p}\right)(1-p)\ln\frac{p(1-\underline{p})}{(1-p)\underline{p}}-\left(p-\underline{p}\right)(1-\bar{p})\ln\frac{\bar{p}(1-\underline{p})}{(1-\bar{p})\underline{p}}\right\}}{p\bar{p}\left(1-\underline{p}\right)\underline{p}^2\left(\frac{1}{r}\left(\frac{1}{p}-\frac{1}{\bar{p}}\right)+\frac{1}{\underline{p}}-\frac{1}{\bar{p}}-\frac{1-\bar{p}}{\bar{p}}\ln\frac{\bar{p}(1-\underline{p})}{(1-\bar{p})\underline{p}}\right)^2} \quad \text{(A.64)}$$

is negative. Indeed, the expression in the curly brackets is positive because it is a concave function of p and equal to 0 at $p = \underline{p}$ and at $p = \bar{p}$.

Since the equilibrium research intensity $x^*(p)$ decreases in $\bar{p}$ and in $\underline{p}$, the comparative statics of $x^*(p)$ follows from the comparative statics of $\bar{p}$ and $\underline{p}$.

A.6 Proof of Lemma 2

The probability of the first innovation depends only on the cumulative investment in research:

$$\Pr(\text{first innovation}) = 1-\exp\left(-\int_0^{+\infty} X(t)\,\mathrm{d}t\right). \quad \text{(A.65)}$$

Given a prior belief p_0 and a stopping threshold p_S, the cumulative investment in research is equal to

$$\int_0^{+\infty} X(t)\,\mathrm{d}t \overset{(7)}{=} \int_{p_S}^{p_0} \frac{\mathrm{d}p}{(1-p)p} = \ln \frac{1/p_S - 1}{1/p_0 - 1}. \tag{A.66}$$

Combining (A.65) and (A.66), we get (23).

Appendix B Microfoundations for the payoff structure

In this section, we provide microfoundations for the parameters π, λ, ϕ.

One innovation

Suppose that only one firm generated an innovation. This innovation allows the firm to produce a product. The inverse demand for the product is

$$p(q) = Q - q, \tag{B.1}$$

where q is the market quantity and $Q > 0$ is a parameter.

Assume that the firm's marginal cost is 0. Then, the firm's profit is $qp(q)$, which is maximized at $q = Q/2$. Hence, the sole innovator's profit is

$$\pi = \frac{Q^2}{4}. \tag{B.2}$$

Two innovations

Suppose that both firms have generated innovations, and, thus, both produce substitute products. Let p_i and q_i be the price and the demand for a product produced by firm i. The inverse demand function for firm i's product is

$$p_i(q_1, q_2) = Q - q_i - \theta q_{-i}, \tag{B.3}$$

where q_{-i} is the demand for the competitor's product. Parameter $\theta \in [0,1]$ is the degree of substitutability between the products. When $\theta = 0$, the products are independent; when $\theta = 1$, the products are perfect substitutes.

Firms' marginal production cost is normalized to 0.

Firms compete by choosing quantities. The firm that innovated first has a natural advantage and so sets its quantity first; that is, the game is a Stackelberg competition with differentiated products. We assume that the firms commit to their quantity choices and then forever receive the flow payoff generated by their choices.

Without loss of generality, suppose that firm 1 is the leader and firm 2 is the follower. The follower maximizes

$$\max_{q_2 \geq 0} q_2 p_2(q_1, q_2) = q_2 \left(Q - q_2 - \theta q_1\right), \tag{B.4}$$

and so, the follower's best response function is

$$q_2(q_1) = \frac{Q - \theta q_1}{2}. \tag{B.5}$$

When choosing its quantity, the leader correctly predicts the follower's best response. Thus, the leader maximizes

$$\max_{q_1 \geq 0} q_1 p_1(q_1, q_2(q_1)) = q_1 \left(\frac{2-\theta}{2} Q - \frac{2-\theta^2}{2} q_1 \right) \tag{B.6}$$

and chooses

$$q_1 = \frac{2-\theta}{2(2-\theta^2)} Q. \tag{B.7}$$

Hence, the leader's payoff is

$$\lambda \pi = \frac{(2-\theta)^2}{2(2-\theta^2)} \frac{Q^2}{4}, \tag{B.8}$$

while the follower's payoff is

$$\phi \pi = \frac{\left(4 - 2\theta - \theta^2\right)^2}{4(2-\theta^2)^2} \frac{Q^2}{4}. \tag{B.9}$$

Interpretation of parameters

Comparing (B.8) and (B.9) with (B.2) yields the expression for λ and ϕ, as functions of θ:

$$\lambda(\theta) = \frac{(2-\theta)^2}{2(2-\theta^2)}, \quad \phi(\theta) = \frac{\left(4 - 2\theta - \theta^2\right)^2}{4(2-\theta^2)^2}. \tag{B.10}$$

When products are independent ($\theta = 0$), both firms get the monopoly profit:

$$\lambda(0) = \phi(0) = 1. \tag{B.11}$$

For all $\theta \in (0,1)$, both $\lambda(\theta)$ and $\phi(\theta)$ are decreasing in θ:

$$\lambda'(\theta) = -\frac{2(2-\theta)(1-\theta)}{(2-\theta^2)^2} < 0, \quad \phi'(\theta) = -\frac{8(1-\theta)^2 + \theta(4 - 2\theta - \theta^3)}{(2-\theta^2)^3} < 0. \tag{B.12}$$

This is intuitive because a higher degree of substitutability intensifies competition. Furthermore, for all $\theta \in (0,1)$, the follower's payoff is lower than the leader's payoff:

$$\lambda(\theta) - \phi(\theta) = \frac{(4-3\theta)\theta^3}{4(2-\theta^2)^2} > 0, \tag{B.13}$$

which is a manifestation of the first-mover advantage, familiar from the classical Stackelberg competition model.

Results (B.11), (B.12) and (B.13) justify restriction (1).

Appendix C On allowing the leader to innovate again

In this appendix, we relax the assumption that each firm can innovate at most once and allow the leader to undertake further R&D after producing the first innovation. The payoff from the second innovation is independent of the identity of the innovator, so that the leader's flow payoff is $(\lambda+\phi)\pi$ if the second innovation also comes from this firm.

To make setups with the competing firms and with the merged entity comparable in this alternative setting, we assume that after the first innovation, the merged entity still has two units of research intensity.

In this appendix, we follow the same structure as in the main text and present the main result first and then go into technical details in the analysis section.

C.1 Main result

Allowing the leader to continue research does not change our main result. In particular, Theorem 1 remains unchanged, except for minor adjustments to comparative statics with respect to ϕ in case 2.

Theorem 1.2*. *If (3) holds, then the merger blocks the second innovation but increases the probability that the first innovation arrives and, moreover, brings it forward in time.*

(a) Suppose that

$$\frac{\pi}{r}\left(\phi-\frac{(1-\lambda)r}{1+r}\right)<c\leq\frac{\phi\pi}{r}. \tag{C.1}$$

Then, the increase in the probability of the first innovation does not depend on ϕ and decreases in λ. The decrease in the expected arrival time of the first innovation (conditional on the first innovation arriving) increases in ϕ and decreases in λ.

(b) Suppose that

$$\frac{(\lambda+\phi-1)\pi}{r}<c\leq\frac{\pi}{r}\left(\phi-\frac{(1-\lambda)r}{1+r}\right). \tag{C.2}$$

Then, the increase in the probability of the first innovation decreases in λ and in ϕ. The decrease in the expected arrival time of the first innovation (conditional on the first innovation arriving) decreases in λ and in ϕ.

The difference between cases (a) and (b) in Theorem 1.2* is due to the leader's behavior after the first innovation. While in case (a), the leader does not attempt to produce the second innovation, in case (b) the leader continues research after producing the first innovation.

When the leader continues research after the first innovation — that is, under the condition (C.2) — the comparative statics with respect to ϕ changes. Intuitively, since the second innovation may come from the leader, a higher ϕ increases the expected payoff from innovating first, thus incentivizing the competing firms to undertake more research in the first stage. Hence, like a higher λ, a higher ϕ (1) lowers the merger-induced increase in the probability of the first innovation, and (2) lowers the merger-induced decrease in the expected arrival time of the first innovation. The latter conclusion is in contrast to the results in our baseline setting, where a higher ϕ aggravates the free-riding problem and, thus, slows down the competing firms' research. Here, a higher ϕ does not induce a strong change in the free-riding incentives because, upon arrival of the first innovation, both firms compete for the second innovation on equal footing — the marginal benefit from the second innovation is the same for the leader and the follower.

In region (C.2), the change in the comparative statics does not alter our policy implication that the benefit of the merger is higher when the first and the second innovations are closer substitutes. As innovations become closer substitutes, both ϕ and λ decrease. According to case (b) in Theorem 1.2*, as ϕ and λ decrease, both the appropriability and the informational effects strengthen. Thus, the benefit of the merger is higher. Noteworthy, unlike our baseline setting, here no ambiguity arises in relation to the informational effect.

C.2 Analysis

C.2.1 Merged entity

Proposition 1 remains the same, except that the expression for the merged entity payoff, (6), changes to

$$V_M = \frac{\pi}{r} + \frac{2}{2+r}\max\left\{\frac{(\lambda+\phi-1)\pi}{r} - c, 0\right\} \tag{C.3}$$

because the merged entity still has two units of research intensity after the first innovation.

Proposition 2 remains the same, except that in case 3, the expression for threshold (12), at which the merged entity aborts research efforts without an innovation, changes to

$$\check{p} = \frac{cr}{\pi}\frac{2+r}{2(\lambda+\phi)+r-2rc/\pi}. \tag{C.4}$$

The expression for $\check{p}$ changes because this threshold is equal to c/V_M and V_M is now defined in (C.3).

C.2.2 Competing firms

To incorporate the possibility that the leader may continue research, Proposition 3 changes more substantially. In particular, case 3 is new; in this case, both the leader and the follower undertake research at full intensity until one of them produces the second innovation.[17]

Proposition 3*.

1. If

$$\frac{\phi\pi}{r} < c, \tag{C.5}$$

[17] We assume that the leader chooses to undertake research whenever indifferent.

then both firms abort research after the first innovation and the second innovation never arrives. The potential follower's and the leader's expected payoffs are

$$V_F = 0, \quad V_L = \frac{\pi}{r}. \tag{C.6}$$

2. *If*

$$\frac{\pi}{r}\left(\phi - \frac{(1-\lambda)r}{1+r}\right) < c \leq \frac{\phi\pi}{r}, \tag{C.7}$$

then the follower undertakes research at full intensity until the second innovation arrives, while the leader aborts research after producing the first innovation. The potential follower's and the leader's expected payoffs are

$$V_F = \frac{1}{1+r}\left(\frac{\phi\pi}{r} - c\right), \quad V_L = \frac{\lambda + r}{1+r}\frac{\pi}{r}. \tag{C.8}$$

3. *If*

$$c \leq \frac{\pi}{r}\left(\phi - \frac{(1-\lambda)r}{1+r}\right), \tag{C.9}$$

then both the leader and the follower undertake research at full intensity until the second innovation arrives. The potential follower's and the leader's expected payoffs are

$$V_F = \frac{1}{2+r}\left(\frac{\phi\pi}{r} - c\right), \quad V_L = \frac{\lambda + r}{1+r}\frac{\pi}{r} + \frac{1}{2+r}\left(\frac{\pi}{r}\left(\phi - \frac{(1-\lambda)r}{1+r}\right) - c\right). \tag{C.10}$$

Proof. For the potential follower's optimal strategy, the proof closely follows the proof of Proposition 3 in Appendix A.3. In particular, (A.24) becomes

$$0 = \frac{\phi\pi}{r} - c - (1 + r + x_L)V_F, \tag{C.11}$$

where x_L is the research intensity of the leader.

For the leader, the optimization problem is

$$V_L = \max_{x_L \in [0,1]} \left\{ \left(\frac{(\lambda+\phi)\pi}{r} - c - V_L \right) x_L \,\mathrm{d}t + \left(1 + \frac{\lambda x}{r}\right) \pi \,\mathrm{d}t + (1 - x\,\mathrm{d}t - r\,\mathrm{d}t)\, V_L \right\}, \tag{C.12}$$

which implies that continuing research ($x_L = 1$) is optimal if and only if

$$\frac{(\lambda+\phi)\pi}{r} - c \geq V_L. \tag{C.13}$$

If $x_L = 1$ is optimal, then (C.12) implies that

$$0 = \frac{(\lambda + \phi + \lambda x + r)\pi}{r} - c - (1 + x + r)V_L. \tag{C.14}$$

If $x_L = 0$ is optimal, then (C.12) implies that

$$0 = \frac{(\lambda x + r)\pi}{r} - (x + r)V_L. \tag{C.15}$$

Substituting V_L from (C.15) into (C.13), we get

$$\frac{(\lambda + \phi)\pi}{r} - c \geq \frac{\lambda x + r}{x + r}\frac{\pi}{r} \quad \Leftrightarrow \quad c \leq \frac{\pi}{r}\left(\phi - \frac{(1-\lambda)r}{x + r}\right). \tag{C.16}$$

Substituting V_L from (C.14) into (C.13), we also get (C.16). Hence, condition (C.16) is necessary and sufficient for $x_L = 1$ to be optimal. From (C.14) and (C.15), the leader's payoff is

$$\begin{aligned} V_L &= \frac{1}{1 + x + r}\left(\frac{(\lambda + \phi + \lambda x + r)\pi}{r} - c\right) \\ &= \frac{\lambda x + r}{x + r}\frac{\pi}{r} + \frac{1}{1 + x + r}\left(\frac{\pi}{r}\left(\phi - \frac{(1-\lambda)r}{x + r}\right) - c\right) \quad \text{if (C.16) holds,} \end{aligned} \tag{C.17}$$

$$V_L = \frac{\lambda x + r}{x + r}\frac{\pi}{r} \quad \text{if (C.16) does not hold.} \tag{C.18}$$

Consider cases 1 and 2, that is, suppose that either (C.5) or (C.7) holds. Under either of these conditions, condition (C.16) does not hold irrespective of the value of x. Hence, the leader does not undertake research, and so, these cases stay the same as in Proposition 3.

Consider case 3, that is, suppose that (C.9) holds. Under condition (C.9), condition (13) also holds. Since condition (13) is necessary and sufficient for $x = 1$ to be optimal, $x = 1$ is optimal in case 3. Thus, condition (C.16) becomes condition (C.9), and so, $x_L = 1$ is optimal. Expressions (C.10) follow from (C.11) and (C.17).

□

In the statement of Proposition 4, there are no changes apart from some adjustments in case 3. In this case, once an innovation arrives, then both the leader and the follower undertake research at full intensity if (C.9) holds; otherwise, only the follower undertakes research in the second stage. Other adjustments in case 3 arise because when the leader undertakes research in the second stage, the expressions for V_F and V_L are different. Expression (20) for $\underline{p}$ becomes

$$\underline{p} = \frac{cr}{\pi}\left(\frac{\lambda + r}{1 + r} + \frac{1}{2 + r}\max\left\{\phi - \frac{(1-\lambda)r}{1 + r} - \frac{cr}{\pi}, 0\right\}\right)^{-1} \tag{C.19}$$

because $\underline{p} = c/V_L$. Equation (21) for $\bar{p}$ becomes

$$\frac{1}{\underline{p}} - \frac{1}{\bar{p}} - \frac{1-\bar{p}}{\bar{p}}\frac{r}{1 + r}\ln\frac{\bar{p}(1-\underline{p})}{(1-\bar{p})\underline{p}} = \frac{\phi\pi - cr}{cr(1 + r)} \times \begin{cases} \frac{1}{2+r}, & \text{if (C.9) holds,} \\ \frac{1}{1+r}, & \text{otherwise.} \end{cases} \tag{C.20}$$

The change in (21) affects only the right-hand side, which is equal to $\frac{V_F}{c(1+r)}$, as follows from

(A.56).

Lemma 1 gives comparative statics of $\underline{p}$, $\bar{p}$ and $x^*(p)$ with respect to λ and ϕ. Since the expressions for $\underline{p}$ and $\bar{p}$ in Proposition 4 are affected only if (C.9) holds, the conclusions of Lemma 1 change under condition (C.9) and remain unchanged otherwise.

Lemma 1*. *Suppose condition (C.9) holds. Threshold $\underline{p}$ defined in (C.19) decreases in λ and in ϕ. Threshold $\bar{p}$ defined in (C.20) also decreases in λ and in ϕ. The equilibrium research intensity $x^*(p)$ defined in (22) increases in λ and in ϕ.*

Proof. The comparative statics of $\underline{p}$ trivially follows from (C.19).

The comparative statics of $\bar{p}$ with respect to λ follows from exactly the same argument as in the proof of Lemma 1 in Appendix A.5.

As compared to Appendix A.5, the comparative statics of $\bar{p}$ with respect to ϕ changes because $\underline{p}$ now depends on ϕ. By the implicit function theorem, since the left-hand side of (C.20) is increasing in $\bar{p}$ by (A.61), the sign of $\bar{p}'(\phi)$ coincides with the sign of

$$\underbrace{\frac{\partial}{\partial \phi}\left(\frac{\phi\pi - cr}{cr(1+r)(2+r)}\right)}_{>0} - \underbrace{\frac{\partial}{\partial \underline{p}}\left(\frac{1}{\underline{p}} - \frac{1}{\bar{p}} - \frac{1-\bar{p}}{\bar{p}}\frac{r}{1+r}\ln\frac{\bar{p}(1-\underline{p})}{(1-\bar{p})\underline{p}}\right)}_{<0 \text{ by (A.62)}} \underbrace{\underline{p}'(\phi)}_{<0}. \tag{C.21}$$

Substituting

$$\underline{p}'(\phi) \overset{(C.19)}{=} -\frac{\pi \underline{p}^2}{cr(2+r)} \tag{C.22}$$

into (C.21), we get

$$\frac{\pi}{cr(1+r)(2+r)} - \frac{\bar{p}(1-\underline{p}) + r(\bar{p}-\underline{p})}{\bar{p}(1+r)(1-\underline{p})\underline{p}^2} \cdot \frac{\pi \underline{p}^2}{cr(2+r)} = -\frac{\pi(\bar{p}-\underline{p})}{c\bar{p}(1+r)(2+r)(1-\underline{p})} < 0. \tag{C.23}$$

Thus, $\bar{p}$ decreases in ϕ.

As we argue in Appendix A.5, the comparative statics of $x^*(p)$ follows from the comparative statics of $\bar{p}$ and $\underline{p}$. □

According to Lemma 1*, while the comparative statics with respect to λ is unaffected by the leader's ability to innovate twice, the comparative statics with respect to ϕ changes the direction. Intuitively, since the leader may also be the producer of the second innovation, higher ϕ increases the payoff from becoming the leader, thereby strengthening the preemption motive in the first stage. At the same time, higher ϕ also increases the follower's payoff and with it, the benefit from free-riding. Thus, now there is a trade-off between a stronger preemption motive which intensifies research in the first stage and a stronger free-riding motive which has the opposite effect. Formally, a stronger preemption motive manifests itself in the second term in (C.21), while the free-riding motive is reflected in the first term in (C.21). As we show in the proof (see (C.23)), the preemption motive takes an upper hand over the free-riding motive. Hence, the comparative statics with respect to ϕ reverses: higher ϕ results in a shorter free-riding belief region and higher intensity of research.

C.2.3 The effect of the merger

Case 1 in Theorem 1 remains unchanged because, under restriction (2), Propositions 2 and 4 do not change.

Case 3 in Theorem 1 also remains unchanged, despite the differences in the expressions for thresholds $\check{p}$ and $\underline{p}$, which are now defined in (C.4) and in (C.19), respectively. The same effects arise because, as in the baseline model, $\underline{p} > \check{p}$.

Case 2 in Theorem 1 has to be split into two regions, (C.1) where condition (C.9) does not hold and (C.2) where condition (C.9) holds. If condition (C.9) does not hold, then the leader does not undertake research after the first innovation, and so, Propositions 2 and 4 do not change. Consequently, the results from the baseline setting remain valid here. If condition (C.9) holds, then threshold $\underline{p}$ is different and now defined in (C.19). Yet, as in case 3, the same effects arise because, as in the baseline model, $\underline{p} > \check{p}$, where $\check{p}$ is defined in (10). The comparative statics with respect to ϕ and λ in case (b) in Theorem 1.2* follows from Lemma 1*.